Hindawi
Discrete Dynamics in Nature and Society
Volume 2018, Article ID 7954263, 11 pages
https://doi.org/10.1155/2018/7954263

Research Article

Distributed Robust Kalman Filtering with Unknown and Noisy Parameters in Sensor Networks

Donghua Chen,[1] Ya Zhang,[2] Cheng-Lin Liu,[3] and Yangyang Chen[2]

[1] *School of Architecture Engineering, Nanjing Institute of Technology, Nanjing 211167, China*
[2] *School of Automation, Southeast University, Nanjing 210096, China*
[3] *Institute of Automation, Jiangnan University, Wuxi 214122, China*

Correspondence should be addressed to Ya Zhang; yazhang@seu.edu.cn

Received 8 August 2018; Revised 9 October 2018; Accepted 18 October 2018; Published 2 December 2018

Academic Editor: Dorota Mozyrska

This paper investigates the distributed filtering for discrete-time-invariant systems in sensor networks where each sensor's measuring system may not be observable, and each sensor can just obtain partial system parameters with unknown coefficients which are modeled by Gaussian white noises. A fully distributed robust Kalman filtering algorithm consisting of two parts is proposed. One is a consensus Kalman filter to estimate the system parameters. It is proved that the mean square estimation errors for the system parameters converge to zero if and only if, for any one system parameter, its accessible node subset is globally reachable. The other is a consensus robust Kalman filter to estimate the system state based on the system matrix estimations and covariances. It is proved that the mean square estimation error of each sensor is upper-bounded by the trace of its covariance. An explicit sufficient stability condition of the algorithm is further provided. A numerical simulation is given to illustrate the results.

1. Introduction

Distributed estimation is a fundamental problem in sensor networks and has attracted broad attention of the researchers. Earlier works on distributed estimation often assume that the system model is precisely known to all sensors. However, in most practical applications, an exact model of the system may not be available to sensors and the robust performance of the distributed filter should be investigated.

Consensus is a simple and feasible strategy for distributed estimation over networks. There have been many consensus estimation algorithms. Olfati-Saber ([1]) proposed consensus Kalman filtering algorithms based on the consensus protocol and standard Kalman filter. Stanković et al. ([2]) gave an algorithm which was composed of decentralized overlapping estimators and a consensus scheme for sensor networks where each sensor had prior knowledge of submatrices of the system matrices. Li et al. ([3]) developed optimal and suboptimal filters and gave sufficient conditions for the stochastic stability of the suboptimal filter. Cattivelli et al. ([4]) studied the networks in which each sensor had observable aggregated measurement matrix and proposed diffusion Kalman filtering algorithms. Yang et al. ([5]) studied sensor networks where each sensor could measure the target and was activated with a certain probability. Ugrinovskii ([6, 7]) proposed a suboptimal H_∞ consensus-based estimation algorithm for continuous-time system with deterministic disturbance. The above-mentioned algorithms are called one time-scale method; i.e., the consensus iteration time interval is equal to the estimation/filtering one. References [8–13] designed consensus Kalman filtering algorithms with two time-scale strategy; i.e., during one estimation/filtering interval, multiple consensus iterations were processed to fuse the prior and novel information. The two time-scale algorithms are fully distributed and their stability just requires the collective detectability of the network. However, in this algorithm vast communication costs are required. In most existing works on the one time-scale algorithm, to design the observer gains or provide stability conditions, global knowledge of the network topology is needed. References [2–7, 14–18] gave stability conditions based on LMIs or the spectral radius of the whole system matrix, which requires global topology information and large computations.

Earlier works on distributed estimation in sensor networks often suppose that an exact model of the target system is known. In most practical problems, it is difficult for all sensors to exactly know the model of the system, and each sensor may just have access to partial model information. The robust performance of the filter with respect to the prior unknown or noisy system parameters is an important issue. Although there have been many works solving the robust state estimation problem ([19–21]), there have been few works designing the consensus-based robust estimation algorithm and investigating the robust performance for sensor networks. Recently, Zhang et al. ([22]) proposed a distributed Kalman/H_∞ filter for sensor networks with partial unknown system parameters. Han et al. ([23]) proposed a distributed H_∞ consensus filtering problem for a class of discrete-time-varying systems with stochastic nonlinearities and gave sufficient stability conditions based on the recursive linear matrix inequalities. However, the linear matrix inequalities of sensors are correlated and their verification requires global knowledge of topology and large computations.

In this paper, we mainly study the distributed filtering for discrete-time-invariant system whose system matrix parameters are previously unknown or not precisely known to the sensors in the network. Each sensor may not be able to obtain the measurement. Even if one sensor can obtain the measurement, its measurement may not be observable. We design a distributed robust Kalman filtering algorithm to estimate the system parameters and system state. To estimate the system parameters, due to noises in the available parameter information, we apply a consensus Kalman filtering algorithm. We prove that the estimation errors of the system parameters converge to zero in mean square sense, if and only if for any one of the system parameters its corresponding node subset in which each node can obtain this parameter information is globally reachable. To estimate the system state, due to parameter estimation errors, we apply a consensus robust Kalman filtering algorithm where the consensus weights are designed based on the covariances of the sensors. We provide a sufficient condition guaranteeing the boundedness of the mean square estimation errors of the sensors. The contributions of this paper include the following: (1) to propose a fully distributed robust Kalman filtering algorithm to deal with the noises in available system parameters; (2) to provide a sufficient stability condition based on some uncorrelated LMIs, whose computation is small and does not require global topology information.

Notation. In this paper, I_p denotes a unit matrix of size p. $\mathrm{tr}(\cdot)$, $\|\cdot\|_*$, and $\|\cdot\|_2$ represent the trace, the trace norm, and the 2-norm of a matrix, respectively. $\|x\|$ denotes the 2-norm of vector x. $\mathrm{diag}\{G_1,\ldots,G_n\}$ denotes a block-diagonal matrix with diagonal blocks $G_1,\ldots,G_n$. Matrix P is positively definite and is written as $P>0$. $E(X)$ denotes the expectation of X. For a finite node set $\mathscr{V}$, $|\mathscr{V}|$ denotes the number of nodes in this set.

2. Problem Formulation

Consider a target system with discrete-time linear dynamics

$$x_0(k+1) = A_0 x_0(k) + v(k), \tag{1}$$

where $x_0 \in R^p$ denotes the state of the system and $v(k) \in R^p$ is a system Gaussian noise with zero mean and covariance matrix $W_0 > 0$. It is supposed that $A_0 \in R^{p\times p}$ is nonsingular.

Suppose target (1) is monitored by a network of n sensors with the following measuring system:

$$y_i(k) = C_i x_0(k) + \omega_i(k), \tag{2}$$

where, for sensor $i \in \{1,2,\ldots,n\}$, $y_i \in R^{q_i}$ is the measurement vector at sensor i, $C_i \in R^{q_i\times p}$ is the measuring matrix, and $\omega_i(k) \in R^{q_i}$ is the Gaussian measurement white noise with zero mean and covariance matrix $W_i > 0$. The measurement noises of different sensors and system noise are uncorrelated. Here (A_0, C_i) may not be observable.

In the network, some sensors may not be able to obtain the measurements (2) due to limited sensing range and energy saving. Let b_i denote whether the measurement is available. If sensor i can obtain its measurement (2), then $b_i = 1$; otherwise $b_i = 0$. Let $\mathscr{B} = \{b_i, 1 \le i \le n\}$ ($b_i \in \{0,1\}$) represent the sensing topology.

In applications, it is a really strong assumption that all sensors in the network precisely know the system parameters of the target. In many cases, if one sensor can just obtain the measurement outputs with respect to partial system state, it is likely to obtain the corresponding partial parameters in system matrix A_0. This paper discusses the case in which the parameters may not be previously known to some sensors and sensors' available system information contains noises. For $(s,l) \in \mathbb{N}$ ($\mathbb{N} = \{(s,l), 1 \le s,l \le p\}$), we apply b_i^{sl} to denote whether the sth rank, l^{th} column element of matrix A_0 is known to sensor i. If the information is unknown to i, $b_i^{sl} = 0$; otherwise, $b_i^{sl} = 1$ and sensor i's corresponding available system information $G_i^{sl}(k)$ is

$$G_i^{sl}(k) = A_0^{sl} + \delta_i^{sl}(k). \tag{3}$$

where $A_0 = [A_0^{sl}]_{p\times p}$ and $\delta_i^{sl}(k)$ is a zero mean Gaussian white noise with variance X_i^{sl}, which is independent of $v(k)$, $\omega_i(k)$, and $\delta_{i_1}^{s_1 l_1}(k)$ for any one different i_1, s_1, or l_1 ($1 \le i_1 \le n, 1 \le s_1, l_1 \le p$).

In this network, even if one sensor can obtain the measurement of the target, it does not completely know the target's system matrix and its measuring system may be also not observable, which makes it infeasible to estimate the state of the target without cooperation. The objective is to construct an estimation algorithm on the basis of the local system parameter information, local measurements (when available), and data received from all adjacent neighbors to estimate the state of the target.

3. Distributed Filtering Algorithm

To explain the principle of graph theory in distributed estimation, preliminaries about graph theory are firstly introduced.

In the cooperative estimation in the sensor network, each sensor is treated as a node and the nodes communicate according to the communication graph. For a communication digraph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathcal{A})$, $\mathcal{V} = \{1, \ldots, n\}$ denotes the node set, $\mathcal{E} = \{(i, j), 1 \leq i, j \leq n\}$ denotes the edge set, and $\mathcal{A}$ denotes the adjacency matrix. If node i can receive information from node j, then there is a corresponding edge in $\mathcal{G}$; i.e., $(i, j) \in \mathcal{E}$. The adjacency elements associated with the edges of the graph are defined as $a_{ij} = 1 \Longleftrightarrow (i, j) \in \mathcal{E}$, $a_{ij} = 0 \Longleftrightarrow (i, j) \notin \mathcal{E}$. Since each node can always get its own information, for all i, $(i, i) \in \mathcal{E}$. Neighbor set of sensor i is denoted as $\mathcal{N}_i = \{j : (i, j) \in \mathcal{E}\}$.

In $\mathcal{G}$, a simple path of length l from i to j is such that there exists a sequence of nodes $i_1, i_2, \ldots, i_{l-1}$ with subsequent edges $(i, i_1), (i_1, i_2), \ldots, (i_{l-1}, j) \in \mathcal{E}$. For node i and a node subset $\mathcal{V}_0$, there exists at least one path from i to the node set $\mathcal{V}_0$ if there exists at least one node $j \in \mathcal{V}_0$ such that there is a path from node i to node j. A node subset V_0 is said to be globally reachable in the communication topology if, for any node *in* in $\mathcal{V}$, there exists at least one path from i to the node set $\mathcal{V}_0$.

Consensus protocol is a simple and effective strategy in cooperation of sensors. In this distributed filtering algorithm, two consensus processes, aimed at estimating the system parameters and system state, respectively, are included. Let us assume that, at time k, each node $i \in \mathcal{V}$ in the sensor network can transmit its system matrix information, including parameter estimation error covariances $\{Y_i^{sl}(k), (s, l) \in \mathbb{N}\}$ and parameter estimations $\{A_i^{sl}(k), (s, l) \in \mathbb{N}\}$, and its system state estimation information, including measurement $y_i(k)$, measurement matrix C_i, measurement noise covariance W_i (when available), state estimation error covariance $P_i(k)$, state covariance $Q_i(k)$, and estimated state $x_i(k)$, to its neighbors. Then the consensus-based distributed robust Kalman filtering algorithm is summarized by Algorithm 1.

For node i, when new data is received, it firstly locally computes the weighted average of its neighbors' estimations and covariances for the parameters in the system matrix:

$$\begin{aligned} \widehat{A}_i(k) &= \sum_{j \in \mathcal{N}_i} m_{ij} A_j(k) \\ \widehat{Y}_i(k) &= \sum_{j \in \mathcal{N}_i} m_{ij} Y_j(k) \end{aligned} \tag{4}$$

where $\widehat{A}_i$ and $\widehat{Y}_i$ denote the fused estimation and covariance for system matrix, respectively, m_{ij} is the weight of edge (i, j) and can be any positive value satisfying $\sum_{j \in \mathcal{N}_i} m_{ij} = 1$, and $\mathcal{N}_i$ denotes the neighbor set of sensor i. And then a local Kalman filtering is constructed to estimate the system parameters.

In the second part, sensor i locally aggregates the measurement vectors, measurement matrices, and measurement noise covariance of its neighbors $i_1, i_2, \ldots, i_{M_i}$, respectively, and obtains $z_i(k) = \left[y_{i_1}^T(k) \;\; y_{i_2}^T(k) \;\; \cdots \;\; y_{i_{M_i}}^T(k)\right]^T$, $\overline{C}_i = \left[C_{i_1}^T \;\; C_{i_2}^T \;\; \cdots \;\; C_{i_{M_i}}^T\right]^T$, $\overline{W}_i = \operatorname{diag}\{W_{i_1}, W_{i_2}, \ldots, W_{i_{M_i}}\}$. Then it computes the weights of in-edges $w_{ij}(k)$ $(j \in \mathcal{N}_i)$ and obtains the weighted average of its neighbors' covariance matrices and estimated states

$$\begin{aligned} \widehat{P}_i(k) &= \sum_{j \in \mathcal{N}_i} w_{ij}(k) P_j(k), \\ \widehat{Q}_i(k) &= \sum_{j \in \mathcal{N}_i} w_{ij}(k) Q_j(k), \\ \widehat{x}_i(k) &= \sum_{j \in \mathcal{N}_i} w_{ij}(k) x_j(k). \end{aligned} \tag{5}$$

In the third part, a local robust Kalman filter is constructed based on the aggregated measurements and weighted average estimate state for each sensor node to estimate the state of the target. For sensor i,

$$x_i(k+1) = \overline{A}_i(k) \widehat{x}_i(k) + F_i(k)\left(z_i(k) - \overline{C}_i \widehat{x}_i(k)\right), \tag{6}$$

$$\begin{aligned} P_i(k+1) &= \overline{A}_i(k) \widehat{P}_i(k) \overline{A}_i^T(k) - \overline{A}_i(k) \widehat{P}_i(k) \\ &\quad \cdot \overline{C}_i^T \left(\overline{C}_i \widehat{P}_i(k) \overline{C}_i^T + \overline{W}_i\right)^{-1} \overline{C}_i \widehat{P}_i(k) \overline{A}_i^T(k) \\ &\quad + \left(\alpha_i I_p + W_0\right) A_i^{-T}(k) R_i A_i^{-1}(k) \left(\alpha_i I_p + W_0\right) + \alpha_i I_p \\ &\quad + W_0, \end{aligned} \tag{7}$$

where $\overline{A}_i(k)$ and $F_i(k)$ are the filter gains to be determined. Assume $x_i(0) = 0$.

The weights in the consensus estimation of the system matrix can be any positive values with row sum being equal to 1, while in the consensus estimation of the system state, due to the dynamical variation of the state, the weights cannot be arbitrary. When the weights are not properly designed, the estimation error of the overall network may be divergent. In [6, 7, 14–18] the weights are constants and concerning conditions are implied in the LMIs-based stability conditions, which is difficult to be computed. Here we apply a fully distributed weight design approach based on the traces of the estimation covariances of the sensors.

At time k, for one node i, if $(A_i(k), \overline{C}_i)$ is observable, then the weight of edge $(i, j) \in \mathcal{E}$ is designed as $w_{ii}(k) = 1$ and $w_{ij}(k) = 0$ for node $j \neq i$, $j \in \mathcal{N}_i$; otherwise

$$w_{ij}(k) = \frac{1/\left\|P_j(k)\right\|_*}{\sum_{s \in \mathcal{N}_i}\left(1/\left\|P_s(k)\right\|_*\right)}, \tag{8}$$

where $\|P_i(k)\|_*$ denotes the trace of covariance matrix $P_i(k)$.

Remark 1. The weighted average consensus-based Kalman filter in Algorithm 1 is a little similar to that proposed in [15] where Matei et al. studied sensor networks with completely known parameters and applied a distributed suboptimal filter with weighted average consensus protocol and standard Kalman filter. For sensor i,

$$\begin{aligned} \widehat{P}_i(k) &= \sum_{j \in \mathcal{N}_i} w_{ij} P_j(k), \\ \widehat{x}_i(k) &= \sum_{j \in \mathcal{N}_i} w_{ij} x_j(k). \end{aligned} \tag{9}$$

1. Initialization: $A_i = A_i(0)$, $P_i(0) = S_1$, $Q_i(0) = S_2$, $x_i = x_i(0)$;
2. while new data is received, do
3. Locally compute the weighted average of estimations and covariances about system parameters:
   $\widehat{A}_i = \sum_{j\in\mathcal{N}_i} m_{ij}A_j$;
   $\widehat{Y}_i = \sum_{j\in\mathcal{N}_i} m_{ij}Y_j$.
4. Compute the Kalman estimate of the system parameters:
   $A_i^{sl} = \widehat{A}_i^{sl} + b_i^{sl}\widehat{Y}_i^{sl}(\widehat{Y}_i^{sl} + X_i^{sl})^{-1}(G_i^{sl} - \widehat{A}_i^{sl})$,
   $A_i = [A_i^{sl}]_{1\le s,l\le p}$,
   $Y_i^{sl} = \widehat{Y}_i^{sl} - b_i^{sl}\widehat{Y}_i^{sl}(\widehat{Y}_i^{sl} + X_i^{sl})^{-1}\widehat{Y}_i^{sl}$, $(s,l) \in \mathbb{N}$.
5. Locally aggregate data and compute the weighted average of estimations and covariances about state estimation:
   $z_i = \left[y_{i_1}^T \quad y_{i_2}^T \quad \cdots \quad y_{i_{M_i}}^T\right]^T$,
   $\overline{C}_i = \left[C_{i_1}^T \quad C_{i_2}^T \quad \cdots \quad C_{i_{M_i}}^T\right]^T$,
   $\overline{W}_i = \mathrm{diag}\{W_{i_1}, W_{i_2}, \cdots, W_{i_{M_i}}\}$,
   $\widehat{P}_i = \sum_{j\in\mathcal{N}_i} w_{ij}P_j$,
   $\widehat{Q}_i = \sum_{j\in\mathcal{N}_i} w_{ij}Q_j$,
   $\widehat{x}_i = \sum_{j\in\mathcal{N}_i} w_{ij}x_j$.
   where $i_s \in \mathcal{N}_i^* = \{j \in \mathcal{V} : a_{ij} = 1 \text{ and } b_j = 1\}$, $M_i = |\mathcal{N}_i^*|$ is the number of nodes in set $\mathcal{N}_i^*$;
   if $(A_i, \overline{C}_i)$ is observable, $w_{ii} = 1$ and for $j \ne i$, $w_{ij} = 0$,
   else
   $$w_{ij} = \frac{1/\|P_j\|_*}{\sum_{s\in\mathcal{N}_i}(1/\|P_s\|_*)}.$$
6. Compute the robust Kalman estimate of the target state:
   if there exists $\alpha_i > 0$ such that
   $\alpha_i I_p - N_i\widehat{Q}_i N_i^T > 0$,
   then $x_i = \overline{A}_i\widehat{x}_i + F_i(z_i - \overline{C}_i\widehat{x}_i)$;
   where,
   $\overline{A}_i = A_i + (\alpha_i I_p + W_0)A_i^{-T}R_i$,
   $F_i = \overline{A}_i\widehat{P}_i\overline{C}_i^T(\overline{C}_i\widehat{P}_i\overline{C}_i^T + \overline{W}_i)^{-1}$,
   $N_i = \max_l \left\{\left(\sum_{s=1}^{p} Y_i^{sl}\right)^{1/2}\right\} I_p$,
   $R_i = \widehat{Q}_i^{-1} - \alpha_i^{-1}N_iN_i^T$,
   P_i and Q_i are updated using the following equation:
   $$\begin{aligned} P_i = {} & \overline{A}_i\widehat{P}_i\overline{A}_i^T \\ & -\overline{A}_i\widehat{P}_i\overline{C}_i^T(\overline{C}_i\widehat{P}_i\overline{C}_i^T + \overline{W}_i)^{-1}\overline{C}_i\widehat{P}_i\overline{A}_i^T \\ & +(\alpha_i I_p + W_0)A_i^{-T}R_iA_i^{-1}(\alpha_i I_p + W_0) \\ & +\alpha_i I_p + W_0, \end{aligned}$$
   $Q_i = A_iR_i^{-1}A_i^T + \alpha_i I_p + W_0$.

Algorithm 1

and

$$\begin{aligned} \widehat{x}_i(k) &= Ax_i(k-1) + F_i(k)\left(y_i(k) - C_iAx_i(k-1)\right), \\ P_i(k) &= A\widehat{P}_i(k-1)A^T + W_0 - A\widehat{P}_i(k-1) \\ &\quad \cdot C_i^T\left(C_i\widehat{P}_i(k-1)C_i^T + W_i\right)^{-1}C_i\widehat{P}_i(k-1)A^T, \\ F_i(k) &= \widehat{P}_i(k-1)C_i^T\left(C_i\widehat{P}_i(k-1)C_i^T + W_i\right)^{-1}. \end{aligned} \tag{10}$$

By Markovian jump linear system, a sufficient stability condition is given based on the feasibility of the following LMIs:

$$\begin{aligned} Q_i > \sum_{j\in\mathcal{N}_i} w_{ij}\Big(&AQ_jA^T + W_0 \\ &- AQ_jC_j^T\left(C_jQ_jC_j^T + W_j\right)^{-1}C_jQ_jA^T\Big). \end{aligned} \tag{11}$$

If the system matrix is not precisely known to sensors and just its estimation can be used, when estimating the system state, standard Kalman filter is not feasible and a robust Kalman filtering algorithm is designed in this paper.

Remark 2. In the network, when one sensor can obtain the measurement outputs of the process, it just obtains partial

parameters of the system matrix with noises. Thus, Algorithm 1 contains two consensus-based filtering processes. One is a consensus Kalman filtering for system parameters. The other is a distributed filtering for system state. Moreover, being different from existing algorithms ([6, 7, 14–18]), the weights in the consensus filtering for system state are adaptively designed. The correction step for filtering is done at sensor levels, and thus the filtering algorithm is completely distributed.

4. Stability Analysis

In this section, we analyze the stability properties of the distributed robust Kalman filtering algorithm in Algorithm 1. In this paper, we assume the network is time-invariant.

Firstly, we analyze the estimation of the system parameters. For sensor $i \in \mathscr{V}$ and $(s,l) \in \mathbb{N}$, from Algorithm 1 we have that

$$\widehat{A}_i^{sl}(k+1) = \sum_{j\in\mathscr{N}_i} m_{ij} A_j^{sl}(k), \tag{12}$$

$$\widehat{Y}_i^{sl}(k+1) = \sum_{j\in\mathscr{N}_i} m_{ij} Y_j^{sl}(k), \tag{13}$$

with

$$\begin{aligned} A_i^{sl}(k) &= \widehat{A}_i^{sl}(k) \\ &\quad + b_i^{sl}\widehat{Y}_i^{sl}(k)\left(\widehat{Y}_i^{sl}(k)+X_i^{sl}\right)^{-1}\left(G_i^{sl}(k)-\widehat{A}_i^{sl}(k)\right), \end{aligned} \tag{14}$$

$$Y_i^{sl}(k) = \widehat{Y}_i^{sl}(k) - b_i^{sl}\widehat{Y}_i^{sl}(k)\left(\widehat{Y}_i^{sl}(k)+X_i^{sl}\right)^{-1}\widehat{Y}_i^{sl}(k), \tag{15}$$

where $b_i^{sl} \in \{0,1\}$ $(1 \le s,l \le p)$. If sensor i has an access to the parameter A_0^{sl} with noise, $b_i^{sl} = 1$; otherwise $b_i^{sl} = 0$.

Define $\Delta A_i^{sl}(k) = A_i^{sl}(k) - A_0^{sl}$, $\Delta\widehat{A}_i^{sl}(k) = \widehat{A}_i^{sl}(k) - A_0^{sl}$ as sensor i's estimation error and fused estimation error of system parameter A_0^{sl}, respectively. From (12) and (14) we have that

$$\begin{aligned} \Delta\widehat{A}_i^{sl}(k+1) &= \sum_{j\in\mathscr{N}_i} m_{ij}(\Delta\widehat{A}_j^{sl}(k) \\ &\quad - b_j^{sl}\widehat{Y}_j^{sl}(k)\left(\widehat{Y}_j^{sl}(k)+X_j^{sl}\right)^{-1}\left(\Delta\widehat{A}_j^{sl}(k)-\delta_j^{sl}(k)\right), \end{aligned} \tag{16}$$

and, obviously, $E((\Delta\widehat{A}_i^{sl}(k))^2) \le \widehat{Y}_i^{sl}(k)$ holds for all $1 \le i \le n, 1 \le s,l \le p$, when $E((\Delta\widehat{A}_i^{sl}(0))^2) = \widehat{Y}_i^{sl}(0)$.

Proposition 3. *For any $(s,l) \in \mathbb{N}$ and any node $i \in \mathscr{V}$, $\lim_{k\longrightarrow\infty}\widehat{Y}_i^{sl}(k) = 0$ holds if and only if in the communication topology there exists at least one path from i to the node set $\mathscr{V}^{sl}$, where $\mathscr{V}^{sl} = \{j : b_j^{sl} = 1, j \in \mathscr{V}\}$.*

Proof.

(Sufficiency). If in the graph there exists one node i_0, $b_{i_0}^{sl} = 1$, and node i_0 is globally reachable from any other node, then we partition the graph into strongly connected subgraphs $(\mathscr{V}_i, \mathscr{E}_i, \mathscr{A}_i)$, $1 \le i \le m$. Renumber the nodes such that $M = [M_{ij}]_{m\times m}$ is an upper block-diagonal matrix and the node subset $\mathscr{V}_1$ corresponding to M_{11} contains node i_0. Define $\xi_i^{sl}(k)$ as a vector stacking the values $\widehat{Y}_j^{sl}(k)$, $j \in \mathscr{V}_i$, $\xi^{sl}(k) = [(\xi_1^{sl}(k))^T, \ldots, (\xi_m^{sl}(k))^T]^T$, and define $\eta^{sl}(k)$ as a vector stacking the values $\widehat{Y}_j^{sl}(k)(\widehat{Y}_j^{sl}(k)+X_j^{sl})^{-1}\widehat{Y}_j^{sl}(k), 1 \le j \le n$, $B^{sl} = \mathrm{diag}\{b_1^{sl}, \ldots, b_n^{sl}\}$. Then,

$$\xi^{sl}(k+1) = M\xi^{sl}(k) - MB^{sl}\eta^{sl}(k). \tag{17}$$

For irreducible stochastic matrix M_{11}, there exists a vector $\beta_1 = [\beta_{11}, \ldots, \beta_{1n_1}]^T$ with positive elements satisfying that $\beta_1^T M_{11} = \beta_1^T$. Thus, $\sum_{i\in\mathscr{V}_1}\beta_{1i}\widehat{Y}_i^{sl}(k+1) = \sum_{i\in\mathscr{V}_1}\beta_{1i}\widehat{Y}_i^{sl}(k) - \sum_{i\in\mathscr{V}_1}\beta_{1i}\sum_{j\in\mathscr{N}_i} b_j^{sl}\widehat{Y}_j^{sl}(k)(\widehat{Y}_j^{sl}(k)+X_j^{sl})^{-1}\widehat{Y}_j^{sl}(k)$. Since the subgraph is strongly connected, if for one node $i \in \mathscr{V}_1$, $\widehat{Y}_i^{sl}(k) = 0$, then from (13), we have that, for any other node $j \in \mathscr{V}_1$, $\widehat{Y}_j^{sl}(k) = 0$ holds. When $\widehat{Y}_{i_0}^{sl}(k) > 0$, since $i_0 \in \mathscr{V}_1$, we have that $\sum_{i\in\mathscr{V}_1}\beta_{1i}\sum_{j\in\mathscr{N}_i} b_j^{sl} > 0$ and then $\sum_{i\in\mathscr{V}_1}\beta_{1i}\widehat{Y}_i^{sl}(k+1) < \sum_{i\in\mathscr{V}_1}\beta_{1i}\widehat{Y}_i^{sl}(k)$, which infers that $\lim_{k\longrightarrow\infty}\sum_{i\in\mathscr{V}_1}\beta_{1i}\widehat{Y}_i^{sl}(k) = 0$, and thus for $i \in \mathscr{V}_1$, $\lim_{k\longrightarrow\infty}\widehat{Y}_i^{sl}(k) = 0$.

For matrix M_{22}, there exists a diagonal matrix D_{22}, in which the diagonal elements are all larger than 1 and there exists at least one diagonal element being strictly larger than 1, such that $D_{22}M_{22}$ is an irreducible stochastic matrix. Define $\beta_2 = [\beta_{21}, \ldots, \beta_{2n_2}]^T$ with positive elements satisfying that $\beta_2^T D_{22} M_{22} = \beta_2^T$, and define $\beta_2^T M_{22} = [\widehat{\beta}_{21}, \ldots, \widehat{\beta}_{2n_2}]^T$; then for $1 \le j \le n_2$, $\beta_{2j} \ge \widehat{\beta}_{2j}$, and there exists at least one $j_0, 1 \le j_0 \le n_2$; $\beta_{2j_0} > \widehat{\beta}_{2j_0}$ holds. Thus, if for all $i \in \mathscr{V}_2$, $\widehat{Y}_i^{sl}(k) > 0$, $\sum_{i\in\mathscr{V}_2}\beta_{2i}\widehat{Y}_i^{sl}(k+1) < \sum_{i\in\mathscr{V}_2}\beta_{2i}\widehat{Y}_i^{sl}(k) + \sum_{i\in\mathscr{V}_2}\beta_{2i}\sum_{j\in\mathscr{N}_i\cap\mathscr{V}_1}\widehat{Y}_j^{sl}(k) - \sum_{i\in\mathscr{V}_2}\beta_{2i}\sum_{j\in\mathscr{N}_i} b_j^{sl}\widehat{Y}_j^{sl}(k)(\widehat{Y}_j^{sl}(k)+X_j^{sl})^{-1}\widehat{Y}_j^{sl}(k)$. From the above analysis, for $i \in \mathscr{V}_1$, $\lim_{k\longrightarrow\infty}\widehat{Y}_i^{sl}(k) = 0$; thus, for $i \in \mathscr{V}_2$, $\lim_{k\longrightarrow\infty}\widehat{Y}_i^{sl}(k) = 0$ also holds.

Similarly, for $i \in \mathscr{V}_s$ $(s > 1)$, $\lim_{k\longrightarrow\infty}\widehat{Y}_i^{sl}(k) = 0$. Therefore, if in the graph there exists one node i_0, $b_{i_0}^{sl} = 1$, and node i_0 is globally reachable from any other node, then $\lim_{k\longrightarrow\infty}\widehat{Y}_i^{sl}(k) = 0$ holds for all i.

If there is no globally reachable node i_0 with $b_{i_0}^{sl} = 1$, then the graph can be partitioned into subgraphs; in each subgraph there exists one node denoted as i_0, $b_{i_0}^{sl} = 1$, and node i_0 is globally reachable from any other node in the subgraph. From the above analysis, it is not difficult to prove that $\lim_{k\longrightarrow\infty}\widehat{Y}_i^{sl}(k) = 0$ holds for all i. The sufficiency part has been proved.

(Necessity). If there exists a node having no path from itself to the node set $\mathscr{V}^{sl}$, then there exists a node subset in which each node has no neighbor in other node subsets and is not be able

to obtain the information of A_0^{sl}. Obviously, the prediction error covariances of the nodes in this subset cannot converge to 0. The necessity part has been proved. □

In the following, we analyze the estimation for the system state. To begin with, we give a lemma about robust Kalman filtering for a system with unknown random uncertainty modeled by white noise.

Lemma 4. *Consider a target with discrete-time system:*

$$\zeta(k+1) = A\zeta(k) + v(k), \tag{18}$$

$$y(k) = C\zeta(k) + \omega(k), \tag{19}$$

where $x(k) \in \mathbb{R}^p$ is the state; $y(k) \in \mathbb{R}^q$ is the measured output with observable measurement matrix C; $v(k)$ and $\omega(k)$ are uncorrelated zero mean Gaussian white noises with variance matrix W_0 and W, respectively. The observer cannot obtain the precise system matrix and just obtains the value with noise; i.e., $A(k) = A + \Delta A(k)$, where $\Delta A(k) = [\Delta A^{sl}]_{p\times p}$ represents the unknown parameter noise and satisfies $E(\Delta A(k)) = 0$, $E(\Delta A^{s_1l_1}(k)\Delta A^{s_2l_2}(k)) = 0$ for any $s_1 \neq s_2$ or $l_1 \neq l_2$, $E((\Delta A^{sl}(k))^2) \leq Y^{sl}(k)$. If there exist $\alpha(k) > 0$ such that

$$\alpha(k) I_p - N(k) Q_2(k) N^T(k) > 0, \tag{20}$$

where $Q_1(k)$, $Q_2(k)$ satisfy the following two Riccati equations

$$\begin{aligned} Q_1(k+1) &= \overline{A}(k) Q_1(k) \overline{A}^T(k) - \overline{A}(k) Q_1(k) \\ &\cdot C^T \left(CQ_1(k) C^T + W\right)^{-1} CQ_1(k) \overline{A}^T(k) \\ &+ \left(\alpha(k) I_p + W_0\right) A^{-T}(k) R(k) A^{-1}(k) \\ &\cdot \left(\alpha(k) I_p + W_0\right) + \alpha(k) I_p + W_0 \end{aligned} \tag{21}$$

$$Q_2(k+1) = A(k) R^{-1}(k) A^T(k) + \alpha(k) I_p + W_0 \tag{22}$$

and

$$\begin{aligned} N(k) &= \max_l \left\{ \left(\sum_{s=1}^{p} Y^{sl}(k) \right)^{1/2} \right\} I_p, \\ R(k) &= Q_2^{-1}(k) - \alpha^{-1}(k) N^T(k) N(k), \\ \overline{A}(k) &= A(k) + \left(\alpha(k) I_p + W_0\right) A^{-T}(k) R(k), \end{aligned} \tag{23}$$

then, the filter

$$\widehat{\zeta}(k+1) = \overline{A}(k) \widehat{\zeta}(k) + F(k) \left(y(k) - C\widehat{\zeta}(k)\right), \tag{24}$$

with $F(k) = \overline{A}(k)Q_1(k)C^T(CQ_1(k)C^T + W)^{-1}$ satisfies

$$E\left(\left(\zeta(k) - \widehat{\zeta}(k)\right)\left(\zeta(k) - \widehat{\zeta}(k)\right)^T\right) \leq \operatorname{tr}\left(Q_1(k)\right) \tag{25}$$

if $E((\zeta(0) - \widehat{\zeta}(0))^T(\zeta(0) - \widehat{\zeta}(0))) < \begin{bmatrix} Q_1(0) & 0 \\ 0 & Q_2(0) \end{bmatrix}$.

Proof. Define $e(k) = \zeta(k) - \widehat{\zeta}(k)$, $e_z(k) = z(k) - \widehat{z}(k)$; the augmented system of (19) and (24) can be expressed by

$$\zeta_e(k+1) = \left(A_e(k) + \Delta A_e(k)\right)\zeta_e(k) + B_e(k)\overline{w}(k), \tag{26}$$

where

$$\begin{aligned} \zeta_e(k) &= \left[e^T(k), \zeta^T(k)\right]^T \\ \overline{w}(k) &= \left[v^T(k), \omega^T(k)\right]^T, \\ A_e(k) &= \begin{bmatrix} \overline{A}(k) - F(k)C & A(k) - \overline{A}(k) \\ 0 & A(k) \end{bmatrix}, \\ \Delta A_e(k) &= \begin{bmatrix} 0 & \Delta A(k) \\ 0 & \Delta A(k) \end{bmatrix}, \\ B_e(k) &= \begin{bmatrix} I_p & -F(k) \\ I_p & 0 \end{bmatrix}. \end{aligned} \tag{27}$$

From the properties of the parameter noises, we have $E(\Delta A^T(k)\Delta A(k)) = \operatorname{diag}\{\sum_{s=1}^{p} Y^{s1}(k), \ldots, \sum_{s=1}^{p} Y^{sp}(k)\}$. Define $N(k) = \max_l\{(\sum_{s=1}^{p} Y^{sl}(k))^{1/2}\}I_p$; it is obvious that $E(\Delta A^T(k)\Delta A(k)) \leq N^T(k)N(k)$.

Define $\Sigma(k) = E(\zeta_e(k)\zeta_e^T(k))$, then from (26),

$$\begin{aligned} &\Sigma(k+1) \\ &= E\left(\left(A_e(k) + \Delta A_e(k)\right)\Sigma(k)\left(A_e(k) + \Delta A_e(k)\right)^T\right) \\ &\quad + B_e(k)\widetilde{W}(k)B_e^T(k) \end{aligned} \tag{28}$$

where

$$\begin{aligned} \Delta A_e(k) &= H\left(\Delta A(k) N^{-1}(k)\right)\overline{N}(k), \\ H &= \begin{bmatrix} I_p \\ I_p \end{bmatrix}, \\ \overline{N}(k) &= \left[0 \;\; N(k)\right], \\ \widetilde{W}(k) &= \begin{bmatrix} W_0 & 0 \\ 0 & W \end{bmatrix}, \end{aligned} \tag{29}$$

and satisfy $E((\Delta A(k)N^{-1}(k))^T(\Delta A(k)N^{-1}(k))) \leq I_p$.

From [19], if there exists $\alpha(k)$ such that $\alpha(k)I_p - \overline{N}(k)\Sigma(k)\overline{N}^T(k) > 0$, then

$$\begin{aligned} &\left(A_e(k) + \Delta A_e(k)\right)\Sigma(k)\left(A_e(k) + \Delta A_e(k)\right)^T \\ &\quad \leq A_e(k)\left(\Sigma^{-1}(k) - \alpha^{-1}(k)\overline{N}^T(k)\overline{N}(k)\right)^{-1} A_e^T(k) \\ &\quad + \alpha(k) HH^T. \end{aligned} \tag{30}$$

Therefore,

$$\begin{aligned} &\Sigma(k+1) \\ &\quad \leq A_e(k)\left(\Sigma^{-1}(k) - \alpha^{-1}(k)\overline{N}^T(k)\overline{N}(k)\right)^{-1} A_e^T(k) \\ &\quad + \alpha(k) HH^T + B_e(k)\widetilde{W}(k)B_e^T(k). \end{aligned} \tag{31}$$

Define an equation

$$\begin{aligned}\overline{\Sigma}(k+1) &= A_e(k)\left(\overline{\Sigma}^{-1}(k)-\alpha^{-1}(k)\overline{N}^T(k)\overline{N}(k)\right)^{-1}A_e^T(k) \\ &\quad + 2\alpha(k)HH^T + B_e(k)\widetilde{W}(k)B_e^T(k)\end{aligned} \tag{32}$$

with $\overline{\Sigma}(0) = \Sigma(0)$. Obviously, $\overline{\Sigma}(k) \geq \Sigma(k)$ holds for all k. Thus, $\alpha(k)I_p - \overline{N}(k)\overline{\Sigma}(k)\overline{N}^T(k) > 0$ implies $\alpha(k)I_p - \overline{N}(k)\Sigma(k)\overline{N}^T(k) > 0$.

If $\overline{\Sigma}(k)$ is assumed to have the form $\begin{bmatrix} Q_1(k) & 0 \\ 0 & Q_2(k)\end{bmatrix}$, from (32) we have $\overline{\Sigma}(k+1) = \begin{bmatrix} \overline{\Sigma}_{11}(k+1) & \overline{\Sigma}_{12}(k+1) \\ \overline{\Sigma}_{21}(k+1) & \overline{\Sigma}_{22}(k+1)\end{bmatrix}$, where

$$\begin{aligned}\overline{\Sigma}_{11}(k+1) &= \left(\overline{A}(k)-F(k)C\right)Q_1(k)\left(\overline{A}(k)-F(k)C\right)^T \\ &\quad + \left(A(k)-\overline{A}(k)\right)R^{-1}(k)\left(A(k)-\overline{A}(k)\right)^T \\ &\quad + \alpha(k)I_p + W_0 + F(k)WF^T(k), \\ \overline{\Sigma}_{12}(k+1) &= \left(A(k)-\overline{A}(k)\right)R^{-1}(k)A^T(k) + W_0 + \alpha(k)I_p, \\ \overline{\Sigma}_{21}(k+1) &= \overline{\Sigma}_{12}^T(k+1), \\ \overline{\Sigma}_{22}(k+1) &= A(k)R^{-1}(k)A^T(k) + \alpha(k)I_p + W_0, \\ R(k) &= Q_2^{-1}(k) - \alpha^{-1}(k)N^T(k)N(k).\end{aligned} \tag{33}$$

When $\overline{A}(k) = A(k) + (\alpha(k)I_p + W_0)A^{-T}(k)R(k)$, there holds $\overline{\Sigma}_{12}(k+1) = 0$, and thus $\overline{\Sigma}(k)$ is a block-diagonal positive definite matrix with diagonal blocks.

$$\begin{aligned}Q_1(k+1) &= \left(\overline{A}(k)-F(k)C\right)Q_1(k)\left(\overline{A}(k)-F(k)C\right)^T \\ &\quad + \left(A(k)-\overline{A}(k)\right)R^{-1}(k)\left(A(k)-\overline{A}(k)\right)^T \\ &\quad + \alpha(k)I_p + W_0 + F(k)WF^T(k)\end{aligned} \tag{34}$$

$$Q_2(k+1) = A(k)R^{-1}(k)A^T(k) + \alpha(k)I_p + W_0 \tag{35}$$

In order to find the optimal filter gains $F(k)$ minimizing $Q_1(k)$, we take the first variation to (34) with respect to $F(k)$ and equal it to zero.

$$\begin{aligned}\frac{\partial Q_1(k+1)}{\partial F(k)} &= \left(\overline{A}(k)-F(k)C\right)Q_1(k)(-C)^T + FW \\ &= 0\end{aligned} \tag{36}$$

From (36) we obtain the optimal gains $F(k) = \overline{A}(k)Q_1(k)C^T(CQ_1(k)C^T + W)^{-1}$. This completes the proof. □

In the following we analyze the stability of the distributed robust Kalman filtering algorithm.

For sensor i, $i \in \mathscr{V}$, it can only use its own estimated system matrix $A_i(k)$, and then the target dynamics can be depicted by

$$x_0(k+1) = \left(A_i(k) - \Delta A_i(k)\right)x_0(k) + \omega_0(k). \tag{37}$$

Here $\Delta A_i(k) = A_i(k) - \mathrm{A}$ and satisfies $E(\Delta A_i^{s_1 l_1}(k)\Delta A_i^{s_2 l_2}(k)) = 0$ for any $s_1 \neq s_2$ or $l_1 \neq l_2$, $E((\Delta A_i^{sl}(k))^2) \leq Y_i^{sl}(k)$.

From Algorithm 1 we have that

$$\widehat{x}_i(k) = \sum_{j\in\mathscr{N}_i} w_{ij}(k)x_j(k), \tag{38}$$

$$\begin{aligned}x_j(k+1) &= \overline{A}_j(k)\widehat{x}_j(k) \\ &\quad + F_j(k)\left(z_j(k) - \overline{C}_j\widehat{x}_j(k)\right),\end{aligned} \tag{39}$$

$$\widehat{P}_i(k) = \sum_{j\in\mathscr{N}_i} w_{ij}(k)P_j(k), \tag{40}$$

$$\widehat{Q}_i(k) = \sum_{j\in\mathscr{N}_i} w_{ij}(k)Q_j(k), \tag{41}$$

where $w_{ij}(k) \geq 0$ with $\sum_{j\in\mathscr{N}_i} w_{ij}(k) = 1$ being the consensus weight of edge (i, j) in the target of state estimation. $F_i(k)$ and $\overline{A}_i(k)$ are filter gains given by Algorithm 1; $P_i(k)$ and $Q_i(k)$ satisfy the following equations:

$$\begin{aligned}P_i(k+1) &= \left(\overline{A}_i(k)-F_i(k)\overline{C}_i\right)\widehat{P}_i(k)\left(\overline{A}_i(k)-F_i(k)\overline{C}_i\right)^T \\ &\quad + \left(A_i(k)-\overline{A}_i(k)\right)R_i^{-1}(k)\left(A_i(k)-\overline{A}_i(k)\right)^T \\ &\quad + \alpha_i(k)I_p + W_0 + F_i(k)\overline{W}_iF_i^T(k)\end{aligned} \tag{42}$$

$$Q_i(k+1) = A_i(k)R_i^{-1}(k)A_i^T(k) + \alpha_i(k)I_p + W_0. \tag{43}$$

Define $e_i(k) = x_i(k) - x_0(k)$, $\widehat{e}_i(k) = \widehat{x}_i(k) - x_0(k)$ as sensor i's estimation error and collective estimation error, respectively. Then, from (38) we have that

$$\widehat{e}_i(k+1) = \sum_{j\in\mathscr{N}_i} w_{ij}(k)e_j(k+1), \tag{44}$$

$$\begin{aligned}e_j(k+1) &= \left(\overline{A}_j(k)-F_j(k)\overline{C}_j\right)\widehat{e}_j(k) \\ &\quad + \left(\overline{A}_j(k)-A_j(k)\right)x_0(k) \\ &\quad + \Delta A_j(k)x_0(k) + F_j(k)\overline{\omega}_j(k) \\ &\quad - v(k).\end{aligned} \tag{45}$$

Let $\xi_i(k) = [e_i^T(k), x_0^T(k)]^T$ and $\widehat{\xi}_i(k) = [\widehat{e}_i^T(k), x_0^T(k)]^T$. Then $E(\widehat{\xi}_i(k)\widehat{\xi}_i^T(k)) \leq \sum_{j\in\mathscr{N}_i} w_{ij}(k)E(\xi_j(k)\xi_j^T(k))$. From the proof of Lemma 4, if there exists $\alpha_i(k) > 0$ such that $\alpha_i(k)I_p - \overline{N}_i(k)\widehat{\Sigma}_i(k)\overline{N}_i^T(k) > 0$, then

$$E\left(\xi_i(k+1)\xi_i^T(k+1)\right) \leq \Sigma_i(k+1) \tag{46}$$

where

$$\begin{aligned}\Sigma_i(k+1) &= \Pi_i(k)\left(\widehat{\Sigma}_i^{-1}(k) - \alpha_i^{-1}(k)\overline{N}_i^T(k)\overline{N}_i(k)\right)^{-1}\Pi_i^T(k) \\ &\quad + \alpha_i(k) I_{2p} + B_i(k)\widetilde{W}_i(k) B_i^T(k)\end{aligned} \tag{47}$$

$$\widehat{\Sigma}_i(k) = \sum_{j\in\mathcal{N}_i} w_{ij}(k)\Sigma_j(k), \tag{48}$$

$$\begin{aligned}\Pi_i(k) &= \begin{bmatrix}\overline{A}_i(k) - F_i(k)\overline{C}_i & A_i(k) - \overline{A}_i(k) \\ 0 & A_i(k)\end{bmatrix}, \\ B_i(k) &= \begin{bmatrix} I_p & -F_i(k) \\ I_p & 0\end{bmatrix}, \\ \widetilde{W}_i(k) &= \begin{bmatrix} W_0 & 0 \\ 0 & \overline{W}_i\end{bmatrix}, \\ \overline{N}_i(k) &= \begin{bmatrix} 0 & N_i(k)\end{bmatrix}.\end{aligned} \tag{49}$$

Furthermore, if $E(\xi_i(0)\xi_i^T(0)) \le \left[\begin{smallmatrix} P_i(0) & 0 \\ 0 & Q_i(0)\end{smallmatrix}\right]$, then $\Sigma_i(k)$ and $\widehat{\Sigma}_i(k)$ are both block-diagonal matrices with diagonal blocks $P_i(k)$, $Q_i(k)$ and $\widehat{P}_i(k)$, $\widehat{Q}_i(k)$, respectively. The inequality $\alpha_i(k)I_p - \overline{N}_i(k)\widehat{\Sigma}_i(k)\overline{N}_i^T(k) > 0$ is equivalent to

$$\alpha_i(k) I_p - N_i(k)\widehat{Q}_i(k) N_i^T(k) > 0. \tag{50}$$

Moreover, $E(e_i^T(k)e_i(k)) \le \|P_i(k)\|_*$. Thus, we can obtain the following result.

Theorem 5. *Consider a system (1) being monitored by sensor networks (2). Then, under Algorithm 1 the mean square estimation error of sensor i is upper-bounded by $\|P_i(k)\|_*$, $i \in \mathcal{V}$.*

In the following, we focus on discussing the boundedness of the sequence $\{P_i(k), k \ge 0\}$. Firstly we define *collaboratively observable* as follows. A node i is said to be *collaboratively observable* if $(A_0, \overline{C}_i)$ is observable.

Suppose the sensor network satisfies the following assumption.

(A1) In the network, for any given $(s,l) \in \mathbb{N}$, for any node $j \in \mathcal{V}$, there exists at least one path from j to the node set $\mathcal{V}^{sl}$, where $\mathcal{V}^{sl} = \{i : b_i^{sl} = 1, i \in \mathcal{V}\}$.

(A2) In the network, there exists at least one collaboratively observable node. Let $\mathcal{V}_0$ denote the set of collaboratively observable nodes; i.e., $\mathcal{V}_0 = \{i \in \mathcal{V} : (A_0, \overline{C}_i)$ *is completely observable*$\}$. The node subset $\mathcal{V}_0$ is globally reachable in the communication topology $\mathcal{G}$.

Notice that in (A2) each node's measurement matrix is not required to be observable. Just one node's aggregated measurement matrix is required to be observable. Although this assumption is a little more conservative than cooperative observability of all sensors, it is easy to be satisfied by properly placing sensors.

Theorem 6. *Consider a system (1) being monitored by sensor networks (2) satisfying (A1) and (A2). For $i \in \mathcal{V}_0$, suppose there exist a constant $\alpha_i > 0$ such that the following three matrix inequalities (51)-(53) have positive definite solutions P_i and Q_i:*

$$\begin{aligned}P_i &> \overline{A}_i P_i \overline{A}_i^T - \overline{A}_i P_i \overline{C}_i^T\left(\overline{C}_i P_i \overline{C}_i^T + \overline{W}_i\right)^{-1}\overline{C}_i P_i \overline{A}_i^T \\ &\quad + \left(\alpha_i I_p + W_0\right) A_0^{-T} R_i A_0^{-1}\left(\alpha_i I_p + W_0\right) + \alpha_i I_p \\ &\quad + W_0,\end{aligned} \tag{51}$$

$$Q_i > A_0 R_i^{-1} A_0^T + \alpha_i I_p + W_0, \tag{52}$$

$$\alpha_i I_p - N_i Q_i N_i^T > 0 \tag{53}$$

where

$$\begin{aligned}\overline{A}_i &= A_0 + \left(\alpha_i I_p + W_0\right) A_0^{-T} R_i, \\ F_i &= \overline{A}_i P_i \overline{C}_i^T\left(\overline{C}_i P_i \overline{C}_i^T + \overline{W}_i\right)^{-1}, \\ N_i &= \lim_{k\longrightarrow\infty}\sup \max_l \left\{\left(\sum_{s=1}^{p} Y_i^{sl}(k)\right)^{1/2}\right\} I_p, \\ R_i &= Q_i^{-1} - \alpha_i^{-1} N_i N_i^T.\end{aligned} \tag{54}$$

Then, under Algorithm 1 the mean square estimation errors of all sensors are bounded.

Proof. From Proposition 3, we have that under (A1), for any node i, $\lim_{k\longrightarrow\infty} E(\Delta A_i^T(k)\Delta A_i(k)) = 0$. Thus, from (A2) there exists a finite time $k_0 \ge 0$, for any node $i \in \mathcal{V}_0$, when $k \ge k_0$, $(A_i(k), \overline{C}_i)$ is observable. From Algorithm 1, in the consensus-based robust Kalman filtering part, if sensor $i \in \mathcal{V}_0$, then $\widehat{x}_i(k) = x_i(k)$, $\widehat{P}_i(k) = P_i(k)$, $\widehat{Q}_i(k) = Q_i(k)$, and

$$x_i(k+1) = \overline{A}_i(k) x_i(k) + F_i(k)\left(z_i(k) - \overline{C}_i x_i(k)\right), \tag{55}$$

$$\begin{aligned}P_i(k+1) &= \left(\overline{A}_i(k) - F_i(k)\overline{C}_i\right) P_i(k)\left(\overline{A}_i(k) - F_i(k)\overline{C}_i\right)^T \\ &\quad + \left(A_i(k) - \overline{A}_i(k)\right) R_i^{-1}(k)\left(A_i(k) - \overline{A}_i(k)\right)^T \\ &\quad + \alpha_i(k) I_p + W_0 + F_i(k) W_i F_i^T(k)\end{aligned} \tag{56}$$

$$Q_i(k+1) = A_i(k) R_i^{-1}(k) A_i^T(k) + \alpha_i(k) I_p + W_0. \tag{57}$$

And the LMI is as follows.

$$\alpha_i(k) I_p - N_i(k) Q_i(k) N_i^T(k) > 0 \tag{58}$$

The above filter is the one discussed in Lemma 4. From Proposition 3, $N_i(k)N_i^T(k) \leq N_iN_i^T$, $N_i = \lim_{k\longrightarrow\infty}\sup \max_l\{(\sum_{s=1}^p Y_i^{sl})^{1/2}\}I_p$ must be bounded, and A_i converges to A_0 in mean square sense, by setting the initial conditions to zero and replacing the Riccati difference equations by the corresponding algebraic Riccati equations ([21]), the stability of the Riccati difference equations can be converted to the feasibility of inequalities (51)-(53). If for $i \in \mathscr{V}_0$, inequalities (51)-(53) are feasible, $\widehat{P}_i(k)$ must be bounded.

In the following, we discuss the other nodes.

Firstly, we prove that if for $i \in \mathscr{V}_0$, (52)-(53) are feasible, for any other node i, there exists $\alpha_i(k) > 0$ such that (50) holds with $Q_i(k)$ satisfying (43).

Define $N_{\max} = \max_i\{N_i\}$; then by replacing N_i with $N_{\max}$, the LMIs (52)-(53) are still feasible, which means there exists $\overline{\alpha}(k) > 0$ such that $\overline{\alpha}(k)I_p - N_{\max}\overline{Q}(k)N_{\max}^T > 0$ and $\overline{Q}(k+1) = A_0(\overline{Q}^{-1}(k) - \overline{\alpha}^{-1}(k)N_{\max}N_{\max}^T)^{-1}A_0^T + \overline{\alpha}(k)I_p + W_0$. When $\overline{Q}(0) \geq Q_i(0)$, since $N_i(k) \leq N_{\max}$, there holds $\overline{Q}(k) \geq Q_i(k)$ and $\overline{\alpha}I_p - N_i(k)Q_i(k)N_i^T(k) > 0$ with $Q_i(k)$ satisfying $Q_i(k+1) = A_0(Q_i^{-1}(k) - \overline{\alpha}^{-1}(k)N_i(k)N_i^T(k))^{-1}A_0^T + \overline{\alpha}(k)I_p + W_0$.

Define $\widehat{Q}_i(k) = \sum_{j\in\mathscr{N}_i} w_{ij}(k)Q_j(k)$. It is obvious that $\overline{Q}(k) \geq \widehat{Q}_i(k)$, $\overline{Q}(k) \geq Q_i(k)$, and $\overline{\alpha}(k)I_p - N_{\max}\widehat{Q}_i(k)N_{\max}^T > 0$ with $Q_i(k)$ satisfying $Q_i(k+1) = A_0(\widehat{Q}_i^{-1}(k) - \overline{\alpha}^{-1}(k)N_i(k)N_i^T(k))^{-1}(k)A_0^T + \overline{\alpha}(k)I_p + W_0$. Since A_i converges to A_0 in mean square sense, there also exists $\alpha_i(k) > 0$ such that (58) holds with $Q_i(k)$ satisfying (43).

Secondly, we investigate the boundedness of the estimation errors. For $i \in \mathscr{V}$, define $\phi_i(k) = \operatorname{tr}(P_i(k))$, $\widehat{\phi}_i(k) = \operatorname{tr}(\widehat{P}_i(k))$. From (40) we obtain that

$$\begin{aligned}\widehat{\phi}_i(k+1) &= \sum_{j\in\mathscr{N}_i} w_{ij}(k)\phi_j(k) = \sum_{j\in\mathscr{N}_i}\frac{1}{\sum_{s\in\mathscr{N}_i}(1/\phi_s(k))} \\ &\leq d_i\phi_j(k).\end{aligned} \tag{59}$$

where $d_i = \sum_{j\in\mathscr{N}_i} a_{ij}$, j is any one of sensor i's neighbors.

Since for $i \in \mathscr{V}_0$, $\widehat{\phi}_i(k)$ and $\phi_i(k)$ are bounded, then for any one of sensor i's out neighbors i_o, $\widehat{\phi}_{i_o}(k)$ must be bounded by $d_{i_o}\phi_i(k)$, which means that the mean square estimation error of node i_o must be also bounded. For node i_o's out neighbor, its mean square estimation error must be also bounded.

Since from (A2), $\mathscr{V}_0$ is globally reachable, for any node i, there exists at least one path from i to some node in $\mathscr{V}_0$. So by the same way, for any node i, $\widehat{\phi}_i(k)$ must be bounded. This completes the proof. □

Remark 7. Compared with the stability results in [6, 7, 14–18], the condition in Theorem 6 is based on some uncorrelated LMIs, which is easy to be verified and does not depend on the global topology information.

Remark 8. The feasibility of (51)-(53) does not depend on the value of N_i as long as N_i is bounded. Since from Proposition 3, $Y_i^{sl}(k)$ must be bounded, the N_i in LMIs (51)-(53) can be replaced by the identity matrix.

5. Numerical Examples

Consider a system with matrix

$$A_0 = \begin{bmatrix} A_{01} & 0 \\ 0 & A_{02}\end{bmatrix}, \tag{60}$$

and system noise covariance $W_0 = I_4$ in a network composed of 5 sensors, where

$$A_{01} = A_{02} = \begin{bmatrix} 0.8 & 0.2 \\ 0 & 0.8\end{bmatrix}. \tag{61}$$

The measurement matrices of sensors are $C_1 = C_3 = C_5 = [1\ \ 1\ \ 0\ \ 0]$, $C_2 = C_4 = [0\ \ 0\ \ 1\ \ 1]$, and the measurement noises have covariances $W_i = 0.01$ for all i. In the network, sensor 1, sensor 2, and sensor 4 can obtain the measurements of the target and partial parameters of system matrix A_0 with noises. The covariance of the noises is equal to 0.01. The adjacency matrix of the communication topology is given by

$$\mathscr{A} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{bmatrix}. \tag{62}$$

Sensor 1 and sensor 3 have access to A_{01} with noises and sensor 2 has an access to A_{02} with noises. Other sensors cannot obtain the system information. Since node 1 and 2 are reachable in the communication topology, from Proposition 3, the mean square estimation errors for the system parameters converge to 0. Figure 1 gives the trajectory of sensor 1's estimation covariance for system parameter A_0^{33} and shows its convergence.

Nodes 1, 2, and 4 are collectively observable, and the LMIs in Theorem 6 are feasible. Set $P_i(0) = Q_i(0) = I_4$ and $\alpha_i(k) = 1$. Under Algorithm 1, 50 Monte Carlo simulations are carried out and the estimation error and covariance of each sensor are shown in Figures 2 and 3. Obviously, each sensor's estimation error is less than its trace of covariance and bounded.

6. Conclusion

This paper mainly proposes a fully distributed robust Kalman filtering algorithm for discrete-time-invariant systems over sensor networks in which each sensor can just obtain partial system parameters and measurement outputs with noises. In this algorithm, each sensor just uses its neighbors' information to update filtering gains and weights. If in the network for any one of system parameters its accessible node subset is globally reachable, and there exists at least one globally reachable collaboratively observable node subset, then as long as a set of uncorrelated LMIs are feasible, the mean square estimation errors of the sensors must be bounded. The distributed filtering problem of time-varying networks is of our research interest in future.

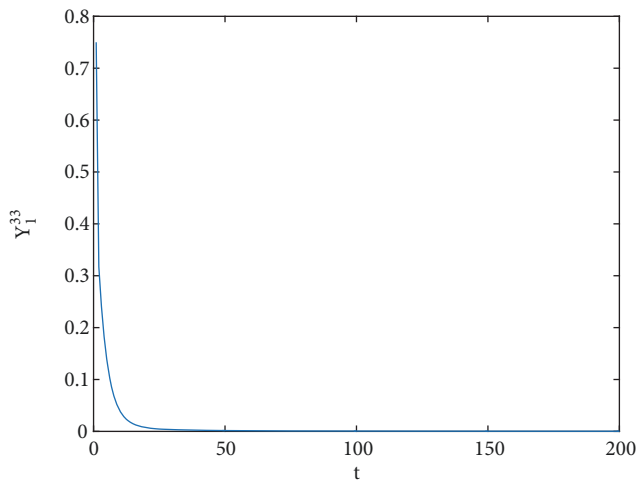

Figure 1: Parameter estimation covariance of sensor 1.

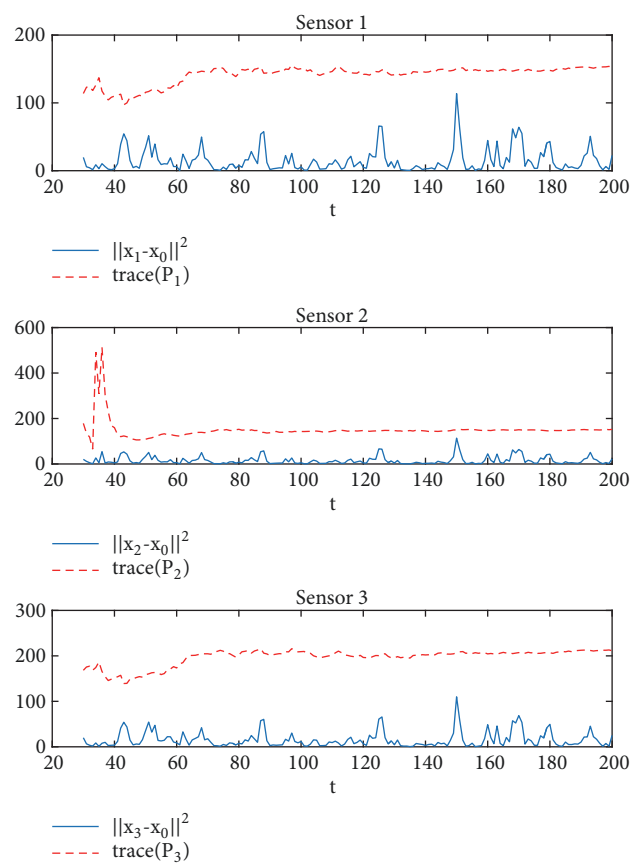

Figure 2: State estimation errors and covariances of sensors 1-3.

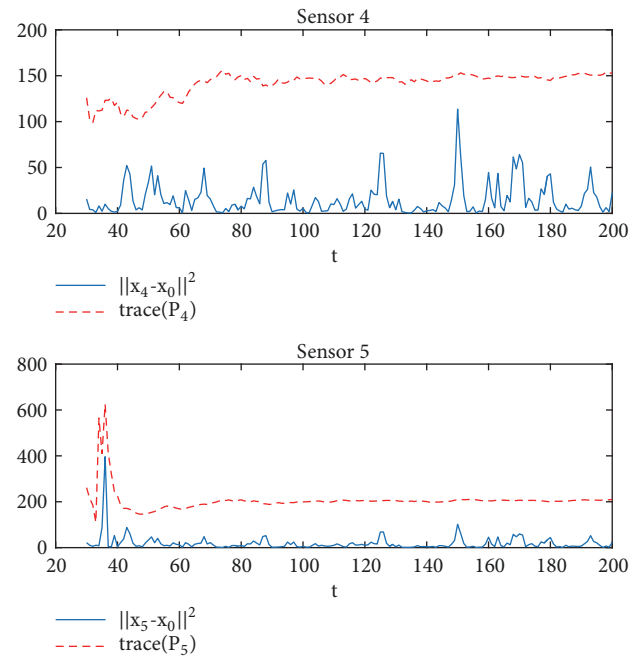

Figure 3: State estimation errors and covariances of sensors 4-5.

Data Availability

(1) The analysis data used to support the findings of this study are included within the article. (2) The programing code of the simulation example used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work is supported by National Natural Science Foundation (NNSF) of China under Grant 61473081, Natural Science Foundation of Nanjing Institute of Technology under Grant YKJ201431, Six Talent Peaks Project in Jiangsu Province under Grant XYDXX-005, Natural Science Foundation of Jiangsu Province under Grant BK20171362.

References

[1] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," in *Proceedings of the 46th IEEE Conference on Decision and Control*, pp. 5492–5498, December 2007.

[2] S. S. Stanković, M. S. Stanković, and D. M. Stipanović, "Consensus based overlapping decentralized estimation with missing observations and communication faults," *Automatica*, vol. 45, no. 6, pp. 1397–1406, 2009.

[3] W. Li, Y. Jia, and J. Du, "Distributed Kalman consensus filter with intermittent observations," *Journal of The Franklin Institute*, vol. 352, no. 9, pp. 3764–3781, 2015.

[4] F. S. Cattivelli and A. H. Sayed, "Diffusion strategies for distributed Kalman filtering and smoothing," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2069–2084, 2010.

[5] W. Yang, G. Chen, X. Wang, and L. Shi, "Stochastic sensor activation for distributed state estimation over a sensor network," *Automatica*, vol. 50, no. 8, pp. 2070–2076, 2014.

[6] V. Ugrinovskii, "Distributed robust filtering with H_∞ consensus of estimates," *Automatica*, vol. 47, no. 1, pp. 1–13, 2011.

[7] V. Ugrinovskii, "Distributed robust estimation over randomly switching networks using H_∞ consensus," *Automatica*, vol. 49, no. 1, pp. 160–168, 2013.

[8] W. Li and Y. Jia, "Distributed consensus filtering for discrete-time nonlinear systems with non-Gaussian noise," *Signal Processing*, vol. 92, no. 10, pp. 2464–2470, 2012.

[9] 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, 2014.

[10] G. Battistelli, L. Chisci, G. Mugnai, A. Farina, and A. Graziano, "Consensus-Based Linear and Nonlinear Filtering," *IEEE Transactions on Automatic Control*, vol. 60, no. 5, pp. 1410–1415, 2015.

[11] G. Battistelli and L. Chisci, "Stability of consensus extended Kalman filter for distributed state estimation," *Automatica*, vol. 68, pp. 169–178, 2016.

[12] W. Li, G. Wei, F. Han, and Y. Liu, "Weighted average consensus-based unscented Kalman filtering," *IEEE Transactions on Cybernetics*, 2015.

[13] M. Diao, Z. Duan, and G. Wen, "A global detectability condition for consensus tracking of linear multi-agent systems with stochastic disturbances," *Asian Journal of Control*, vol. 18, no. 1, pp. 357–366, 2016.

[14] P. Millán, L. Orihuela, C. Vivas, and F. R. Rubio, "Distributed consensus-based estimation considering network induced delays and dropouts," *Automatica*, vol. 48, no. 10, pp. 2726–2729, 2012.

[15] I. Matei and J. S. Baras, "Consensus-based linear distributed filtering," *Automatica*, vol. 48, no. 8, pp. 1776–1782, 2012.

[16] M. Doostmohammadian and U. A. Khan, "On the genericity properties in distributed estimation: Topology design and sensor placement," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 2, pp. 195–204, 2013.

[17] B. Açikmeşe, M. Mandić, and J. L. Speyer, "Decentralized observers with consensus filters for distributed discrete-time linear systems," *Automatica*, vol. 50, no. 4, pp. 1037–1052, 2014.

[18] C. E. de Souza, D. Coutinho, and M. Kinnaert, "Mean square state estimation for sensor networks," *Automatica*, vol. 72, pp. 108–114, 2016.

[19] Z. Dong and Z. You, "Finite-horizon robust Kalman filtering for uncertain discrete time-varying systems with uncertain-covariance white noises," *IEEE Signal Processing Letters*, vol. 13, no. 8, pp. 493–496, 2006.

[20] S. M. Mohamed and S. Nahavandi, "Robust finite-horizon Kalman filtering for uncertain discrete-time systems," *IEEE Transactions on Automatic Control*, vol. 57, no. 6, pp. 1548–1552, 2012.

[21] Y. S. Hung and F. Yang, "Robust H_∞ filtering with error variance constraints for discrete time-varying systems with uncertainty," *Automatica*, vol. 39, no. 7, pp. 1185–1194, 2003.

[22] Y. Zhang, D. Chen, C. Liu, and Y. Chen, "Distributed Kalman/H∞ Filtering for Sensor Networks with Partial Unknown System Parameters," in *Proceedings of the 2018 IEEE 14th International Conference on Control and Automation (ICCA)*, pp. 1–6, IEEE, Anchorage, AK, USA, June 2018.

[23] F. Han, G. Wei, D. Ding, and Y. Song, "Local condition based consensus filtering with stochastic nonlinearities and multiple missing measurements," *IEEE Transactions on Automatic Control*, vol. 62, no. 9, pp. 4784–4790, 2017.