Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Convergence analysis of distributed Kalman filtering for relative sensing networks*

Che LIN[1], Rong-hao ZHENG[‡1,2], Gang-feng YAN[1], Shi-yuan LU[1]

[1]*College of Electrical Engineering, Zhejiang University, Hangzhou 310027, China*

[2]*Zhejiang Province Marine Renewable Energy Electrical Equipment and System Technology Research Laboratory, Hangzhou 310027, China*

E-mail: linche@zju.edu.cn; rzheng@zju.edu.cn; ygf@zju.edu.cn; 917964950@qq.com

Received Jan. 24, 2017; Revision accepted Mar. 23, 2017; Crosschecked July 8, 2018; Published online Sept. 12, 2018

Abstract: We study the distributed Kalman filtering problem in relative sensing networks with rigorous analysis. The relative sensing network is modeled by an undirected graph while nodes in this network are running homogeneous dynamical models. The sufficient and necessary condition for the observability of the whole system is given with detailed proof. By local information and measurement communication, we design a novel distributed suboptimal estimator based on the Kalman filtering technique for comparison with a centralized optimal estimator. We present sufficient conditions for its convergence with respect to the topology of the network and the numerical solutions of n linear matrix inequality (LMI) equations combining system parameters. Finally, we perform several numerical simulations to verify the effectiveness of the given algorithms.

Key words: Relative sensing network; Distributed Kalman filter; Schur stable; Linear matrix inequality
https://doi.org/10.1631/FITEE.1700066 **CLC number:** TP13

1 Introduction

Wireless sensor networks (WSNs), which include tens of thousands of wireless sensor nodes geographically distributed in certain areas, have recently attracted many researchers' attention. Generally, the sensors involved in these networks are intelligent nodes with limited computational ability, constrained power supply, and communication ranges. Due to their insights in a variety of real situations, they have a wide range of applications, such as environmental monitoring (Garcia-Sanchez et al., 2011; Li et al., 2013), intelligent transportation (Tubaishat et al., 2009), cooperative robotics (Lin et al., 2014), surveillance missions (Chen et al., 2013), smart grids (Pasqualetti et al., 2012; Fadel et al., 2015), and cyber-physical systems (CPSs) monitoring (Chen et al., 2015).

One important problem with WSNs is how to extract information about the state vectors of all the nodes from observations that are contaminated with external disturbances. It is generally known that the traditional Kalman filter algorithm is optimal for linear systems with exact system models, but it needs a fusion center to gather all the system information. However, in large-scale WSNs, a centralized estimator for the whole system is both restrictive and infeasible, which motivates us to focus on distributed algorithms. Consensus methodology is widely adopted in distributed algorithm design. Kalman consensus filtering (KCF), which combines Kalman filtering and the consensus algorithm, is to

‡ Corresponding author

* Project supported by the National Natural Science Foundation of China (No. 61503335) and the Key Laboratory of System Control and Information Processing, China (No. Scip201504)

ORCID: Rong-hao ZHENG, https://orcid.org/0000-0002-9095-5905

solve the common noisy estimate problem (Olfati-Saber, 2009; Olfati-Saber and Jalalkamali, 2012). According to the consensus algorithm, Chen et al. (2015) designed a distributed tree-based broadcasting communication algorithm and a consensus-based distributed estimation method for ubiquitous monitoring in CPSs. More recently, Gao et al. (2016) designed a distributed estimator with local unreliable measurements, and the stability condition of the estimation error covariance was established. In addition, Duan and Li (2011) designed fusion-based distributed estimators to reduce the calculation complexity of previous works.

In this paper, we concentrate on the distributed state estimation problem for relative sensing networks. A relative sensing network generally means that a node in the network can measure only the relative states between itself and its neighbors. A small number of nodes can also measure their absolute states; these nodes are called 'anchor nodes'. The relative sensing networks, which are widely considered in network localization and formation problems, are a common type of network. Morbidi et al. (2010), Piovan et al. (2013), Ravazzi et al. (2013), and Lin et al. (2014) proposed distributed algorithms to address localization and formation problems in relative sensing networks.

Instead of solving the specific localization/formation problem, we aim to solve a more general state-estimation problem in relative sensing networks with measurement noises. In contrast to the common state estimation problem (Olfati-Saber and Jalalkamali, 2012; Zhou et al., 2013), distributed estimation for relative sensing networks is designed to estimate the absolute state of each node individually, using its own available measurements and information possibly exchanged with its neighbors. Suppose that there is a group of sensor nodes in the network; each node's state represents some specific information, for instance, its position. An undirected graph is used to describe the sensing and communication relationship of the network. It is assumed that all nodes can measure the relative states of its neighbors and itself, and only a small number of anchor nodes can additionally know their own absolute state measurement. A centralized optimal estimator is then constructed using the standard Kalman filter, which minimizes the estimation error of the whole system, while the graph is connected and there is at least one anchor node. However, because the optimal estimator requires all-to-all communication, which is not feasible in large-scale networks, a distributed estimator is proposed that ignores the correlation between the estimates by non-neighboring nodes. The distributed algorithm reduces the communication rate, improves the estimate scalability of the system, but still includes the basic idea of the Kalman filter. It can intuitively perform quite well, but its convergence seems challenging because it does not have a standard Kalman filter form and is affected by the system topology and parameters. Hence, we design a novel method to prove the convergence of the distributed algorithm, which effectively solves this problem while traditional analysis methods fail. Compared with our previous work (Lu et al., 2015), this paper contains two significant contributions: (1) it presents the sufficient and necessary condition of the observability of the whole system with detailed proof; (2) it gives a rigorous analysis of the convergence of the distributed estimator with a sufficient condition related to the topology of the system and the system's parameters. We perform simulations to demonstrate the performance of the proposed estimation algorithms.

Notations: $\mathrm{tr}(\cdot)$ represents the trace of a matrix. $\boldsymbol{I}_n \in \mathbb{R}^n$ is the unit matrix of n dimensions. The number of elements in a set N_i is represented by $|N_i|$. $\mathbf{0}$ is a matrix with all incidence being 0 and is of proper size if mentioned.

2 Preliminaries and problem statement

2.1 Preliminaries

We deal with the problem of distributed estimation in a relative sensing network. An undirected graph $G = (V, E)$ is adopted to describe the interconnected structure of the network. We denote the set of vertices and the set of edges in graph G by V and E, respectively. Each vertex i corresponds to a node in the network. We write $(i, j) \in E$ as the edge between vertices i and j. If (i, j) exists, nodes i and j are neighbors and are able to have relative measurements of and communication with each other. We denote N_i as the neighbor set of node i in G.

In a relative sensing network, there are two types of nodes, called 'anchor nodes' and 'sensing nodes', respectively. Both anchor nodes and sensing nodes are called sensor nodes in this paper. Each node, which is either an anchor node or a sensing node, has the relative state measurements between itself and its neighbors. However, anchor nodes can additionally have the measurements of their own absolute states. A relative sensing network in Fig. 1 is modeled by an undirected graph. Each vertex represents a node in the network and data are exchanged between neighbors. For consistency, a virtual node 0 is introduced and has edges to the anchor nodes. Therefore, the relative measurements between an anchor node and the virtual node are actually the measurements of the absolute states of the anchor node. For example, in Fig. 1, node 2 is the neighbor of node 1, which means that node 2 has the relative measurements between nodes 1 and 2, and also receives data from node 3. Nodes 3 and 5 are anchor nodes and have additional measurements of their own absolute states.

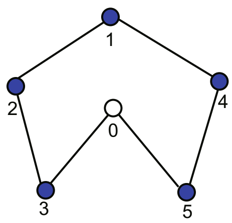

Fig. 1 Example of a relative sensing network modeled by an undirected graph

Next, we introduce a modified edge-vertex incidence matrix $\boldsymbol{F}$. Every entry $F_{e,i}$ in the matrix with rows labeled by the edge and columns labeled by the vertex in G satisfies

$$F_{e,i} = \begin{cases} 1, & \text{if } e=(i,j), \\ -1, & \text{if } e=(j,i), \\ 0, & \text{otherwise.} \end{cases}$$

For two edges $e_k = (i_k, j_k)$ and $e_l = (i_l, j_l)$, the row corresponding to edge e_k is indexed with a smaller row number if $j_k < j_l$ or if $j_k = j_l$ and $i_k < i_l$. Given an example graph in Fig. 2, the incidence matrix $\boldsymbol{F}$ is

$$\boldsymbol{F} = \begin{array}{c|cccc} E\backslash V & 0 & 1 & 2 & 3 \\ \hline (1,0) & -1 & 1 & 0 & 0 \\ (0,1) & 1 & -1 & 0 & 0 \\ (2,1) & 0 & -1 & 1 & 0 \\ (3,1) & 0 & -1 & 0 & 1 \\ (1,2) & 0 & 1 & -1 & 0 \\ (1,3) & 0 & 1 & 0 & -1 \end{array}.$$

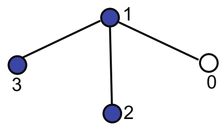

Fig. 2 Sample relative sensing network to illustrate the incidence matrix and observation matrix

2.2 Problem statement

For a relative sensing network with n nodes, suppose that each node $i \in \{1, 2, \ldots, n\}$ has a dynamical model:

$$\boldsymbol{x}_i = \boldsymbol{A}_i \boldsymbol{x}_i^- + \boldsymbol{w}_i, \quad \boldsymbol{w}_i \sim N(\boldsymbol{0}, \boldsymbol{Q}_i), \tag{1}$$

where $\boldsymbol{x}_i \in \mathbb{R}^p$ is the state of the node i at time step k. $\boldsymbol{w}_i \in \mathbb{R}^p$ is the Gaussian white process noise and its covariance matrix is $\boldsymbol{Q}_i$. We use briefly $\boldsymbol{x}_i$ instead of $\boldsymbol{x}_i(k)$ and $\boldsymbol{x}_i^-$ instead of $\boldsymbol{x}_i(k-1)$.

Because homogeneous networks widely exist in many applications, we assume that the relative sensing network under study is homogeneous, i.e., $\boldsymbol{A}_i = \boldsymbol{A}_j = \boldsymbol{A}$ for all i, j, and $\boldsymbol{w}_j$ and $\boldsymbol{w}_i$ are uncorrelated. Let $\boldsymbol{w}_{ij} = \boldsymbol{w}_j - \boldsymbol{w}_i$ and $\boldsymbol{Q}_{ij} = \boldsymbol{Q}_j + \boldsymbol{Q}_i$. The relative state $\boldsymbol{x}_{ij} = \boldsymbol{x}_j - \boldsymbol{x}_i$ can be described as

$$\begin{aligned} \boldsymbol{x}_{ij} &= \boldsymbol{A}_j \boldsymbol{x}_j^- - \boldsymbol{A}_i \boldsymbol{x}_i^- + \boldsymbol{w}_j - \boldsymbol{w}_i \\ &= \boldsymbol{A} \boldsymbol{x}_{ij}^- + \boldsymbol{w}_{ij}, \quad \boldsymbol{w}_{ij} \sim N(\boldsymbol{0}, \boldsymbol{Q}_{ij}). \end{aligned}$$

In a relative sensing network, all nodes have relative measurements. The measurement $\boldsymbol{y}_{ij} \in \mathbb{R}^{q_{ij}}$ of the relative state $\boldsymbol{x}_{ij}$ in node i is modeled as

$$\boldsymbol{y}_{ij} = \boldsymbol{C}_{ij} \boldsymbol{x}_{ij} + \boldsymbol{v}_{ij}, j \in N_i, \quad \boldsymbol{v}_{ij} \sim N(\boldsymbol{0}, \boldsymbol{R}_{ij}), \tag{2}$$

where $\boldsymbol{v}_{ij}$ is the Gaussian white measurement noise and is independent of any other measurement noise.

Moreover, anchor nodes can have additional measurements of their own absolute states. By introducing the virtual node 0 and $\boldsymbol{x}_0 \equiv \boldsymbol{0}, \boldsymbol{w}_0 \equiv \boldsymbol{0}$, the additional measurements of the absolute states by the anchor nodes are described as

$$\boldsymbol{y}_{i0} = \boldsymbol{C}_{i0} \boldsymbol{x}_{i0} + \boldsymbol{v}_{i0}, \quad \boldsymbol{v}_{i0} \sim N(\boldsymbol{0}, \boldsymbol{R}_{i0}), \tag{3}$$

where $\boldsymbol{x}_{i0} = -\boldsymbol{x}_i$ and $\boldsymbol{v}_{i0}$ is the Gaussian white measurement noise with covariance matrix $\boldsymbol{R}_{i0}$.

Now we are ready to present our problem statement after giving several technical assumptions.

Assumption 1 Suppose that each pair $(\boldsymbol{A}, \boldsymbol{C}_{ij})$, $i \in V, i \neq 0, j \in N_i$, is observable.

Assumption 2 The undirected graph of the system is connected, and the topology of the graph is fixed.

Problem 1 Consider a homogeneous relative sensing network modeled by an undirected graph and suppose that (a) each node is governed by the dynamical model in Eq. (1), (b) each node has the measurements of the relative states between itself and its neighbors, modeled in Eq. (2), and (c) each anchor node also has the measurements of its own absolute state, modeled in Eq. (3). Find a distributed estimation scheme to reduce the overall estimation error of the large-scale network.

3 Optimal Kalman estimator

In this section, we formulate an optimal estimator based on the standard Kalman filter to estimate the global states of all nodes in a centralized form. Then we prove the full state observability condition of the relative sensing network.

3.1 Centralized optimal estimator

Define $\bar{\boldsymbol{x}}_i$ and $\hat{\boldsymbol{x}}_i$ as the priori estimate and posteriori estimate of state $\boldsymbol{x}_i$, respectively. Let

$$\bar{\boldsymbol{\eta}}_i = \bar{\boldsymbol{x}}_i - \boldsymbol{x}_i \text{ and } \boldsymbol{\eta}_i = \hat{\boldsymbol{x}}_i - \boldsymbol{x}_i \tag{4}$$

denote the priori and posteriori estimation errors, respectively. Then the error covariance matrices $\bar{\boldsymbol{P}}_i$ and $\boldsymbol{P}_i$ associated with the estimates $\bar{\boldsymbol{x}}_i$ and $\hat{\boldsymbol{x}}_i$ are given by

$$\bar{\boldsymbol{P}}_i = E[\bar{\boldsymbol{\eta}}_i\bar{\boldsymbol{\eta}}_i^{\mathrm{T}}] \text{ and } \boldsymbol{P}_i = E[\boldsymbol{\eta}_i\boldsymbol{\eta}_i^{\mathrm{T}}]. \tag{5}$$

To formulate the optimal estimator, $\boldsymbol{X} \in \mathbb{R}^{np\times 1}$ denotes the overall state and is organized as

$$\boldsymbol{X} = [\boldsymbol{x}_1^{\mathrm{T}}, \boldsymbol{x}_2^{\mathrm{T}}, \ldots, \boldsymbol{x}_n]^{\mathrm{T}},$$

where $\boldsymbol{x}_i$ is the state of node i. $\boldsymbol{Y} \in \mathbb{R}^{q\times 1}, q = \sum_i q_i$ denotes the overall measurement and is organized as

$$\boldsymbol{Y} = [\boldsymbol{Y}_1^{\mathrm{T}}, \boldsymbol{Y}_2^{\mathrm{T}}, \ldots, \boldsymbol{Y}_n^{\mathrm{T}}]^{\mathrm{T}},$$

where $\boldsymbol{Y}_i \in \mathbb{R}^{q_i\times 1}, q_i = \sum_{j\in N_i} q_{ij}$ is the total measurement (including both absolute and relative measurements) observed by node i and is organized as

$$\boldsymbol{Y}_i = [y_{ij_1}, y_{ij_2}, \ldots, y_{ij_{|N_i|}}]^{\mathrm{T}},$$

where $j_\bullet \in N_i$. Assume that the Gaussian white noises are independent. Then the aggregate noises are

$$\begin{cases} \boldsymbol{W} = [\boldsymbol{w}_1^{\mathrm{T}}, \boldsymbol{w}_2^{\mathrm{T}}, \ldots, \boldsymbol{w}_n^{\mathrm{T}}]^{\mathrm{T}}, & \boldsymbol{W} \in \mathbb{R}^{np\times 1}, \\ \boldsymbol{V} = [\boldsymbol{V}_1^{\mathrm{T}}, \boldsymbol{V}_2^{\mathrm{T}}, \ldots, \boldsymbol{V}_n^{\mathrm{T}}]^{\mathrm{T}}, & \boldsymbol{W} \in \mathbb{R}^{q\times 1}, \\ \boldsymbol{V}_i = [v_{ij_1}, v_{ij_2}, \ldots, v_{ij_{|N_i|}}]^{\mathrm{T}}, \boldsymbol{V}_i \in \mathbb{R}^{q_i\times 1}, \end{cases}$$

and the covariances of noises are

$$\begin{cases} \boldsymbol{Q} = \mathrm{diag}(\boldsymbol{Q}_1, \boldsymbol{Q}_2, \ldots, \boldsymbol{Q}_n), & \boldsymbol{Q} \in \mathbb{R}^{np\times np}, \\ \boldsymbol{R} = \mathrm{diag}(\boldsymbol{R}_1, \boldsymbol{R}_2, \ldots, \boldsymbol{R}_n), & \boldsymbol{R} \in \mathbb{R}^{q\times q}, \\ \boldsymbol{R}_i = \mathrm{diag}(\boldsymbol{R}_{ij_1}, \boldsymbol{R}_{ij_2}, \ldots, \boldsymbol{R}_{ij_{|N_i|}}), & \boldsymbol{R}_i \in \mathbb{R}^{q_i\times q_i}. \end{cases}$$

Denote $\boldsymbol{A}^* \in \mathbb{R}^{np\times np}$ as the overall system matrix and $\boldsymbol{H} \in \mathbb{R}^{q\times np}$ as the observation matrix. Here $\boldsymbol{A}^* = \boldsymbol{I}_n \otimes \boldsymbol{A}$ and $\boldsymbol{H}$ is a block matrix, which will be explained below. Moreover, denote the overall estimation error covariance $\boldsymbol{P}, \bar{\boldsymbol{P}} \in \mathbb{R}^{np\times np}$, and the Kalman filter gain as $\boldsymbol{K} \in \mathbb{R}^{np\times q}$. Therefore, the centralized optimal estimator is described as

$$\begin{cases} \bar{\boldsymbol{X}} = \boldsymbol{A}^*\hat{\boldsymbol{X}}^-, \\ \hat{\boldsymbol{X}} = \bar{\boldsymbol{X}} + \boldsymbol{K}(\boldsymbol{Y} - \boldsymbol{H}\bar{\boldsymbol{X}}), \\ \boldsymbol{K} = \bar{\boldsymbol{P}}\boldsymbol{H}^{\mathrm{T}}(\boldsymbol{R} + \boldsymbol{H}\bar{\boldsymbol{P}}\boldsymbol{H}^{\mathrm{T}})^{-1}, \\ \bar{\boldsymbol{P}} = \boldsymbol{A}^*\boldsymbol{P}^-\boldsymbol{A}^{*\mathrm{T}} + \boldsymbol{Q}, \\ \boldsymbol{P} = (\boldsymbol{I} - \boldsymbol{K}\boldsymbol{H})\bar{\boldsymbol{P}}. \end{cases} \tag{6}$$

Now we present the observation matrix $\boldsymbol{H}$. Because $\boldsymbol{x}_{ij} = \boldsymbol{x}_j - \boldsymbol{x}_i$, from Eq. (2), we have $\boldsymbol{y}_{ij} = -\boldsymbol{C}_{ij}\boldsymbol{x}_i + \boldsymbol{C}_{ij}\boldsymbol{x}_j + \boldsymbol{v}_{ij}$. $\boldsymbol{H}$ is thus generated based on the incidence matrix $\boldsymbol{F}$ defined in Section 2.1. Because the state $\boldsymbol{x}_0$ of the virtual node is not included in $\boldsymbol{X}$ and is always $\boldsymbol{0}$, to generate $\boldsymbol{H}$, divide $\boldsymbol{F}$ into four blocks as

$$\boldsymbol{F} = \begin{bmatrix} \boldsymbol{F}^a & \boldsymbol{F}^b \\ \boldsymbol{F}^c & \boldsymbol{F}^d \end{bmatrix},$$

where $[\boldsymbol{F}^a, \boldsymbol{F}^b]$ are rows corresponding to edge $(i, 0)$, i is an anchor node, and $\left[(\boldsymbol{F}^a)^{\mathrm{T}}, (\boldsymbol{F}^c)^{\mathrm{T}}\right]^{\mathrm{T}}$ are columns corresponding to node 0. Let $\boldsymbol{h}_{l,k}$ be the blocked element of matrix $\boldsymbol{H}$. Then $\boldsymbol{h}_{l,k} = F_{l,k}^d\boldsymbol{C}_{ij}$, where for any given l corresponding to edge (i, j), $F_{l,i}^d = -1$, $F_{l,j}^d = 1, j \neq 0$ and $F_{l,k}^d = 0, k \neq i$ or j.

Example 1 Assume the network is modeled by Fig. 2 with its incidence matrix being divided as

$$\boldsymbol{F} = \left[\begin{array}{c:ccc} -1 & 1 & 0 & 0 \\ \hdashline 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 1 & 0 & -1 \end{array}\right].$$

The measurements in the network are represented by

$$\boldsymbol{Y} = \left[\begin{array}{c} \boldsymbol{Y}_1 \\ \hdashline \boldsymbol{Y}_2 \\ \hdashline \boldsymbol{Y}_3 \end{array}\right] = \left[\begin{array}{c} \boldsymbol{y}_{10} \\ \boldsymbol{y}_{12} \\ \boldsymbol{y}_{13} \\ \hdashline \boldsymbol{y}_{21} \\ \hdashline \boldsymbol{y}_{31} \end{array}\right].$$

Thus, the corresponding observation matrix $\boldsymbol{H}$ is of the following form:

$$\boldsymbol{H}=\begin{bmatrix} -\boldsymbol{C}_{10} & 0 & 0 \\ -\boldsymbol{C}_{12} & \boldsymbol{C}_{12} & 0 \\ -\boldsymbol{C}_{13} & 0 & \boldsymbol{C}_{13} \\ \boldsymbol{C}_{21} & -\boldsymbol{C}_{21} & 0 \\ \boldsymbol{C}_{31} & 0 & -\boldsymbol{C}_{31} \end{bmatrix}.$$

3.2 Full state observability

Because Eq. (6) is a standard Kalman filter and it is optimal if $(\boldsymbol{A}^*, \boldsymbol{H})$ is observable, the following theorem shows the full state observability condition in terms of graphical connectivity:

Theorem 1 Under Assumptions 1 and 2, a relative sensing network is full state observable if and only if graph G is connected.

To prove Theorem 1, we need to analyze $\boldsymbol{F}^d$ first because $\boldsymbol{H}$ is constructed based on $\boldsymbol{F}^d$. Denote $\boldsymbol{F}_{l,:}$ as a row in $\boldsymbol{F}$. Because the network between the anchor nodes and their neighboring sensing nodes sense is undirected, for any row $\boldsymbol{F}_{l,:}$ in matrix $[\boldsymbol{F}^a, \boldsymbol{F}^b]$, there is a row $\boldsymbol{F}_{l',:}$ in matrix $[\boldsymbol{F}^c, \boldsymbol{F}^d]$ with $\boldsymbol{F}_{l,:} = -\boldsymbol{F}_{l',:}$. Therefore, $\operatorname{rank}(\boldsymbol{F}) = \operatorname{rank}([\boldsymbol{F}^c, \boldsymbol{F}^d])$. Denote $\boldsymbol{F}_{:,k}$ as a column in $[\boldsymbol{F}^c, \boldsymbol{F}^d]$. Because $\sum \boldsymbol{F}_{:,k} = \boldsymbol{0}$, $\operatorname{rank}([\boldsymbol{F}^c, \boldsymbol{F}^d]) = \operatorname{rank}(\boldsymbol{F}^d)$. We then introduce a lemma regarding the relationship between the rank of the incidence matrix and the number of connected components in the graph.

Lemma 1 (Godsil and Royle, 2001) Let G be a graph with n vertices and c connected components. If $\boldsymbol{F}$ is the incidence matrix of G, then $\operatorname{rank}(\boldsymbol{F}) = n - c$.

Because the rank of the modified incidence matrix in this study is the same as that in Godsil and Royle (2001), it is ready to provide the proof for the theorem.

Proof There are n sensor nodes in the network and by introducing the virtual node, the vertex number reaches $n+1$. If the graph is connected, $c=1$ and $\operatorname{rank}(\boldsymbol{F}) = n+1-1 = n$ according to Lemma 1. Therefore, $\operatorname{rank}(\boldsymbol{F}^d) = \operatorname{rank}(\boldsymbol{F}) = n$.

Note that $\boldsymbol{H}$ is a block matrix with $\boldsymbol{h}_{l,k}$ that is either $\boldsymbol{C}_{ij}$, $-\boldsymbol{C}_{ij}$, or $\boldsymbol{0}$. Denote a block matrix $\boldsymbol{U}$ with each instance:

$$\boldsymbol{U}_{l,k}=\begin{bmatrix} \boldsymbol{h}_{l,k} \\ \boldsymbol{h}_{l,k}\boldsymbol{A} \\ \vdots \\ \boldsymbol{h}_{l,k}\boldsymbol{A}^{p-1} \end{bmatrix},$$

$$\operatorname{rank}(\boldsymbol{U}_{l,k}) = \begin{cases} p, & \text{if } \boldsymbol{h}_{l,k} \neq \boldsymbol{0}, \\ 0, & \text{if } \boldsymbol{h}_{l,k} = \boldsymbol{0}, \end{cases}$$

because each pair $(\boldsymbol{A}, \boldsymbol{C}_{ij})$ is observable. Through elementary row operations, we have

$$\operatorname{rank}(\boldsymbol{U}) = \operatorname{rank}\left(\begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H}\boldsymbol{A}^* \\ \vdots \\ \boldsymbol{H}\boldsymbol{A}^{*p-1} \end{bmatrix}\right).$$

Because $\boldsymbol{A}^* = \boldsymbol{I}_n \otimes \boldsymbol{A}$, it can also be inferred that

$$\begin{aligned} \operatorname{rank}(\boldsymbol{U}) &= \operatorname{rank}\left(\begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H}\boldsymbol{A}^* \\ \vdots \\ \boldsymbol{H}\boldsymbol{A}^{*p} \end{bmatrix}\right) = \ldots \\ &= \operatorname{rank}\left(\begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H}\boldsymbol{A}^* \\ \vdots \\ \boldsymbol{H}\boldsymbol{A}^{*np-1} \end{bmatrix}\right). \end{aligned} \tag{7}$$

Therefore, if $\operatorname{rank}(\boldsymbol{U}) = np$, then $(\boldsymbol{A}^*, \boldsymbol{H})$ is observable.

Now we want to show that $\operatorname{rank}(\boldsymbol{U}) = np$. Notice that $\boldsymbol{U}$ is a matrix constructed in a way similar to $\boldsymbol{H}$ with each row $\boldsymbol{U}_{l,:}$ corresponding to an edge in the graph. Through elementary row operations, the rows are rearranged and $\boldsymbol{U}$ can be transformed into $\bar{\boldsymbol{U}} = \left[(\bar{\boldsymbol{U}}^a)^{\mathrm{T}}, (\bar{\boldsymbol{U}}^b)^{\mathrm{T}}\right]^{\mathrm{T}}$, where $\bar{\boldsymbol{U}}^a$ is a matrix with rows corresponding to all edges in the spanning tree of the graph and $\bar{\boldsymbol{U}}^b$ contains the remaining rows. All rows in $\bar{\boldsymbol{U}}^a$ are arranged as follows: For two rows $\boldsymbol{U}_{l_1,:}$ and $\boldsymbol{U}_{l_2,:}$ corresponding to edges (i_1, j_1) and (i_2, j_2), respectively, $\boldsymbol{U}_{l_1,:}$ has a smaller row number in $\bar{\boldsymbol{U}}^a$ if the distance in the graph between node i_1 and the virtual node is less than that of node i_2. Therefore, $\bar{\boldsymbol{U}}^a$ becomes a lower triangle block matrix. Because $\operatorname{rank}(\boldsymbol{U}_{l,k}) = p$ when $\boldsymbol{U}_{l,k} \neq \boldsymbol{0}$ and there are n blocks in a row, $\operatorname{rank}(\bar{\boldsymbol{U}}^a) = np$.

Referring to the fact that for any two matrices $\boldsymbol{A} \in \mathbb{R}^{a\times n}, \boldsymbol{B} \in \mathbb{R}^{b\times n}$, if $n = \operatorname{rank}(\boldsymbol{A}) \geqslant \operatorname{rank}(\boldsymbol{B})$, then $\operatorname{rank}(\left[\boldsymbol{A}^{\mathrm{T}}, \boldsymbol{B}^{\mathrm{T}}\right]) = n$, and we have $\operatorname{rank}(\boldsymbol{U}) = \operatorname{rank}(\bar{\boldsymbol{U}}) = \operatorname{rank}(\bar{\boldsymbol{U}}^a) = np$. Consequently, $(\boldsymbol{A}^*, \boldsymbol{H})$ is observable if the graph is connected.

We prove it in a contrapositive form.

Assume that the graph is not connected and is divided into $c > 1$ components. If the nodes are labeled properly, then the incidence matrix $\boldsymbol{F}$ becomes a block diagonal matrix $\boldsymbol{F} = \mathrm{diag}(\boldsymbol{F}^1, \boldsymbol{F}^2, \ldots, \boldsymbol{F}^c)$, so does the matrix $\boldsymbol{U}$ corresponding to $\boldsymbol{F}$ with $\boldsymbol{U} = \mathrm{diag}(\boldsymbol{U}^1, \boldsymbol{U}^2, \ldots, \boldsymbol{U}^c)$. Assume that each component contains n_i sensor nodes and that there is a virtual node in one of the components. We then have $\mathrm{rank}(\boldsymbol{F}^1) = n_1$ and $\mathrm{rank}(\boldsymbol{F}^i) = n_i - 1, i = 2, 3, \ldots, c$ without loss of generality. From the above discussion, $\mathrm{rank}(\boldsymbol{U}_1) = n_1 p$ and $\mathrm{rank}(\boldsymbol{U}_i) < n_i p, i = 2, 3, \ldots, c$. Thus, $\mathrm{rank}(\boldsymbol{U}) < np$.

From Eq. (7), we have

$$\mathrm{rank}\left(\begin{bmatrix} \boldsymbol{H} \\ \boldsymbol{H}\boldsymbol{A}^* \\ \vdots \\ \boldsymbol{H}\boldsymbol{A}^{*np-1} \end{bmatrix}\right) = \mathrm{rank}(\boldsymbol{U}) < np,$$

so $(\boldsymbol{A}^*, \boldsymbol{H})$ is not observable if the graph is not connected.

Remark 1 We will illustrate Theorem 1 by introducing an example in Fig. 3. It can be seen that the topology of Case 1 is connected, whereas the topology of Case 2 is not. Intuitively, the sensing nodes in Case 2 may never receive information from the anchor nodes, so no estimator would be effective for estimating the sensing nodes. This explains why Theorem 1 requires the graph topology be connected.

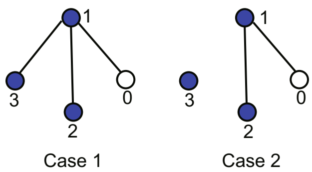

Fig. 3 Intuitive explanation for Theorem 1

Notice that in the optimal estimator (6), the Kalman filter gain $\boldsymbol{K}$ is not always a block diagonal matrix. This indicates that the state estimation at a certain node requires the information of the others even if they are not neighbors. In other words, for optimal estimation, all-to-all communications are necessary. Hence, a feasible distributed estimator should be invented to meet the need for state estimation in large-scale networks.

4 Distributed suboptimal estimators

In this section, we propose a suboptimal estimator for distributed estimation of the relative sensing network. To run this estimator, every node in the network uses its local measurements and information that it receives from its neighbors. We analyze the convergence of the distributed estimator under LMI and topology conditions rigorously.

4.1 Design of estimators

To meet the distributed nature, the Kalman similar estimator at every node in the network is designed with the following form:

$$\begin{cases} \bar{\boldsymbol{x}}_i &= A\hat{\boldsymbol{x}}_i^-, \\ \hat{\boldsymbol{x}}_i &= \bar{\boldsymbol{x}}_i + \boldsymbol{K}_i(\boldsymbol{Y}_i - \boldsymbol{H}_i\bar{\boldsymbol{X}}), \end{cases} \tag{8}$$

where $\boldsymbol{x}_i$ is the state at node i and $\bar{X}$ is defined as in the preceding section, in which the states of others are achieved through communications. The observation matrix $\boldsymbol{H}_i \in \mathbb{R}^{q_i \times np}$ is organized as $\boldsymbol{H}_i = [\boldsymbol{H}_{i1}, \boldsymbol{H}_{i2}, \ldots, \boldsymbol{H}_{in}]$ with $\boldsymbol{H}_{ij} \in \mathbb{R}^{q_i \times p}$. Note that $\boldsymbol{H}_{ij}$ is different from $\boldsymbol{h}_{l,k}$ in the last section and is actually a column of blocked elements $\boldsymbol{h}_{l,k}$ with some labels l. In addition, there is $\boldsymbol{K}_i \in \mathbb{R}^{p \times q_i}$.

Because when $j \notin N_i$ or $j \neq i$, $\boldsymbol{H}_{ij} = \boldsymbol{0}$, the states of non-neighbor nodes are not required in Eq. (8). Then, $\boldsymbol{H}_i\bar{\boldsymbol{X}} = \sum_{j=1}^n \boldsymbol{H}_{ij}\bar{\boldsymbol{x}}_j = \sum_{j\in N_i, j\neq i} \boldsymbol{H}_{ij}\bar{\boldsymbol{x}}_j + \boldsymbol{H}_{ii}\bar{\boldsymbol{x}}_i$. For brevity, hereinafter we use $\sum_j$ as a shortened form of $\sum_{j=1}^n$.

To derive the optimal $\boldsymbol{K}_i$, subtract $\boldsymbol{x}_i$ from both sides of the two expressions in Eq. (8), and we have

$$\begin{aligned} \bar{\boldsymbol{\eta}}_i &= \boldsymbol{A}\boldsymbol{\eta}_i^- - \boldsymbol{w}_i, \\ \boldsymbol{\eta}_i &= \bar{\boldsymbol{x}}_i - \boldsymbol{x}_i + \boldsymbol{K}_i(Y_i - \boldsymbol{H}_i\bar{\boldsymbol{X}}) \\ &= \bar{\boldsymbol{\eta}}_i - \boldsymbol{K}_i\sum_j \boldsymbol{H}_{ij}\bar{\boldsymbol{\eta}}_j + \boldsymbol{K}_i\boldsymbol{v}_i. \end{aligned} \tag{9}$$

Denote $\bar{\boldsymbol{P}}_{i,j} = E[\bar{\boldsymbol{\eta}}_i\bar{\boldsymbol{\eta}}_j^{\mathrm{T}}]$ and $\boldsymbol{P}_{i,j} = E[\boldsymbol{\eta}_i\boldsymbol{\eta}_j^{\mathrm{T}}]$ as the edge covariances. Consequently, $\bar{\boldsymbol{P}}_{i,j} = \bar{\boldsymbol{P}}_{j,i}^{\mathrm{T}}$ and $\boldsymbol{P}_{i,j} = \boldsymbol{P}_{j,i}^{\mathrm{T}}$. According to the definitions in Eq. (5), we also have $\bar{\boldsymbol{P}}_i = \bar{\boldsymbol{P}}_{i,i}$ and $\boldsymbol{P}_i = \boldsymbol{P}_{i,i}$. Therefore,

$$\begin{aligned} \bar{\boldsymbol{P}}_i =& \boldsymbol{A}\boldsymbol{P}_i^-\boldsymbol{A}^{\mathrm{T}} + \boldsymbol{Q}_i, \\ \boldsymbol{P}_i =& \bar{\boldsymbol{P}}_i - \sum_j \bar{\boldsymbol{P}}_{j,i}^{\mathrm{T}}\boldsymbol{H}_{ij}^{\mathrm{T}}\boldsymbol{K}_i^{\mathrm{T}} - \sum_j \boldsymbol{K}_i\boldsymbol{H}_{ij}\bar{\boldsymbol{P}}_{j,i} \\ &+ \sum_j\sum_k \boldsymbol{K}_i\boldsymbol{H}_{ij}\bar{\boldsymbol{P}}_{j,k}\boldsymbol{H}_{ik}^{\mathrm{T}}\boldsymbol{K}_i^{\mathrm{T}} + \boldsymbol{K}_i\boldsymbol{R}_i\boldsymbol{K}_i^{\mathrm{T}}. \end{aligned} \tag{10}$$

To calculate $\boldsymbol{K}_i$, let $\frac{\partial \mathrm{tr}(\boldsymbol{P}_i)}{\partial \boldsymbol{K}_i} = \boldsymbol{0}$. From matrix calculus, for any two matrices $\boldsymbol{A}$ and $\boldsymbol{X}$, the following hold:

$$\frac{\partial \mathrm{tr}(\boldsymbol{A}\boldsymbol{X})}{\partial \boldsymbol{X}} = \boldsymbol{A}^{\mathrm{T}} \text{ and } \frac{\partial \mathrm{tr}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{X})}{\partial \boldsymbol{X}} = \boldsymbol{X}(\boldsymbol{A} + \boldsymbol{A}^{\mathrm{T}}).$$

Then we have

$$\begin{aligned}&-2\textstyle\sum_j \bar{\boldsymbol{P}}_{j,i}^{\mathrm{T}}\boldsymbol{H}_{ij}^{\mathrm{T}}+2\boldsymbol{K}_i\boldsymbol{R}_i\\&+2\boldsymbol{K}_i\textstyle\sum_j\sum_k \boldsymbol{H}_{ij}\bar{\boldsymbol{P}}_{j,k}\boldsymbol{H}_{ik}^{\mathrm{T}}=\mathbf{0}.\end{aligned}\tag{11}$$

Hence,

$$\boldsymbol{K}_i=(\sum_j \bar{\boldsymbol{P}}_{j,i}^{\mathrm{T}}\boldsymbol{H}_{ij}^{\mathrm{T}})(\boldsymbol{R}_i+\sum_j\sum_k \boldsymbol{H}_{ij}\bar{\boldsymbol{P}}_{j,k}\boldsymbol{H}_{ik}^{\mathrm{T}})^{-1},$$

where

$$\begin{aligned}\bar{\boldsymbol{P}}_{i,j}=&\boldsymbol{A}\boldsymbol{P}_{i,j}^{-}\boldsymbol{A}^{\mathrm{T}},\\ \boldsymbol{P}_{i,j}=&\bar{\boldsymbol{P}}_{i,j}-\sum_l \bar{\boldsymbol{P}}_{l,i}^{\mathrm{T}}\boldsymbol{H}_{jl}^{\mathrm{T}}\boldsymbol{K}_j^{\mathrm{T}}-\sum_k \boldsymbol{K}_i\boldsymbol{H}_{ik}\bar{\boldsymbol{P}}_{k,i}\\&+\sum_k\sum_l \boldsymbol{K}_i\boldsymbol{H}_{ik}\bar{\boldsymbol{P}}_{k,l}\boldsymbol{H}_{jl}^{\mathrm{T}}\boldsymbol{K}_j^{\mathrm{T}}.\end{aligned}$$

In estimations, $\bar{\boldsymbol{P}}_j$ is achieved through communications with neighbor nodes. Note that in the above equations, l is the neighbor of j and k is the neighbor of i.

Therefore, to estimate all edge covariances $\bar{\boldsymbol{P}}_{i,j}$ and $\boldsymbol{P}_{i,j}$, all-to-all communications are still required, which is not feasible in applications. We provide below a suboptimal distributed estimator that requires only local communications.

Suboptimal estimator: In a relative sensing network, sensing nodes obtain their global information iteratively from the anchor nodes through communications. The shared information also helps improve the precision of estimations and makes the estimates appear to be correlated. However, the states of all nodes are actually independent according to their dynamic models. We therefore surmise that the correlation is weak and has little effect on the performance of estimations. Assuming $\bar{\boldsymbol{P}}_{i,j}=\boldsymbol{P}_{i,j}=\mathbf{0}$ for all $i\neq j$ in Eq. (10), then

$$\begin{aligned}\boldsymbol{P}_i=&\bar{\boldsymbol{P}}_i-\boldsymbol{K}_i\boldsymbol{H}_{ii}\bar{\boldsymbol{P}}_i-\bar{\boldsymbol{P}}_i\boldsymbol{H}_{ii}^{\mathrm{T}}\boldsymbol{K}_i^{\mathrm{T}}+\boldsymbol{K}_i\boldsymbol{R}_i\boldsymbol{K}_i^{\mathrm{T}}\\&+\sum_j \boldsymbol{K}_i\boldsymbol{H}_{ij}\bar{\boldsymbol{P}}_j\boldsymbol{H}_{ij}^{\mathrm{T}}\boldsymbol{K}_i^{\mathrm{T}}.\end{aligned}$$

Because $\bar{\boldsymbol{P}}_{i,j}=\boldsymbol{P}_{i,j}=\mathbf{0}$ and according to Eq. (11), we have $\boldsymbol{K}_i\boldsymbol{R}_i\boldsymbol{K}_i^{\mathrm{T}}+\sum_j \boldsymbol{K}_i\boldsymbol{H}_{ij}\bar{\boldsymbol{P}}_j\boldsymbol{H}_{ij}^{\mathrm{T}}\boldsymbol{K}_i^{\mathrm{T}}=\bar{\boldsymbol{P}}_i\boldsymbol{H}_{ii}^{\mathrm{T}}\boldsymbol{K}_i^{\mathrm{T}}$. Then a fully distributed estimator is realized as follows:

$$\begin{cases}\bar{\boldsymbol{x}}_i=\boldsymbol{A}\hat{\boldsymbol{x}}_i^{-},\\ \hat{\boldsymbol{x}}_i=\bar{\boldsymbol{x}}_i+\boldsymbol{K}_i(\boldsymbol{Y}_i-\boldsymbol{H}_i\bar{\boldsymbol{X}}),\\ \boldsymbol{K}_i=\bar{\boldsymbol{P}}_i\boldsymbol{H}_{ii}^{\mathrm{T}}(\boldsymbol{R}_i+\sum_j \boldsymbol{H}_{ij}\bar{\boldsymbol{P}}_j\boldsymbol{H}_{ij}^{\mathrm{T}})^{-1},\\ \bar{\boldsymbol{P}}_i=\boldsymbol{A}\boldsymbol{P}_i^{-}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{Q}_i,\\ \boldsymbol{P}_i=(\boldsymbol{I}-\boldsymbol{K}_i\boldsymbol{H}_{ii})\bar{\boldsymbol{P}}_i,\end{cases}\tag{12}$$

where $\bar{\boldsymbol{x}}_j$ and $\bar{\boldsymbol{P}}_j$ are achieved through communications between neighbors.

4.2 Convergence analysis

In this subsection, we focus on the analysis of convergence of the suboptimal estimator. Because $\boldsymbol{P}_i\neq E[\boldsymbol{\eta}_i\boldsymbol{\eta}_i^{\mathrm{T}}]$, the convergence of $\boldsymbol{P}_i$ is not enough to guarantee the convergence of $\boldsymbol{\eta}_i$. Hence, we want to present the process in two steps:

1. Prove that $\boldsymbol{P}_i$ is convergent under some prescribed conditions. As long as $\boldsymbol{P}_i$ converges, according to the relationship of $\boldsymbol{P}_i$ and $\boldsymbol{K}_i$, we know that $\boldsymbol{K}_i$ is also convergent.

2. Using the conclusion in the first step, we aim to prove that $\boldsymbol{\eta}_i$ will finally converge if the network topology and system parameters are under certain conditions.

Step 1: convergence of $\boldsymbol{P}_i$. To state the proof more clearly, define

$$\begin{aligned}\Phi_i(P,\boldsymbol{K}_i)=&(\boldsymbol{I}-\boldsymbol{K}_i\boldsymbol{H}_{ii})\bar{\boldsymbol{P}}_i(\boldsymbol{I}-\boldsymbol{K}_i\boldsymbol{H}_{ii})^{\mathrm{T}}\\&+\boldsymbol{K}_i\boldsymbol{R}_i\boldsymbol{K}_i^{\mathrm{T}}+\sum_{j,j\neq i}\boldsymbol{K}_i\boldsymbol{H}_{ij}\bar{\boldsymbol{P}}_j\boldsymbol{H}_{ij}^{\mathrm{T}}\boldsymbol{K}_i^{\mathrm{T}},\end{aligned}\tag{13}$$

where P denotes the set of matrices $\{\boldsymbol{P}_1,\boldsymbol{P}_2,\ldots,\boldsymbol{P}_n\}$, $\bar{\boldsymbol{P}}_i=\boldsymbol{A}\boldsymbol{P}_i^{-}\boldsymbol{A}^{\mathrm{T}}+\boldsymbol{Q}_i$, and $\boldsymbol{K}_i$ is the Kalman matrix. Because $\bar{\boldsymbol{P}}_i$ is positive semi-definite, it can be verified that $\Phi_i(\cdot,\cdot)\geq 0$.

Then define

$$\Phi_i^*(P)=\Phi_i(P,\phi_i(P)),\tag{14}$$

where $\phi_i(P)=\bar{\boldsymbol{P}}_i\boldsymbol{H}_{ii}^{\mathrm{T}}(\boldsymbol{R}_i+\sum_j \boldsymbol{H}_{ij}\bar{\boldsymbol{P}}_j\boldsymbol{H}_{ij}^{\mathrm{T}})^{-1}$ and we have $\boldsymbol{P}_i(k+1)=\Phi_i^*(P(k))$. Notice that $\Phi_i^*(P)=\min_{\boldsymbol{K}_i}\{\Phi_i(P,\boldsymbol{K}_i)\}$ because $\frac{\partial\mathrm{tr}(\boldsymbol{P}_i)}{\partial\boldsymbol{K}_i}=\mathbf{0}$.

The convergence analysis of $\Phi_i^*(P)$ starts from its monotonicity discussion. That is, for any i, if $\boldsymbol{X}_i\leq\boldsymbol{Y}_i$, we have

$$\Phi_i^*(\boldsymbol{X})\leq\Phi_i(\boldsymbol{X},\phi_i(\boldsymbol{Y}))\leq\Phi_i(\boldsymbol{Y},\phi_i(\boldsymbol{Y}))=\Phi_i^*(\boldsymbol{Y}).\tag{15}$$

We then show that $\Phi_i^*(P)$ is bounded for all i's under certain conditions, but before that, the following lemma is introduced:

Lemma 2 Define the operator

$$\begin{aligned}f_i(P)=&(\boldsymbol{I}-\boldsymbol{K}_i\boldsymbol{H}_{ii})\boldsymbol{A}\boldsymbol{P}_i\boldsymbol{A}^{\mathrm{T}}(\boldsymbol{I}-\boldsymbol{K}_i\boldsymbol{H}_{ii})^{\mathrm{T}}\\&+\sum_{j,j\neq i}\boldsymbol{K}_i\boldsymbol{H}_{ij}\boldsymbol{A}\boldsymbol{P}_j\boldsymbol{A}^{\mathrm{T}}\boldsymbol{H}_{ij}^{\mathrm{T}}\boldsymbol{K}_i^{\mathrm{T}},\end{aligned}\tag{16}$$

with

$$f_i^{k+1}(P) =(\boldsymbol{I}-\boldsymbol{K}_i\boldsymbol{H}_{ii})\boldsymbol{A}f_i^k(P)\boldsymbol{A}^{\rm T}(\boldsymbol{I}-\boldsymbol{K}_i\boldsymbol{H}_{ii})^{\rm T} + \sum_{j,j\neq i}\boldsymbol{K}_i\boldsymbol{H}_{ij}\boldsymbol{A}f_j^k(P)\boldsymbol{A}^{\rm T}\boldsymbol{H}_{ij}^{\rm T}\boldsymbol{K}_i^{\rm T}. \tag{17}$$

If there exists a set of positive definite matrices P^* such that $\boldsymbol{P}_i^* > f_i(P^*)$ for all $i = 1, 2, \ldots, n$, then: (a) for all positive semi-definite matrices $\boldsymbol{P}_i, i \in \{1, 2, \ldots, n\}$, $\lim_{k\to\infty} f_i^k(P) = 0$ holds for all $i = 1, 2, \ldots, n$; (b) for all positive semi-definite matrices $\boldsymbol{U}_i$ and $\boldsymbol{P}_i(0)$, the sequences $\boldsymbol{P}_i(k+1)$ satisfying $\boldsymbol{P}_i(k+1) \leq f_i(P(k)) + \boldsymbol{U}_i$ are bounded.

Refer to the Appendix for further proof of Lemma 2.

Now we come to the convergence proof of $\boldsymbol{P}_i$.

Theorem 2 $\boldsymbol{P}_i$ in the distributed algorithm is convergent if there exist gains $\boldsymbol{K}_i$ and positive definite matrices $\boldsymbol{P}_i^*, i = 1, 2, \ldots, n$, such that

$$\boldsymbol{P}_i^* >(\boldsymbol{I}-\boldsymbol{K}_i\boldsymbol{H}_{ii})\boldsymbol{A}\boldsymbol{P}_i^*\boldsymbol{A}^{\rm T}(\boldsymbol{I}-\boldsymbol{K}_i\boldsymbol{H}_{ii})^{\rm T} + \sum_{j,j\neq i}\boldsymbol{K}_i\boldsymbol{H}_{ij}\boldsymbol{A}\boldsymbol{P}_j^*\boldsymbol{A}^{\rm T}\boldsymbol{H}_{ij}^{\rm T}\boldsymbol{K}_i^{\rm T}. \tag{18}$$

Also, $\boldsymbol{K}_i$ is convergent.

Proof Eq. (18) is equivalent to $\boldsymbol{P}_i^* > f_i(P^*)|_{\boldsymbol{K}_i=\boldsymbol{K}_i^*}$, $i = 1, 2, \ldots, n$. Then, according to Lemma 2, we have

$$\begin{aligned}\varPhi_i^*(P) &= \varPhi_i^*(P, \phi_i(P)) \leq \varPhi_i(P, \boldsymbol{K}_i^*)\\ &= f_i(P)|_{\boldsymbol{K}_i=\boldsymbol{K}_i^*} + (\boldsymbol{I}-\boldsymbol{K}_i^*\boldsymbol{H}_{ii})\boldsymbol{Q}_i(\boldsymbol{I}-\boldsymbol{K}_i^*\boldsymbol{H}_{ii})^{\rm T}\\ &\quad + \sum_{j,j\neq i}\boldsymbol{K}_i^*\boldsymbol{H}_{ij}\boldsymbol{Q}_j\boldsymbol{H}_{ij}^{\rm T}\boldsymbol{K}_i^{*\rm T} + \boldsymbol{K}_i^*\boldsymbol{R}_i\boldsymbol{K}_i^{*\rm T}.\end{aligned} \tag{19}$$

Let $\boldsymbol{U}_i = (\boldsymbol{I}-\boldsymbol{K}_i^*\boldsymbol{H}_{ii})\boldsymbol{Q}_i(\boldsymbol{I}-\boldsymbol{K}_i^*\boldsymbol{H}_{ii})^{\rm T} + \sum_{j,j\neq i}(\boldsymbol{K}_i^*\boldsymbol{H}_{ij})\boldsymbol{Q}_j(\boldsymbol{K}_i^*\boldsymbol{H}_{ij})^{\rm T} + \boldsymbol{K}_i^*\boldsymbol{R}_i\boldsymbol{K}_i^{*\rm T}$. From Lemma 2b, we conclude that $\varPhi_i^*(P)$ is bounded.

Now we are able to prove the convergence of $\varPhi_i^*(P)$ under the above conditions.

First, consider the case with zero initial conditions, i.e., $\boldsymbol{P}_i(0) = E[\boldsymbol{\eta}_i(0)\boldsymbol{\eta}_i(0)^{\rm T}] = \boldsymbol{0}$ for all i's. Because $\boldsymbol{P}_i(0) = \boldsymbol{0}$, then for all $j \in N_i$, it can be easily verified that $\bar{\boldsymbol{P}}_{i,j}(1) = E(\bar{\boldsymbol{\eta}}_i(1)\bar{\boldsymbol{\eta}}_j(1)) = \boldsymbol{0}$. Therefore, $\boldsymbol{P}_{i,j}(1) = \boldsymbol{0}$, although assuming $\bar{\boldsymbol{P}}_{ij} = \boldsymbol{P}_{i,j} = \boldsymbol{0}$ for all $i \neq j$, $\boldsymbol{P}_i(1)$ still represents the real estimation error covariance matrix. Thus, $\boldsymbol{P}_i(1) = E[\boldsymbol{\eta}_i(1)\boldsymbol{\eta}_i(1)^{\rm T}] \geq \boldsymbol{0} = \boldsymbol{P}_i(0)$. According to the monotonicity of $\varPhi_i^*(P)$ in Eq. (15), we then have $\boldsymbol{P}_i(k+1) \geq \boldsymbol{P}_i(k)$ for all $k > 0$. Also, we have that $\varPhi_i^*(P)$ is upper bounded from Eq. (19). Therefore, $\varPhi_i^*(P)$ converges for all i's. We thus assume that the convergent value of $\boldsymbol{P}_i(k)$ under the condition is $\boldsymbol{D}_i$ with $\varPhi_i^*(\boldsymbol{D}) = \boldsymbol{D}_i$.

Second, consider the case with initial conditions $\boldsymbol{P}_i(0) \geq \boldsymbol{D}_i$ for all i's. Because $\varPhi_i^*(\boldsymbol{X})$ is monotonic, we have $\varPhi_i^*(P(0)) \geq \varPhi_i^*(\boldsymbol{D})$ and $\lim_{k\to\infty}\boldsymbol{P}_i(k) \geq \boldsymbol{D}_i$. Also, because

$$\begin{aligned}\boldsymbol{P}_i(k) - \boldsymbol{D}_i &= \varPhi_i^*(P(k-1)) - \varPhi_i^*(\boldsymbol{D})\\ &\leq \varPhi_i(P(k-1), \phi_i(\boldsymbol{D})) - \varPhi_i(\boldsymbol{D}, \phi_i(\boldsymbol{D}))\\ &= f_i(P(k-1)) - f_i(\boldsymbol{D}),\end{aligned}$$

let $P(k-1) - \boldsymbol{D}$ denote the set of positive semi-definite matrices $\{\boldsymbol{P}_1(k) - \boldsymbol{D}_1, \boldsymbol{P}_2(k) - \boldsymbol{D}_2, \ldots, \boldsymbol{P}_n(k) - \boldsymbol{D}_n\}$. From Lemma 2 we have $\lim_{k\to\infty}\boldsymbol{P}_i(k) - \boldsymbol{D}_i \leq \lim_{k\to\infty} f_i(P(k-1) - \boldsymbol{D}) = \boldsymbol{0}$. Therefore, $\lim_{k\to\infty}\boldsymbol{P}_i(k) - \boldsymbol{D}_i = \boldsymbol{0}$ and $\boldsymbol{P}_i(k)$ converges for all i's.

Finally, consider the case with initial conditions $\boldsymbol{P}_i(0) \geq 0$ for any i and neither of the above cases is satisfied. Assume $0 \leq \boldsymbol{P}_i(0) \leq \delta\boldsymbol{D}_i$ for all i's, where $\delta > 1$ is a certain constant. We then have $\boldsymbol{D}_i = \lim_{k\to\infty}\varPhi_i^*(\boldsymbol{0}) \leq \lim_{k\to\infty}\varPhi_i^*(P(0)) \leq \lim_{k\to\infty}\varPhi_i^*(\delta\boldsymbol{D}) = \boldsymbol{D}_i$, i.e., $\lim_{k\to\infty}\boldsymbol{P}_i(k) = \boldsymbol{D}_i$, and the convergence of $\boldsymbol{P}_i$ is proved.

According to the relationship between $\boldsymbol{P}_i$ and $\boldsymbol{K}_i$ in Eq. (12), we know that $\boldsymbol{K}_i$ is also convergent. Define $\lim_{k\to\infty}\boldsymbol{K}_i = \boldsymbol{K}_i^*$.

Remark 2 Now we come to the condition in Lemma 2 that makes $\boldsymbol{P}_i$ converge. The condition in Lemma 2 is composed of n LMIs. The conditions are the solutions to these n LMIs. After multiplying both left and right sides of Eq. (18) by $\boldsymbol{P}_i^{*-1}$ and defining $\boldsymbol{\varTheta}_i = \boldsymbol{P}_i^{*-1}, \boldsymbol{\varUpsilon}_i = \boldsymbol{P}_i^{*-1}\boldsymbol{K}_i$, we have the following convergence condition by recursively using the Schur complement lemma. If there exist gains $\boldsymbol{\varUpsilon}_i$ and positive-definite matrices $\boldsymbol{\varTheta}_i^*, \boldsymbol{\varTheta}_j^*, j \in N_i$, such that matrix

$$\begin{bmatrix} \boldsymbol{\varTheta}_i^*, & (\boldsymbol{\varTheta}_i^* - \boldsymbol{\varUpsilon}_i\boldsymbol{H}_{ii})\boldsymbol{A}, & \boldsymbol{\varUpsilon}_i\boldsymbol{H}_{ij_1}\boldsymbol{A}, & \cdots, & \boldsymbol{\varUpsilon}_i\boldsymbol{H}_{ij_{|N_i|}}\boldsymbol{A}\\ * & \boldsymbol{\varTheta}_i^* & & & \\ * & & \boldsymbol{\varTheta}_{j_1}^* & & \\ \vdots & & & \ddots & \\ * & & & & \boldsymbol{\varTheta}_{j_{|N_i|}}^* \end{bmatrix}$$

is positive definite, then $\boldsymbol{P}_i$ is convergent.

Step 2: convergence of $\boldsymbol{\eta}_i$. In this step, we want to prove that the estimation error $\boldsymbol{\eta}_i$ is also convergent. Referring to Eq. (9), we have

$$\begin{aligned}\boldsymbol{\eta}_i =&\boldsymbol{A}\boldsymbol{\eta}_i^- - \boldsymbol{K}_i\sum_j \boldsymbol{H}_{ij}\boldsymbol{A}\boldsymbol{\eta}_j^- \\ &+ (\boldsymbol{K}_i\boldsymbol{v}_i + \boldsymbol{K}_i\sum_j \boldsymbol{H}_{ij}\boldsymbol{w}_j - \boldsymbol{w}_i) \\ =&(\boldsymbol{I} - \boldsymbol{K}_i\boldsymbol{H}_{ii})\boldsymbol{A}\boldsymbol{\eta}_i^- - \sum_{j,j\neq i}\boldsymbol{K}_i\boldsymbol{H}_{ij}\boldsymbol{A}\boldsymbol{\eta}_j^- + \boldsymbol{u}_i,\end{aligned} \tag{20}$$

where $\boldsymbol{u}_i = \boldsymbol{K}_i\boldsymbol{v}_i + \boldsymbol{K}_i\sum_j \boldsymbol{H}_{ij}\boldsymbol{w}_j - \boldsymbol{w}_i$ can be treated as the sum of white noises. Then we can aggregate all the $\boldsymbol{\eta}_i$ to form the error equation of the system:

$$\boldsymbol{\eta} = \boldsymbol{S}_e\boldsymbol{\eta}^- + \boldsymbol{u}. \tag{21}$$

$\boldsymbol{\eta} = [\boldsymbol{\eta}_1^{\mathrm{T}}, \boldsymbol{\eta}_2^{\mathrm{T}}, \ldots, \boldsymbol{\eta}_n^{\mathrm{T}}]^{\mathrm{T}}$, $u = [\boldsymbol{u}_1^{\mathrm{T}}, \boldsymbol{u}_2^{\mathrm{T}}, \ldots, \boldsymbol{u}_n^{\mathrm{T}}]^{\mathrm{T}}$, and $\boldsymbol{S}_e$ is a block matrix of the same structure as the Laplacian matrix $\boldsymbol{L}$. Define the $(i^{\mathrm{th}}, j^{\mathrm{th}})$ block of $\boldsymbol{S}_e$ as $\boldsymbol{S}_e(i,j)$. Then we have

$$\boldsymbol{S}_e(i,j) := \begin{cases}(\boldsymbol{I} - \boldsymbol{K}_i\boldsymbol{H}_{ii})\boldsymbol{A}, & i = j, \\ -\boldsymbol{K}_i\boldsymbol{H}_{ij}\boldsymbol{A}, & i \neq j, \boldsymbol{L}(i,j) \neq 0, \\ 0, & \text{otherwise.}\end{cases} \tag{22}$$

As mentioned above, $\lim_{k\to\infty}\boldsymbol{K}_i = \boldsymbol{K}_i^*$. Thus, we conclude that $\lim_{k\to\infty}\boldsymbol{S}_e = \boldsymbol{S}_e^*$. Instead of directly proving the convergence of $\boldsymbol{\eta}$, we aim to prove the convergence of $\boldsymbol{\eta}$'s covariance matrix because it is a random variable.

Because $\boldsymbol{u}$ combines white noises at time k and $\boldsymbol{\eta}^-$ is related to white noises at time $k-1$, then $E[\boldsymbol{\eta}^-\boldsymbol{u}^{\mathrm{T}}] = E[\boldsymbol{u}\boldsymbol{\eta}^{-\mathrm{T}}] = \boldsymbol{0}$. Denote $\boldsymbol{P}_{\mathrm{dis}} = E[\boldsymbol{\eta}\boldsymbol{\eta}^{\mathrm{T}}]$, $\boldsymbol{P}_{\boldsymbol{u}} = E[\boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}]$. Then $E[\boldsymbol{\eta}\boldsymbol{\eta}^{\mathrm{T}}] = E[\boldsymbol{S}_e\boldsymbol{\eta}^-\boldsymbol{\eta}^{-\mathrm{T}}\boldsymbol{S}_e^{\mathrm{T}} + \boldsymbol{S}_e\boldsymbol{\eta}^-\boldsymbol{u}^{\mathrm{T}} + \boldsymbol{u}\boldsymbol{\eta}^{-\mathrm{T}}\boldsymbol{S}_e^{\mathrm{T}} + \boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}]$, which can be denoted as

$$\boldsymbol{P}_{\mathrm{dis}} = \boldsymbol{S}_e\boldsymbol{P}_{\mathrm{dis}}^-\boldsymbol{S}_e^{\mathrm{T}} + \boldsymbol{P}_{\boldsymbol{u}}. \tag{23}$$

Before the analysis of Eq. (23), we introduce a definition and a lemma.

Definition 1 The matrix $\boldsymbol{\Gamma}$ is Schur stable or asymptotically stable, if and only if

$$|\lambda_i| < 1, \forall i = 1, 2, \ldots, n,$$

where λ_i's are the eigenvalues of matrix $\boldsymbol{\Gamma}$.

Lemma 3 Consider the linear discrete-time system $\boldsymbol{x}(k+1) = \boldsymbol{F}\boldsymbol{x}(k)$; the following conditions are equivalent:

(1) The matrix $\boldsymbol{\Gamma}$ is Schur stable;

(2) Given any matrix $\boldsymbol{Q} = \boldsymbol{Q}^{\mathrm{T}} > 0$, there exists a unique positive definite matrix $\boldsymbol{P} = \boldsymbol{P}^{\mathrm{T}}$ satisfying the equation

$$\boldsymbol{\Gamma}\boldsymbol{P}\boldsymbol{\Gamma}^{\mathrm{T}} - \boldsymbol{P} = -\boldsymbol{Q}. \tag{24}$$

The detailed proof of Lemma 3 is presented in the Appendix.

Here comes the theorem for convergence of $\boldsymbol{P}_{\mathrm{dis}}$.

Theorem 3 Suppose Eq. (18) is satisfied. Then the real estimation error covariance matrix $\boldsymbol{P}_{\mathrm{dis}}$ will finally converge if and only if $\boldsymbol{S}_e^*$ is Schur stable.

Proof After defining $\boldsymbol{\Delta} = \boldsymbol{S}_e - \boldsymbol{S}_e^*$, Eq. (23) can be transformed into

$$\begin{aligned}\boldsymbol{P}_{\mathrm{dis}} =&(\boldsymbol{S}_e^* + \boldsymbol{\Delta})\boldsymbol{P}_{\mathrm{dis}}^-(\boldsymbol{S}_e^* + \boldsymbol{\Delta})^{\mathrm{T}} + \boldsymbol{P}_{\boldsymbol{u}} \\ =&\boldsymbol{S}_e^*\boldsymbol{P}_{\mathrm{dis}}^-\boldsymbol{S}_e^{*\mathrm{T}} + \boldsymbol{P}_{\boldsymbol{u}}^{\boldsymbol{\Delta}},\end{aligned} \tag{25}$$

where $\boldsymbol{P}_{\boldsymbol{u}}^{\boldsymbol{\Delta}} = \boldsymbol{P}_{\boldsymbol{u}} + \boldsymbol{S}_e^*\boldsymbol{P}_{\mathrm{dis}}^-\boldsymbol{\Delta}^{\mathrm{T}} + \boldsymbol{\Delta}\boldsymbol{P}_{\mathrm{dis}}^-\boldsymbol{S}_e^{*\mathrm{T}} + \boldsymbol{\Delta}\boldsymbol{\Delta}^{\mathrm{T}}$. Next we focus on matrix $\boldsymbol{P}_{\boldsymbol{u}}^{\boldsymbol{\Delta}}$. Facing the fact that $\lim_{k\to\infty}\boldsymbol{K}_i = \boldsymbol{K}_i^*$, then we have

$$\begin{aligned}&\lim_{k\to\infty}\boldsymbol{P}_{\boldsymbol{u}} = \boldsymbol{P}_{\boldsymbol{u}}^* = E[\boldsymbol{u}^*\boldsymbol{u}^{*\mathrm{T}}] \\ &= \sum_{i=1}^{n}(\boldsymbol{K}_i^*\boldsymbol{R}_i\boldsymbol{K}_i^{*\mathrm{T}} + \boldsymbol{K}_i^*\sum_j(\boldsymbol{H}_{ij}\boldsymbol{Q}_j\boldsymbol{H}_{ij}^{\mathrm{T}})\boldsymbol{K}_i^{*\mathrm{T}} + \boldsymbol{Q}_i).\end{aligned} \tag{26}$$

Thus, $\boldsymbol{P}_{\boldsymbol{u}}^*$ is a symmetric positive definite matrix. Because $\lim_{k\to\infty}\boldsymbol{\Delta} = \boldsymbol{0}$, then $\lim_{k\to\infty}\boldsymbol{P}_{\boldsymbol{u}}^{\boldsymbol{\Delta}} = \boldsymbol{P}_{\boldsymbol{u}}^* > 0$. According to Lemma 3 and because $\boldsymbol{S}_e^*$ is Schur stable, we conclude that $\lim_{k\to\infty}\boldsymbol{P}_{\mathrm{dis}} = \boldsymbol{P}_{\mathrm{dis}}^*$, and $\boldsymbol{P}_{\mathrm{dis}}^*$ is the unique solution of equation $\boldsymbol{P}_{\mathrm{dis}}^* = \boldsymbol{S}_e^*\boldsymbol{P}_{\mathrm{dis}}^*\boldsymbol{S}_e^{*\mathrm{T}} + \boldsymbol{P}_{\boldsymbol{u}}^*$. Then, the solution of $\boldsymbol{P}_{\mathrm{dis}}^*$ can be calculated as

$$\boldsymbol{P}_{\mathrm{dis}}^* = \sum_{i=0}^{\infty}(\boldsymbol{S}_e^*)^i\boldsymbol{P}_{\boldsymbol{u}}^*(\boldsymbol{S}_e^{*\mathrm{T}})^i.$$

Remark 3 $\boldsymbol{P}_{\mathrm{dis}}^*$ can be proved to be bounded according to

$$\begin{aligned}\boldsymbol{P}_{\mathrm{dis}}^* =&\sum_{i=0}^{\infty}(\boldsymbol{S}_e^*)^i\boldsymbol{P}_{\boldsymbol{u}}^*(\boldsymbol{S}_e^{*\mathrm{T}})^i \leq \sum_{i=0}^{\infty}(\boldsymbol{S}_e^*)^i\lambda_{\max}\boldsymbol{I}(\boldsymbol{S}_e^{*\mathrm{T}})^i \\ \leq&\lambda_{\max}\sum_{i=0}^{\infty}(\boldsymbol{S}_e^*\boldsymbol{S}_e^{*\mathrm{T}})^i = \lambda_{\max}(\boldsymbol{I} - \boldsymbol{S}_e^*\boldsymbol{S}_e^{*\mathrm{T}})^{-1}, \\ &\text{because } \boldsymbol{S}_e^*\boldsymbol{S}_e^{*\mathrm{T}} < \boldsymbol{I}.\end{aligned}$$

Remark 4 Although the stability of $\boldsymbol{S}_e^*$ could not be verified beforehand, its stability is equivalent to the convergence of $\boldsymbol{P}_{\mathrm{dis}}$. In other words, if $\boldsymbol{S}_e^*$ is not Schur stable, we can definitely conclude that the suboptimal algorithm will not be effective. Thus, it is reasonable to suppose that $\boldsymbol{S}_e^*$ is Schur stable, but the future work is to find a more specific condition to make the error system stable.

5 Simulation

In this section, the performances of the optimal and suboptimal estimators are illustrated. Without loss of generality, a localization example is considered in simulations. Two cases with different topologies are tested with 1000 groups of Monte Carlo samples.

In our simulations, six buoys floating on calm waters need to be localized. The network topologies are presented in Figs. 4 and 5, respectively. In these networks, only anchor nodes have measurements of their own absolute positions. Each node, however, can measure the relative position and receive local information from its neighbors. For comparison, we assume that the system models of two simulations are the same and are described as

$$\begin{cases} \boldsymbol{x}_i = \boldsymbol{A}\boldsymbol{x}_i^- + \boldsymbol{w}_i, \\ \boldsymbol{y}_{ij} = \boldsymbol{C}_{ij}\boldsymbol{x}_{ij} + \boldsymbol{v}_{ij}, \end{cases} \tag{27}$$

where $\boldsymbol{A} = \boldsymbol{C} = \boldsymbol{R} = \boldsymbol{I}_2$, $\boldsymbol{Q} = 0.01\boldsymbol{I}_2$, $\boldsymbol{C}_i = \boldsymbol{C}_{ij} = \boldsymbol{C}$, $\boldsymbol{R}_{ij} = \boldsymbol{R}_i$, $\boldsymbol{Q}_1 = \boldsymbol{Q}_2 = \boldsymbol{Q}_3 = 5\boldsymbol{Q}$, $\boldsymbol{Q}_4 = \boldsymbol{Q}_6 = 2\boldsymbol{Q}$, $\boldsymbol{Q}_5 = 12\boldsymbol{Q}$, $\boldsymbol{R}_1 = \boldsymbol{R}_5 = \boldsymbol{R}$, $\boldsymbol{R}_2 = \boldsymbol{R}_4 = 2\boldsymbol{R}$, and $\boldsymbol{R}_3 = \boldsymbol{R}_6 = 4\boldsymbol{R}$.

Given initial $\boldsymbol{x}_i(0) = \boldsymbol{0}, \boldsymbol{P}_i(0) = \boldsymbol{P}_{ij}(0) = \boldsymbol{I}_2, \forall i, j$, estimators are used to estimate the state. It can be verified that six LMI equations have positive-definite solutions. We calculate the estimation error covariance $\boldsymbol{P}_{\text{dis}}$ of the entire system in two examples. $\text{tr}(\boldsymbol{P}_{\text{dis}})$ is then computed and the results are presented in Figs. 6 and 7, respectively.

According to the simulations, errors in all estimators converge to steady states. Because the suboptimal estimator has ignored the edge covariance although the states of nodes are actually weakly correlated, its estimation errors are larger than those of the optimal estimator. For these two simulations, $\boldsymbol{S}_e^*$ can be verified as Schur stable.

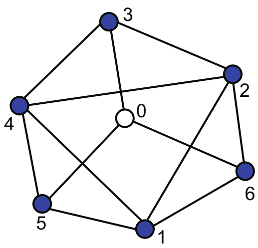

Fig. 4 Network topology of simulation 1

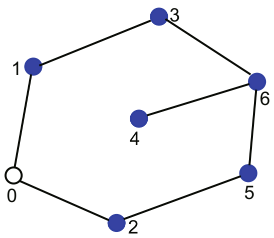

Fig. 5 Network topology of simulation 2

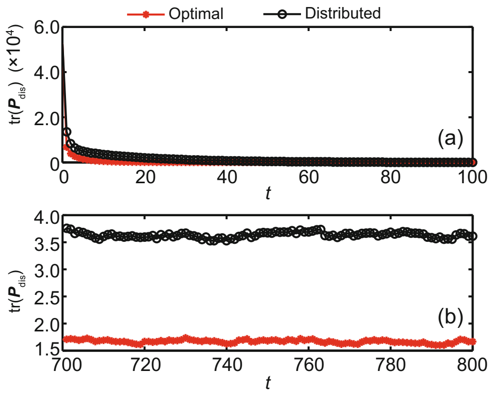

Fig. 6 Results of tr($\boldsymbol{P}_{\text{dis}}$) at each time step in simulation 1: (a) from time step 0 to time step 100; (b) from time step 700 to time step 800

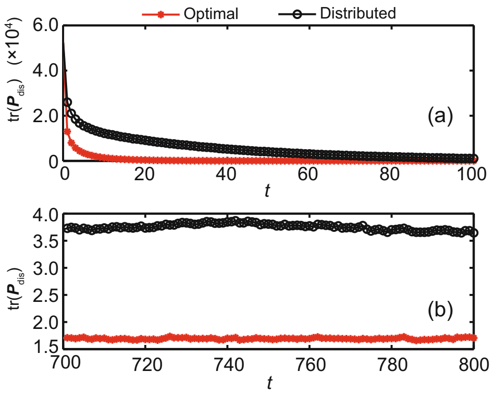

Fig. 7 Results of tr($\boldsymbol{P}_{\text{dis}}$) at each time-step in simulation 2: (a) from time step 0 to time step 100; (b) from time step 700 to time step 800

To see the estimates more clearly, we plot the trajectory evolution of the real state, centralized estimator, and the distributed estimator in Figs. 8 and 9, respectively. Although starting from different initial positions, all the estimates will finally catch up with the real state.

Moreover, the communication delay phenomenon always exists in WSNs. Although we do not consider the time-delay situation in our assumption, it is interesting to consider it in our simulation and determine its influence. We assume that each node in the first simulation would tolerate a fixed time delay of three time steps, which means that the a priori estimate received from its neighbors is $\bar{\boldsymbol{X}}(k-3)$ instead of $\bar{\boldsymbol{X}}$. We compare its performance with those under a no-delay distributed algorithm and an optimal algorithm. As shown in Fig. 10, with a time delay, the convergence rate would become slower but it is still convergent with a larger trace of the estimation error.

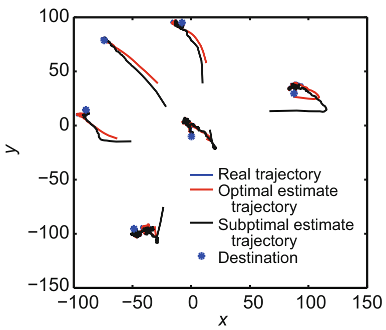

Fig. 8 Trajectory evolution of simulation 1

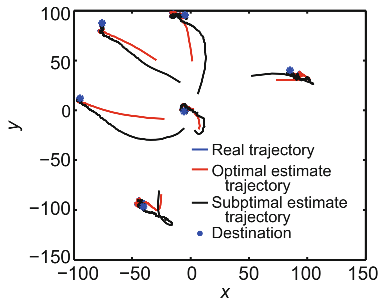

Fig. 9 Trajectory evolution of simulation 2

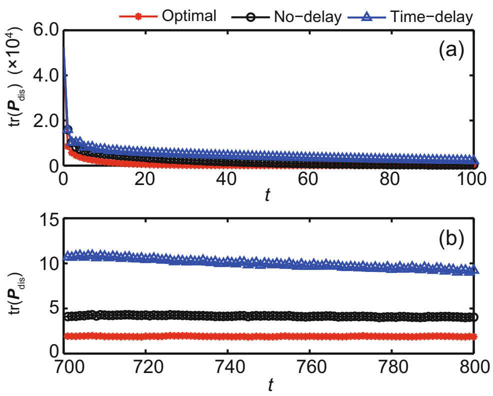

Fig. 10 Comparisons between a no-delay algorithm and a time-delay algorithm: (a) from time step 0 to time step 100; (b) from time step 700 to time step 800

6 Conclusions and future work

In this paper, we explore the problem of distributed Kalman filtering in a homogeneous relative sensing network. According to our assumptions, each node has relative state measurements between itself and its neighbors, but only anchor nodes have direct measurements of their own states. The sufficient and necessary condition for observability of the whole system is presented with detailed analysis. A centralized optimal estimator is then constructed with standard Kalman filtering. Because the optimal estimator requires all-to-all communications, which is infeasible in large-scale applications, a distributed suboptimal estimator is designed by ignoring the correlated edge covariance. The solutions of n LMIs are prepared as the condition for convergence, and the system topology and parameters also influence the stability of the estimation error. Simulation cases with different topologies are carried out to verify the performance of the proposed estimators.

In future work, the emergent task is to determine the equivalent condition for the stability of $\boldsymbol{S}_e^*$. In addition, some other distributed algorithms will be discussed, and analyzing the performance differences between them would be interesting and necessary. The effect of time-delay transmission has been simulated in this paper, and further theoretical analysis would be interesting and important. Relative sensing networks modeled by directed graphs and switching topology will also be considered.

References

Chen C, Yan J, Lu N, et al., 2015. Ubiquitous monitoring for industrial cyber-physical systems over relay-assisted wireless sensor networks. *IEEE Trans Emerg Top Comput*, 3(3):352-362. https://doi.org/10.1109/tetc.2014.2386615

Chen M, González-Valenzuela S, Cao H, et al., 2013. Enabling low bit-rate and reliable video surveillance over practical wireless sensor network. *J Supercomput*, 65(1):287-300. https://doi.org/10.1007/s11227-010-0475-2

Duan Z, Li X, 2011. Lossless linear transformation of sensor data for distributed estimation fusion. *IEEE Trans Signal Proc*, 59(1):362-372. https://doi.org/10.1109/tsp.2010.2084574

Fadel E, Gungor V, Nassef L, et al., 2015. A survey on wireless sensor networks for smart grid. *Comput Commun*, 71:22-33. https://doi.org/10.1016/j.comcom.2015.09.006

Gao Y, Li X, Song E, 2016. Robust linear estimation fusion with allowable unknown cross-covariance. *IEEE Trans Syst Man Cybern Syst*, 46(9):1314-1325. https://doi.org/10.1109/tsmc.2015.2487882

Garcia-Sanchez AJ, Garcia-Sanchez F, Garcia-Haro J, 2011. Wireless sensor network deployment for integrating video-surveillance and data-monitoring in precision agriculture over distributed crops. *Comput Electron Agric*, 75(2):288-303. https://doi.org/10.1016/j.compag.2010.12.005

Godsil C, Royle G, 2001. Algebraic Graph Theory. Springer-Verlag. https://doi.org/10.1007/978-1-4613-0163-9

Li S, Peng S, Chen W, et al., 2013. Income: practical land monitoring in precision agriculture with sensor networks. *Comput Commun*, 36(4):459-467. https://doi.org/10.1016/j.comcom.2012.10.011

Lin Z, Wang L, Han Z, et al., 2014. Distributed formation control of multi-agent systems using complex Laplacian. *IEEE Trans Autom Contr*, 59(7):1765-1777. https://doi.org/10.1109/TAC.2014.2309031

Lu S, Lin C, Lin Z, et al., 2015. Distributed kalman filter for relative sensing networks. 34th Chinese Control Conf, p.7541-7546. https://doi.org/10.1109/ChiCC.2015.7260835

Morbidi F, Mariottini G, Prattichizzo D, 2010. Observer design via immersion and invariance for vision-based leader-follower formation control. *Automatica*, 46(1):148-154. https://doi.org/10.1016/j.automatica.2009.10.016

Olfati-Saber R, 2009. Kalman-consensus filter: optimality, stability, and performance. Proc 48th IEEE Conf on Decision and Control held jointly with the 28th Chinese Control Conf, p.7036-7042. https://doi.org/10.1109/cdc.2009.5399678

Olfati-Saber R, Jalalkamali P, 2012. Coupled distributed estimation and control for mobile sensor networks. *IEEE Trans Autom Contr*, 57(10):2609-2614. https://doi.org/10.1109/tac.2012.2190184

Pasqualetti F, Carli R, Bullo F, 2012. Distributed estimation via iterative projections with application to power network monitoring. *Automatica*, 48(5):747-758. https://doi.org/10.1016/j.automatica.2012.02.025

Piovan G, Shames I, Fidan B, et al., 2013. On frame and orientation localization for relative sensing networks. *Automatica*, 49(1):206-213. https://doi.org/10.1016/j.automatica.2012.09.014

Ravazzi C, Frasca P, Tempo R, et al., 2013. Almost sure convergence of a randomized algorithm for relative localization in sensor networks. Proc 52nd IEEE Conf on Decision and Control, p.4778-4783. https://doi.org/10.1109/cdc.2013.6760638

Tubaishat M, Zhuang P, Qi Q, et al., 2009. Wireless sensor networks in intelligent transportation systems. *Wirel Commun Mob Comput*, 9(3):287-302. https://doi.org/10.1002/wcm.v9:3

Zhou Z, Fang H, Hong Y, 2013. Distributed estimation for moving target based on state-consensus strategy. *IEEE Trans Autom Contr*, 58(8):2096-2101. https://doi.org/10.1109/tac.2013.2246476

Appendix: Proof of Lemmas 2 and 3

Proof of Lemma 2

(a) According to the definitions, the following properties hold: (1) If $\boldsymbol{X}_i \leq \boldsymbol{Y}_i$ for all i's, then $f_i(\boldsymbol{X}) \leq f_i(\boldsymbol{Y})$; (2) If $\boldsymbol{X}_i = \alpha \boldsymbol{Y}_i$ for all i's, then $f_i(\boldsymbol{X}) = \alpha f_i(\boldsymbol{Y})$ and $f_i^k(\boldsymbol{X}) = \alpha f_i^k(\boldsymbol{Y})$.

Define constants $\beta_i = \min\{\alpha : \alpha \boldsymbol{P}_i^* \geq f_i(\boldsymbol{P}^*)\}$; there always exists a constant $\beta = \max_i\{\beta_i\}$ with $0 < \beta < 1$ such that $f_i(\boldsymbol{P}^*) \leq \beta \boldsymbol{P}_i^*$ for all i's. Therefore, $f_i^2(\boldsymbol{P}^*) \leq \beta f_i(\boldsymbol{P}^*) \leq \beta^2 \boldsymbol{P}_i^*$ and sequentially $f_i^k(\boldsymbol{P}^*) \leq \beta^k \boldsymbol{P}_i^*$.

Then assume $\boldsymbol{P}_i \leq \gamma \boldsymbol{P}_i^*$ for all i's with a certain constant γ. Obviously, $f_i^k(\boldsymbol{P}) \leq \gamma f_i^k(\boldsymbol{P}^*)$. We thus have $\lim_{k\to\infty} f_i^k(\boldsymbol{P}) \leq \lim_{k\to\infty} \gamma\beta^k \boldsymbol{P}_i^* = 0$.

(b) To prove the statement, we first assume that

$$\boldsymbol{P}_i(k) \leq f_i^k(\boldsymbol{P}(0)) + \sum_{l=1}^{k-1} f_i^l(\boldsymbol{U}) + \boldsymbol{U}_i, \tag{A1}$$

where $\boldsymbol{U} = \{\boldsymbol{U}_1, \boldsymbol{U}_2, \ldots, \boldsymbol{U}_n\}$. Then an inductive method is applied to verify the equation. Assume that at step k the above equation is satisfied. We substitute Eq. (A1) into $\boldsymbol{P}_i$ at step $k+1$, and have

$$\begin{aligned}
&\boldsymbol{P}_i(k+1) \leq f_i(\boldsymbol{P}(k)) + \boldsymbol{U}_i \\
=&[(\boldsymbol{I} - \boldsymbol{K}_i\boldsymbol{H}_{ii})\boldsymbol{A}]\boldsymbol{P}_i(k)[(\boldsymbol{I} - \boldsymbol{K}_i\boldsymbol{H}_{ii})\boldsymbol{A}]^{\mathrm{T}} + \boldsymbol{U}_i \\
&+ \sum_{j,j\neq i} (\boldsymbol{K}_i\boldsymbol{H}_{ij}\boldsymbol{A})\boldsymbol{P}_j(k)(\boldsymbol{K}_i\boldsymbol{H}_{ij}\boldsymbol{A})^{\mathrm{T}} \\
\leq&[\cdot](f_i^k(\boldsymbol{P}(0)) + \sum_{l=1}^{k-1} f_i^l(\boldsymbol{U}) + \boldsymbol{U}_i)[\cdot]^{\mathrm{T}} + \boldsymbol{U}_i \\
&+ \sum_{j,j\neq i} (\cdot)(f_j^k(\boldsymbol{P}(0)) + \sum_{l=1}^{k-1} f_j^l(\boldsymbol{U}) + \boldsymbol{U}_j)(\cdot)^{\mathrm{T}} \\
=&[\cdot]f_i^k(\boldsymbol{P}(0))[\cdot]^{\mathrm{T}} + \sum_{j,j\neq i}(\cdot)f_j^k(\boldsymbol{P}(0)(\cdot)^{\mathrm{T}} + \boldsymbol{U}_i \\
&+ [\cdot](\sum_{l=1}^{k-1} f_i^l(\boldsymbol{U}) + \boldsymbol{U}_i)[\cdot]^{\mathrm{T}} \\
&+ \sum_{j,j\neq i}(\cdot)(\sum_{l=1}^{k-1} f_j^l(\boldsymbol{U}) + \boldsymbol{U}_j)(\cdot)^{\mathrm{T}} \\
=&f_i^{k+1}(\boldsymbol{P}(0)) + \sum_{l=1}^{k} f_i^l(\boldsymbol{U}) + \boldsymbol{U}_i.
\end{aligned}$$

Therefore, the equation is satisfied.

Similar to the proof of (a), but assume that $f_i^k(\boldsymbol{P}^*) \leq \beta^k \boldsymbol{P}_i^*$, $\boldsymbol{P}_i(0) \leq \gamma \boldsymbol{P}_i^*$, and $\boldsymbol{U}_i \leq \delta \boldsymbol{P}_i^*$ for all i's with certain constants β, γ, and δ. Notice that $0 < \beta < 1$ and we then have

$$\begin{aligned}
\boldsymbol{P}_i(k+1) \leq& \gamma\beta^{k+1}\boldsymbol{P}_i^* + \sum_{l=1}^{k} \delta\beta^l \boldsymbol{P}_i^* + \delta\boldsymbol{P}_i^* \\
<&(\gamma + \frac{\delta}{1-\beta})\boldsymbol{P}_i^*.
\end{aligned}$$

Thus, the lemma is proved.

Proof of Lemma 3

(1. $\Rightarrow$ 2.) Take the matrix $\boldsymbol{P} = \sum_{i=0}^{\infty} \boldsymbol{\Gamma}^i \boldsymbol{Q} (\boldsymbol{\Gamma}^{\mathrm{T}})^i$, which is well defined by the asymptotic stability of $\boldsymbol{\Gamma}$, and $\boldsymbol{P} = \boldsymbol{P}^{\mathrm{T}} > 0$ by definition. Now, substitute $\boldsymbol{P}$ into Eq. (24):

$$\begin{aligned}\boldsymbol{\Gamma P \Gamma}^{\mathrm{T}} - \boldsymbol{P} =& \boldsymbol{\Gamma}(\sum_{i=0}^{\infty} \boldsymbol{\Gamma}^i \boldsymbol{Q} (\boldsymbol{\Gamma}^{\mathrm{T}})^i) \boldsymbol{F}^{\mathrm{T}} - \sum_{i=0}^{\infty} \boldsymbol{\Gamma}^i \boldsymbol{Q} (\boldsymbol{\Gamma}^{\mathrm{T}})^i \\ =& \sum_{i=1}^{\infty} \boldsymbol{\Gamma}^i \boldsymbol{Q} (\boldsymbol{\Gamma}^{\mathrm{T}})^i - \sum_{i=0}^{\infty} \boldsymbol{\Gamma}^i \boldsymbol{Q} (\boldsymbol{\Gamma}^{\mathrm{T}})^i = -\boldsymbol{Q}.\end{aligned}$$

To show uniqueness, suppose that there is another matrix $\bar{\boldsymbol{P}}$ that satisfies Eq. (24). After some recursion, we can show that if both $\boldsymbol{P}$ and $\bar{\boldsymbol{P}}$ satisfy Eq. (24), then

$$\boldsymbol{\Gamma}^N (\boldsymbol{P} - \bar{\boldsymbol{P}}) (\boldsymbol{\Gamma}^{\mathrm{T}})^N = \boldsymbol{P} - \bar{\boldsymbol{P}}.$$

Letting $N \to \infty$ yields the result.

(2. $\Rightarrow$ 1.) Now we consider the Lyapunov function $V(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{P} \boldsymbol{x}$, and have

$$\begin{aligned}V(\boldsymbol{x}(k+1)) - V(\boldsymbol{x}(k)) &= -\boldsymbol{x}(k)^{\mathrm{T}} \boldsymbol{Q} \boldsymbol{x}(k) \\ &\leq -\lambda_{\min}(\boldsymbol{Q}) \|\boldsymbol{x}(k)\|_2^2.\end{aligned}$$

According to Lyapunov stability, the linear system is exponentially stable. Therefore, $\boldsymbol{F}$ is Schur stable.