Distributed Kalman Filtering for Interconnected Dynamic Systems

Yuchen Zhang, *Graduate Student Member, IEEE*, Bo Chen, *Member, IEEE*, Li Yu, *Member, IEEE*, and Daniel W. C. Ho, *Fellow, IEEE*

***Abstract*—This article is concerned with the distributed Kalman filtering problem for interconnected dynamic systems, where the local estimator of each subsystem is designed only by its own information and neighboring information. A decoupling strategy is developed to minimize the impact of interconnected terms on the estimation performance, and then the recursive and distributed Kalman filter is derived in the minimum mean-squared error sense. Moreover, by using Lyapunov criterion for linear time-varying systems, stability conditions are presented such that the designed estimator is bounded. Finally, a heavy duty vehicle platoon system is employed to show the effectiveness and advantages of the proposed methods.**

***Index Terms*—Decoupling strategy, distributed Kalman filtering, interconnected dynamic systems (IDSs), stability analysis.**

I. Introduction

INTERCONNECTED dynamic systems (IDSs) consist of a variety of subsystems that are deployed over a large geographical region, where each subsystem involves physical interactions with its neighboring subsystems [1], [2]. With the development of the sensor and communication technology as well as the emergence of some special system tasks that require subsystem collaborations, IDSs have found applications in many fields, such as multiagent systems [2], [3]; complex networks [4]; and smart grids [5]. As a basic topic of IDSs, the state estimation problem has attracted considerable research interest because real-time monitoring of systems is helpful to design efficient control or decision strategies [6]. One way to extract state information from measurements is to augment the states of subsystems as a unified model to design a centralized estimator, where all the local measurements are transmitted to a fusion center. However, the dimension of the state of the entire system is usually huge, which will lead to large computational burden for the centralized estimation method [7]. Even though all the local information can be collected at the expense of the global communication burden, it is still very difficult to obtain the intricate structure information (such as dynamic behavior) of the interconnection of the entire system [8]. The complexity of the interconnected systems can be staggering, especially those systems with varying interconnected terms. On the other hand, a distributed estimation method estimates the state of each subsystem using only local and neighboring information [9] and, thus, can overcome the disadvantages of the centralized method. Therefore, more attention has been focused on distributed estimation methods [10], [11].

Different physical interconnection structures and communication structures directly affect the performance and the stability of estimators deployed in IDSs. Therefore, the strategy to deal with interconnected terms in IDSs is of the most importance for distributed estimator designing. When IDSs have special structures, various estimation approaches have been developed. For example, local unbiased filters were designed in [12] and [13] with specific structures to exchange information, while a decentralized estimation algorithm was developed in [14] based on the consensus strategies for sparsely connected systems with special measurement structures. Meanwhile, a distributed moving horizon estimation method [15] and a distributed Kalman filtering algorithm [16] based on the approaches of the distributed matrix inversion were developed for sparse systems decomposed into coupled subsystems, where the coupled states can be locally observed. Recently, a suboptimal distributed Kalman filtering problem was addressed in [17] for a class of sequentially interconnected systems. In addition, the fusion estimation problem for subsystems with overlapping states was investigated in [18], and several fusion methods were developed under the situation of unknown correlation [19]. Although the interconnected terms for these system structures can be completely decoupled under certain conditions, it is difficult for most IDSs to transform into these structures. Therefore, the ϵ-technique was introduced in [20] to obtain the near-optimal estimation, while a decoupling strategy was proposed in [21] to design stable and distributed estimators for general IDSs. Without any structure constraints of the interconnected terms, distributed estimation methods have been designed in [22] and [23] for the time-invariant IDSs based on the stability analysis of the small gain

Manuscript received 28 September 2020; revised 30 January 2021; accepted 25 March 2021. Date of publication 16 June 2021; date of current version 17 October 2022. This work was supported in part by the National Natural Science Funds of China under Grant 61973277 and Grant 62073292; in part by the Zhejiang Provincial Natural Science Foundation of China under Grant LR20F030004; in part by the Research Grants Council of the Hong Kong Special Administrative Region, China, (CityU) under Grant 11200717 and Grant 11202819; and in part by the CityU Strategic Research Grant under Grant 7005511. This article was recommended by Associate Editor B. Jiang. *(Corresponding author: Bo Chen.)*

Yuchen Zhang, Bo Chen, and Li Yu are with the Department of Automation, Zhejiang University of Technology, Hangzhou 310023, China, and also with the Institute of Cyberspace Security, Zhejiang University of Technology, Hangzhou 310023, China (e-mail: YuchenZhang95@163.com; bchen@aliyun.com; lyu@zjut.edu.cn).

Daniel W. C. Ho is with the Department of Mathematics, City University of Hong Kong, Hong Kong (e-mail: madaniel@cityu.edu.hk).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TCYB.2021.3072198.

Digital Object Identifier 10.1109/TCYB.2021.3072198

theorem. At the same time, the state estimation problem of complex networks, modeled as general IDSs, was addressed in [4] and [24], while a robust distributed state estimator was proposed in [5] for interconnected power systems with quasisteady states. Different constraints on the interconnected structure were assumed in [12]–[18] to guarantee the viability of distributed estimator design. However, without structure constraints of IDSs, how to design distributed estimation methods based on the local and neighboring information is still challenging.

On the other hand, the design of a stable and distributed estimator without any global communication structures can significantly alleviate the communication and computational burden among sensors, especially for high-dimensional systems [16] and, thus, can meet the requirement of increasing network nodes. With only local communication capacity, the work in [17] designed a distributed Kalman filter (DKF) to optimally estimate subsystems' states in a recursive way, and the work in [21] designed a distributed estimator with the decoupling strategy to address the estimation problem for general interconnected systems. However, the DKF in [17] was designed for sequentially interconnected systems, while the distributed estimator in [21] cannot achieve optimal performance under Gaussian noises. Motivated by the above analysis, we shall study the distributed Kalman filtering problem for general IDSs based on the local and neighboring information. The main contributions of this article can be summarized as follows.

1) After utilizing the decoupling strategy that can minimize the effect of interconnected terms, a locally optimal distributed estimator is designed based on local and neighboring information in the minimum mean-squared error sense.
2) By exploring the topological structure of estimation error systems and the Lyapunov criterion for linear time-varying systems, the stability conditions are derived such that the designed DKF is bounded.

Notations: Define $\mathbb{N}_L = \{1, 2, \dots, L\}$, where L is a natural number excluding zero, and denote the set of n-dimensional real vectors by $\mathbb{R}^n$. Given sets A and B, $|A|$ means the number of elements of set A, while $A\backslash B$ represents the set of all elements of A that are not in B. The superscript "T" represents the transpose, while "I" is the identity matrix with appropriate dimensions. The symmetric terms in a symmetric matrix are denoted by "$*$," while $X > (<)0$ is a positive-definite (negative-definite) matrix. The notation col$\{a_1, \dots, a_n\}$ means a column vector, whose elements are $a_1, \dots, a_n$, while diag$\{\cdot\}$ stands for a block diagonal matrix. The mathematical expectation is denoted by E$\{\cdot\}$, and $\|A\|_2$ is the 2-norm of matrix A. The orthogonality is denoted by $\perp$, while Tr(A) represents the trace of the matrix A.

II. Problem Formulation

In this section, we recall some definitions of graph theory and formulate the problem of distributed Kalman filtering for IDSs.

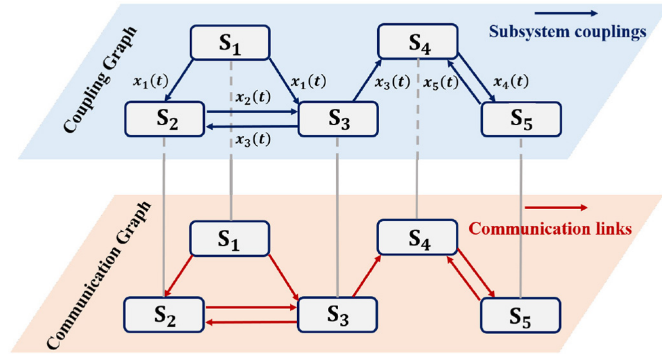


Fig. 1. Example of an interconnected system with separate coupling and communication layers.

A. Graph Theory

Definition 1 (Directed Graph): A directed graph $\mathcal{G}$ is an ordered pair $(\mathcal{N}(\mathcal{G}), \mathcal{E}(\mathcal{G}))$ consisting of a set $\mathcal{N} = \{\mathfrak{n}_1, \mathfrak{n}_2, \dots, \mathfrak{n}_L\}$ of nodes/vertices and a set $\mathcal{E} \subset \mathcal{N} \times \mathcal{N}$ of edges/links, together with an incidence function that is associated with each edge an ordered pair of nodes $(\mathfrak{n}_i, \mathfrak{n}_j)$.

A directed edge $(\mathfrak{n}_i, \mathfrak{n}_j)$ indicates $\mathfrak{n}_i$ couples/interconnects with $\mathfrak{n}_j$. The nodes, which couple with a node $\mathfrak{n}_i$, are its in-neighbors, denoted by $\overline{\Omega}_i = \{\mathfrak{n}_j \in \mathcal{N} : (\mathfrak{n}_j, \mathfrak{n}_i) \in \mathcal{E}\}$. Those which are coupled by the node $\mathfrak{n}_i$ are its out neighbors, denoted by $\underline{\Omega}_i = \{\mathfrak{n}_j \in \mathcal{N} : (\mathfrak{n}_i, \mathfrak{n}_j) \in \mathcal{E}\}$. The directed graph $\mathcal{G}$ can be fully represented by its adjacency matrix, where the elements of the adjacency matrix $\mathcal{A}$ are defined as

$$a_{i,j} = \begin{cases} 1, & \text{if } (\mathfrak{n}_i, \mathfrak{n}_j) \in \mathcal{E} \\ 0, & \text{otherwise.} \end{cases} \tag{1}$$

Then, we recall the definitions of the directed acyclic graph and acyclic ordering.

Definition 2 (Directed Acyclic Graph): A directed acyclic graph is defined as a directed graph that has no cycles.

A cycle in a graph is a directed path that starts and ends at the same node and contains no repeated nodes, except for the initial and the final nodes.

Definition 3 (Acyclic Ordering): An acyclic ordering of a directed graph $\mathcal{G}$ is an ordering of its vertices, that is, $\{\tilde{\mathfrak{n}}_1, \tilde{\mathfrak{n}}_2, \dots, \tilde{\mathfrak{n}}_L\}$, which satisfies $i < j$ for every edge $(\tilde{\mathfrak{n}}_i, \tilde{\mathfrak{n}}_j)$ in graph $\mathcal{G}$.

Lemma 1 [25]: Every directed acyclic graph has an acyclic ordering of its vertices.

B. Interconnected System Model

Consider a discrete-time interconnected system $\mathbf{S}$ comprised of L subsystems (see Fig. 1 for an example), where the state dynamics of the ith subsystem $\mathbf{S}_i, i \in \mathbb{N}_L$ is described as follows:

$$\mathbf{S}_i : \begin{cases} x_i(k+1) = A_i x_i(k) + B_i u_i(k) + \Gamma_i w_i(k) \\ \qquad + \sum_{j \in \mathbb{N}_L \backslash \{i\}} A_{i,j} x_j(k) \\ i = 1, 2, \dots, L \end{cases} \tag{2}$$

where $x_i(k) \in \mathbb{R}^{n_i}$ is the state of subsystem $\mathbf{S}_i$, while A_i, Γ_i, and $A_{i,j}$ are matrices with appropriate dimensions, and $w_i(k)$ is the system noise. Meanwhile, the ith subsystem's sensor measurement $y_i(k) \in \mathbb{R}^{m_i}$ is modeled by

$$y_i(k) = C_i x_i(k) + v_i(k) \tag{3}$$

where C_i is the measurement matrix with full rank, and $v_i(k)$ is the measurement noise.

The subsystem couplings are characterized by matrices $A_{i,j}$, where $A_{i,j} \neq 0$ indicates that subsystem $\mathbf{S}_i$ is coupled by subsystem $\mathbf{S}_j$. Thus, a directed coupling graph $\mathcal{G}_I = (\mathcal{N}, \mathcal{E}_I)$ is introduced to better describe subsystem dynamics, where $\mathcal{N} = \{\mathbf{S}_1, \ldots, \mathbf{S}_L\}$ and $\mathcal{E}_I \subset \mathcal{N} \times \mathcal{N}$ represent the set of nodes (i.e., subsystems) and the set of edges (i.e., couplings), respectively. The adjacency matrix of $\mathcal{G}_I$ is $\mathcal{A} = [a_{i,j}]_{i,j\in \mathrm{N}_L}$ with $a_{i,i} = 0$ and $a_{i,j}$ $(j \in \mathbb{N}_L \backslash \{i\})$ obtained by

$$a_{i,j} = \begin{cases} 1, & \text{if } A_{j,i} \neq 0 \\ 0, & \text{otherwise.} \end{cases} \tag{4}$$

Moreover, the set of the indexes for in-neighbors of subsystem $\mathbf{S}_i$ can be obtained by

$$\overline{\Omega}_i = \left\{ i_\kappa^\rho : A_{i,i_\kappa^\rho} \neq 0,\ i_\kappa^\rho \in \mathbb{N}_L \backslash \{i\} \right\} \tag{5}$$

where the number of elements in $\overline{\Omega}_i$ is $|\overline{\Omega}_i| = \theta_i < L$ and, thus, this set can be represented as $\overline{\Omega}_i = \{i_1^\rho, \ldots, i_\kappa^\rho, \ldots, i_{\theta_i}^\rho\}$. Moreover, to develop a decoupling strategy later, the set $\overline{\Omega}_i$ is divided into two nonoverlapping subsets $\overline{\Omega}_{c_i}$ and $\overline{\Omega}_{r_i}$

$$\begin{cases} \overline{\Omega}_{c_i} = \left\{ i_\kappa^c : i_\kappa^c \in \overline{\Omega}_i, C_{i_\kappa^c} \text{has full column rank} \right\} \\ \overline{\Omega}_{r_i} = \left\{ i_\kappa^r : i_\kappa^r \in \overline{\Omega}_i, C_{i_\kappa^r} \text{does not have full column rank} \right\} \end{cases} \tag{6}$$

where the numbers of elements in $\overline{\Omega}_{c_i}$ and $\overline{\Omega}_{r_i}$ are $|\overline{\Omega}_{c_i}| = \theta_{c_i}$ and $|\overline{\Omega}_{r_i}| = \theta_{r_i}$, respectively. Therefore, these two subsets can be clearly represented as $\overline{\Omega}_{c_i} = \{i_1^c, \ldots, i_\kappa^c, \ldots, i_{\theta_{c_i}}^c\}$ and $\overline{\Omega}_{r_i} = \{i_1^r, \ldots, i_\kappa^r, \ldots, i_{\theta_{r_i}}^r\}$. It is also found from (6) that $\theta_{c_i} + \theta_{r_i} = \theta_i$.

As is well known, each subsystem needs to exchange its information via a communication network to collaboratively accomplish predetermined system tasks (i.e., the distributed estimation task in this work). The communication among subsystems is analogously described as a directed graph $\mathcal{G}_c = \{\mathcal{N}, \mathcal{E}_c\}$, where $\mathcal{N} = \{\mathbf{S}_1, \ldots, \mathbf{S}_L\}$ and $\mathcal{E}_c \subset \mathcal{N} \times \mathcal{N}$ represent the set of nodes (i.e., subsystems) and the set of edges (i.e., communication link), respectively. An edge $(\mathbf{S}_i, \mathbf{S}_j)$ of $\mathcal{G}_c$ indicates there exists a communication link from $\mathbf{S}_i$ to $\mathbf{S}_j$.

Assumption 1 (Communication): All the edges of the directed coupling graph $\mathcal{G}_I$ are contained in the directed communication graph $\mathcal{G}_c$.

Assumption 2 (Gaussian Noise): The system noise $w_i(k)$ and measurement noise $v_i(k)$ are assumed as uncorrelated Gaussian white noises satisfying

$$\begin{cases} \mathrm{E}\left[w_i(k) w_j(k_1)\right] = \delta_{i,j} \delta_{k,k_1} Q_{w_i} \\ \mathrm{E}\left[v_i(k) v_j(k_1)\right] = \delta_{i,j} \delta_{k,k_1} Q_{v_i} \\ \mathrm{E}\left[w_i(k) v_j(k_1)\right] = 0 \quad (\forall i, j, k, k_1) \end{cases} \tag{7}$$

where Q_{w_i} and Q_{v_i} are the known covariances of $w_i(k)$ and $v_i(k)$, respectively. $\delta_{k,k_1} = 0$ if $k \neq k_1$ and $\delta_{k,k_1} = 1$ otherwise.

Remark 1: Assumption 1 ensures that subsystem $\mathbf{S}_i$ can receive information of measurements and estimates from its in-neighbors $\mathbf{S}_{i_\kappa^\rho} (i_\kappa^\rho \in \overline{\Omega}_i)$ (see the communication structure in Fig. 1). For convenience, the neighboring information in this article is used to refer to the information from in-neighbors. Although the communication link in most situations is bidirectional, the directed communication graph assumption is helpful to construct a minimum information exchange strategy for further analysis. On the other hand, the communication structure is limited to the energy of subsystems and network bandwidth. Therefore, a centralized communication structure with one subsystem that communicates with all the other subsystems is not practical.

C. Distributed Kalman Filtering Problem

To estimate subsystems' states in a distributed manner, the estimator designed for each subsystem can only leverage local information and the information from a communication network. Therefore, the following DKF for the ith subsystem is proposed:

$$\begin{cases} \hat{x}_i^p(k) = A_i \hat{x}_i(k-1) + B_i u_i(k-1) \\ \qquad + \sum_{i_\kappa^\rho \in \overline{\Omega}_i} A_{i,i_\kappa^\rho} \hat{x}_{i_\kappa^\rho}(k-1) \\ \hat{x}_i(k) = \hat{x}_i^p(k) + K_i(k)\left[y_i(k) - C_i \hat{x}_i^p(k)\right] \\ \qquad + \sum_{i_\kappa^\rho \in \overline{\Omega}_i} K_{i,i_\kappa^\rho}(k) \left[y_{i_\kappa^\rho}(k-1) - C_{i_\kappa^\rho} \hat{x}_{i_\kappa^\rho}(k-1) \right] \end{cases} \tag{8}$$

where $\hat{x}_i^p(k)$ and $\hat{x}_i(k)$ denote the one-step prediction and filtered state estimate of $x_i(k)$, respectively. Notice that local estimation errors will remain coupled and, thus, the innovation term $y_{i_\kappa^\rho}(k-1) - C_{i_\kappa^\rho} \hat{x}_{i_\kappa^\rho}(k-1)$ in the estimator is used for estimation errors decoupling. In addition, the estimator gain $K_i(k)$ and the decoupling gain $K_{i,i_\kappa^\rho}(k) (i_\kappa^\rho \in \overline{\Omega}_i)$ will be designed later. Defining the one-step prediction error $\tilde{x}_i^p(k) \triangleq x_i(k) - \hat{x}_i^p(k)$ and local estimation error $\tilde{x}_i(k) \triangleq x_i(k) - \hat{x}_i(k)$, then by (2) and (8), it is found that

$$\begin{cases} \tilde{x}_i^p(k) = A_i \tilde{x}_i(k-1) + \sum_{i_\kappa^\rho \in \overline{\Omega}_i} A_{i,i_\kappa^\rho} \tilde{x}_{i_\kappa^\rho}(k-1) \\ \qquad + \Gamma_i w_i(k-1) \\ \tilde{x}_i(k) = K_{C_i}(k) \tilde{x}_i^p(k) - K_i(k) v_i(k) \\ \qquad - \sum_{i_\kappa^\rho \in \overline{\Omega}_i} K_{i,i_\kappa^\rho}(k) \left[C_{i_\kappa^\rho} \tilde{x}_{i_\kappa^\rho}(k-1) + v_{i_\kappa^\rho}(k-1) \right] \end{cases} \tag{9}$$

where $K_{C_i}(k) \triangleq [I - K_i(k) C_i]$. Thus, the local estimation error system is obtained as

$$\tilde{\mathbf{S}}_i : \begin{cases} \tilde{x}_i(k) = K_{C_i}(k) A_i \tilde{x}_i(k-1) - K_i(k) v_i(k) \\ \quad + K_{C_i}(k) \Gamma_i w_i(k-1) \\ \quad + \sum_{i_\kappa^\rho \in \overline{\Omega}_i} \Big\{ \left[K_{C_i}(k) A_{i,i_\kappa^\rho} - K_{i,i_\kappa^\rho}(k) C_{i_\kappa^\rho} \right] \\ \qquad \tilde{x}_{i_\kappa^\rho}(k-1) - K_{i,i_\kappa^\rho}(k) v_{i_\kappa^\rho}(k-1) \Big\} \\ i = 1, 2, \ldots, L \end{cases} \tag{10}$$

where $\tilde{x}_{i_\kappa^\rho}(k-1)$ is the interconnected estimation error of subsystem $\mathbf{S}_i$. Define the coefficient matrix of $\tilde{x}_{i_\kappa^\rho}(k-1) (i_\kappa^\rho \in \overline{\Omega}_i)$

as $A^e_{i,i^\rho_\kappa}(k) \triangleq K_{C_i}(k)A_{i,i^\rho_\kappa} - K_{i,i^\rho_\kappa}(k)C_{i^\rho_\kappa}$, then it is said that the estimation error system $\tilde{\mathbf{S}}_{i^\rho_\kappa}$ interacts/couples with $\tilde{\mathbf{S}}_i$ when $A^e_{i,i^\rho_\kappa}(k)$ is not designed to be 0.

To describe the estimation error interactions among local estimation error systems, a directed estimation error relationship graph $\mathcal{G}_e = \{\mathcal{N}_e, \mathcal{E}_e\}$ is introduced, where $\mathcal{N}_e = \{\tilde{\mathbf{S}}_1, \ldots, \tilde{\mathbf{S}}_L\}$ and $\mathcal{E}_e \subset \mathcal{N}_e \times \mathcal{N}_e$ represent the set of nodes (i.e., local estimation error systems) and the set of edges (i.e., estimation error interactions), respectively. The adjacency matrix of $\mathcal{G}_e$ is $\mathcal{A}_e = [e_{i,j}]_{i,j\in\mathbb{N}_L}$ with $e_{i,i} = 0$ and $e_{i,j}$ $(j \neq i)$ obtained by

$$e_{i,j} = \begin{cases} 1, & \text{if } A^e_{j,i}(k) \neq 0 \ \left(i \in \overline{\Omega}_j\right) \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

Note that interconnected estimation errors are actually caused by couplings in IDSs and can be decoupled by carefully selecting the decoupling matrix $K_{i,i^\rho_\kappa}(k)$ to minimize the coefficient matrix $A^e_{i,i^\rho_\kappa}(k)$. When the system $\mathbf{S}$ is not interconnected, the adjacency matrix $\mathcal{A}_e$ is reduced to the null matrix. In this case, the optimal local estimator for the noninterconnected subsystem can be designed as the celebrated Kalman filter with minimum estimation error covariance. However, without any decoupling strategies for interconnected systems, the estimation error relationship graph has the same topology as the coupling graph. These interconnected terms in estimation error systems increase the difficulty of ensuring the stability of each local estimator and, thus, the design of a decoupling strategy is necessary.

Consequently, the aims of this article are described as follows.

1) Design a locally optimal distributed estimator gain $K_i(k)$ in the minimum mean-squared error sense, while the decoupling gain $K_{i,i^\rho_\kappa}(k)(i^\rho_\kappa \in \overline{\Omega}_i)$ is to be properly chosen such that the impact of the interconnected estimation error is minimized.
2) Find the stability conditions of the designed estimator such that the estimator is bounded.

Remark 2: Compared with the IDSs modeled in [12]–[18], there are no structure constraints in (2) and (3) for the addressed IDSs, and the coupled states cannot be locally measured. Notice that the distributed estimation problem for general IDSs without structure constraints is more challenging because the local estimation error in (10) will be largely influenced by their in-neighbors' estimation errors. In this case, the boundedness of each local estimation error is also determined by the boundedness of its in-neighbors' estimation error. On the other hand, the DKF design, unlike the approach in [16], is in the subsystem level and only use local information and neighboring information.

III. Main Results

A. Design of the Distributed Kalman Filter

In this section, we will present the design method for estimator gains $K_{i,i^\rho_\kappa}(k)(i^\rho_\kappa \in \overline{\Omega}_i)$ and $K_i(k)$. Before deriving the main results, let us define

$$\begin{cases} C_{N,i^c_\kappa} \triangleq \left(C^{\mathrm{T}}_{i^c_\kappa} C_{i^c_\kappa}\right)^{-1} C^{\mathrm{T}}_{i^c_\kappa} \\ C_{M,i^r_\kappa} \triangleq C^{\mathrm{T}}_{i^r_\kappa}\left(C_{i^r_\kappa} C^{\mathrm{T}}_{i^r_\kappa}\right)^{-1} \\ \chi_i(k) \triangleq A_i\tilde{x}_i(k-1) \\ \qquad + \sum_{i^r_\kappa\in\overline{\Omega}_{r_i}} A_{i,i^r_\kappa}\left(I - C_{M,i^r_\kappa}C_{i^r_\kappa}\right)\tilde{x}_{i^r_\kappa}(k-1) \\ \xi_i(k) \triangleq \sum_{i^r_\kappa\in\overline{\Omega}_{r_i}} A_{i,i^r_\kappa} C_{M,i^r_\kappa} v_{i^r_\kappa}(k-1) \\ \qquad + \sum_{i^c_\kappa\in\overline{\Omega}_{c_i}} A_{i,i^c_\kappa} C_{N,i^c_\kappa} v_{i^c_\kappa}(k-1) \\ \Upsilon_{i,j} \triangleq \mathrm{E}\left[\xi_i(k)\xi_j^{\mathrm{T}}(k)\right] \\ \quad = \sum_{i^r_\kappa\in\overline{\Omega}_{r_i}}\sum_{j^r_\kappa\in\overline{\Omega}_{r_j}} \delta_{i^r_\kappa,j^r_\kappa} A_{i,i^r_\kappa} C_{M,i^r_\kappa} Q_{v_{i^r_\kappa}} C^{\mathrm{T}}_{M,j^r_\kappa} A^{\mathrm{T}}_{j,j^r_\kappa} \\ \qquad + \sum_{i^c_\kappa\in\overline{\Omega}_{c_i}}\sum_{j^c_\kappa\in\overline{\Omega}_{c_j}} \delta_{i^c_\kappa,j^c_\kappa} A_{i,i^c_\kappa} C_{N,i^c_\kappa} Q_{v_{i^c_\kappa}} C^{\mathrm{T}}_{N,j^c_\kappa} A^{\mathrm{T}}_{j,j^c_\kappa} \\ \Phi_{i,j}(k) \triangleq \mathrm{E}\left[\chi_i(k)\xi_j^{\mathrm{T}}(k)\right] \\ \quad = -\sum_{j^r_\kappa\in\overline{\Omega}_{r_j}} \delta_{i,j^r_\kappa} A_i K_i(k-1) Q_{v_i} C^{\mathrm{T}}_{M,i} A^{\mathrm{T}}_{j,i} \\ \qquad - \sum_{j^c_\kappa\in\overline{\Omega}_{c_j}} \delta_{i,j^c_\kappa} A_i K_i(k-1) Q_{v_i} C^{\mathrm{T}}_{N,i} A^{\mathrm{T}}_{j,i} \\ \qquad - \sum_{i^r_\kappa\in\overline{\Omega}_{r_i}}\sum_{j^r_\kappa\in\overline{\Omega}_{r_j}} \Big\{\delta_{i^r_\kappa,j^r_\kappa} A_{i,i^r_\kappa}\left(I - C_{M,i^r_\kappa}C_{i^r_\kappa}\right) \\ \qquad\qquad \times K_{i^r_\kappa}(k-1) Q_{v_{i^r_\kappa}} C^{\mathrm{T}}_{M,j^r_\kappa} A^{\mathrm{T}}_{j,j^r_\kappa}\Big\} \\ \hat{Q}_{w_i} \triangleq \Gamma_i Q_{w_i}\Gamma_i^{\mathrm{T}}. \end{cases} \tag{12}$$

Theorem 1: For the distributed filter (8), an optimal estimator decoupling gain $K_{i,i^\rho_\kappa}(k)$ that minimizes the effect of the interconnected estimation error is calculated by

$$K_{i,i^\rho_\kappa}(k) = \begin{cases} K_{i,i^c_\kappa}(k), & \text{if } i^\rho_\kappa \in \overline{\Omega}_{c_i} \\ K_{i,i^r_\kappa}(k), & \text{if } i^\rho_\kappa \in \overline{\Omega}_{r_i} \end{cases} \tag{13}$$

where $K_{i,i^c_\kappa}(k)$ and $K_{i,i^r_\kappa}(k)$ are given as

$$\begin{cases} K_{i,i^c_\kappa}(k) = K_{C_i}(k)A_{i,i^c_\kappa}C_{N,i^c_\kappa} \\ K_{i,i^r_\kappa}(k) = K_{C_i}(k)A_{i,i^r_\kappa}C_{M,i^r_\kappa}. \end{cases} \tag{14}$$

Meanwhile, a locally optimal estimator gain $K_i(k)$ in linear minimum mean-squared error sense is calculated by

$$\begin{aligned} K_i(k) = &\left[\Lambda_{i,i}(k) + \Upsilon_{i,i} - \Phi_{i,i}(k) - \Phi^{\mathrm{T}}_{i,i}(k) + \hat{Q}_{w_i}\right]C_i^{\mathrm{T}} \\ &\times \left\{C_i\left[\Lambda_{i,i}(k) + \Upsilon_{i,i} - \Phi_{i,i}(k) - \Phi^{\mathrm{T}}_{i,i}(k) + \hat{Q}_{w_i}\right]C_i^{\mathrm{T}} \right. \\ &\left. + Q_{v_i}\right\}^{-1} \end{aligned} \tag{15}$$

where $\Lambda_{i,j}(k) \triangleq \mathrm{E}[\chi_i(k)\chi_j^{\mathrm{T}}(k)]$ and the formula of $\Lambda_{i,i}(k)$ is as follows:

$$\begin{aligned} \Lambda_{i,i}(k) = &A_i P_{i,i}(k-1)A_i^{\mathrm{T}} \\ &+ \sum_{i^r_\kappa\in\overline{\Omega}_{r_i}} A_{i,i^r_\kappa}\left(I - C_{M,i^r_\kappa}C_{i^r_\kappa}\right)P_{i^r_\kappa,i}(k-1)A_i^{\mathrm{T}} \\ &+ \sum_{i^r_\kappa\in\overline{\Omega}_{r_i}} A_i P_{i,i^r_\kappa}(k-1)\left(I - C_{M,i^r_\kappa}C_{i^r_\kappa}\right)^{\mathrm{T}} A^{\mathrm{T}}_{i,i^r_\kappa} \\ &+ \sum_{i^r_{\kappa_1}\in\overline{\Omega}_{r_i}}\sum_{i^r_{\kappa_2}\in\overline{\Omega}_{r_i}} A_{i,i^r_{\kappa_1}}\left(I - C_{M,i^r_{\kappa_1}}C_{i^r_{\kappa_1}}\right) \\ &\times P_{i^r_{\kappa_1},i^r_{\kappa_2}}(k-1)\left(I - C_{M,i^r_{\kappa_2}}C_{i^r_{\kappa_2}}\right)^{\mathrm{T}} A^{\mathrm{T}}_{i,i^r_{\kappa_2}} \end{aligned} \tag{16}$$

where $P_{i,j}(k-1) \triangleq \mathrm{E}\{\tilde{x}_i(k-1)\tilde{x}_j^{\mathrm{T}}(k-1)\}$ is the estimation error cross-covariance at the instant $k-1$, and it can be calculated as

$$\begin{cases} P_{i,i}(k-1) = K_{C_i}(k-1)\Big[\Lambda_{i,i}(k-1) + \Upsilon_{i,i} - \Phi_{i,i}(k-1) \\ \qquad\qquad - \Phi_{i,i}^{\mathrm{T}}(k-1) + \hat{Q}_{w_i}\Big] \\ P_{i,j}(k-1) = K_{C_i}(k-1) \\ \quad \times \Big[\Lambda_{i,j}(k-1) + \Upsilon_{i,j} - \Phi_{i,j}(k-1) - \Phi_{i,j}^{\mathrm{T}}(k-1)\Big] \\ \quad \times K_{C_j}^{\mathrm{T}}(k-1) \quad (j \neq i) \end{cases} \tag{17}$$

with

$$\begin{aligned} \Lambda_{i,j}(k-1) = {} & A_i P_{i,j}(k-2) A_j^{\mathrm{T}} \\ & + \sum_{i_\kappa^r \in \overline{\Omega}_{r_i}} A_{i,i_\kappa^r}\big(I - C_{M,i_\kappa^r} C_{i_\kappa^r}\big) P_{i_\kappa^r,j}(k-2) A_j^{\mathrm{T}} \\ & + \sum_{j_\kappa^r \in \overline{\Omega}_{r_j}} A_i P_{i,j_\kappa^r}(k-2)\big(I - C_{M,j_\kappa^r} C_{j_\kappa^r}\big)^{\mathrm{T}} A_{i,j_\kappa^r}^{\mathrm{T}} \\ & + \sum_{i_\kappa^r \in \overline{\Omega}_{r_i}} \sum_{j_\kappa^r \in \overline{\Omega}_{r_j}} A_{i,i_\kappa^r}\big(I - C_{M,i_\kappa^r} C_{i_\kappa^r}\big) \\ & \times P_{i_\kappa^r,j_\kappa^r}(k-2)\big(I - C_{M,j_\kappa^r} C_{j_\kappa^r}\big)^{\mathrm{T}} A_{i,j_\kappa^r}^{\mathrm{T}}. \end{aligned} \tag{18}$$

Proof: According to (6), the local estimation error in (10) can be further written as

$$\begin{aligned} \tilde{x}_i(k) = {} & K_{C_i}(k) A_i \tilde{x}_i(k-1) - K_i(k) v_i(k) \\ & + K_{C_i}(k)\Gamma_i w_i(k-1) \\ & + \sum_{i_\kappa^r \in \overline{\Omega}_{r_i}} \Big\{\Big[K_{C_i}(k) A_{i,i_\kappa^r} - K_{i,i_\kappa^r}(k) C_{i_\kappa^r}\Big] \\ & \qquad \tilde{x}_{i_\kappa^r}(k-1)\Big\} - \sum_{i_\kappa^r \in \overline{\Omega}_{r_i}} K_{i,i_\kappa^r}(k) v_{i_\kappa^r}(k-1) \\ & + \sum_{i_\kappa^c \in \overline{\Omega}_{c_i}} \Big\{\Big[K_{C_i}(k) A_{i,i_\kappa^c} - K_{i,i_\kappa^c}(k) C_{i_\kappa^c}\Big] \\ & \qquad \tilde{x}_{i_\kappa^c}(k-1)\Big\} - \sum_{i_\kappa^c \in \overline{\Omega}_{c_i}} K_{i,i_\kappa^c}(k) v_{i_\kappa^c}(k-1) \end{aligned} \tag{19}$$

where the estimation error $\tilde{x}_i(k)$ can be directly affected by the disturbance noises [i.e., $w_i(k-1)$, $v_i(k)$, and $v_{i_\kappa^\rho}(k-1)$ ($i_\kappa^\rho \in \overline{\Omega}_i$)] and interconnected estimation errors (i.e., $\tilde{x}_{i_\kappa^r}(k-1)$ and $\tilde{x}_{i_\kappa^c}(k-1)$). Although the interconnected estimation errors can be viewed as disturbance noises of $\tilde{x}_i(k)$, it is still difficult to know whether these interconnected terms are bounded or not in advance.

Following the idea of reducing the impact of neighboring information to the maximum extent [21], the coefficient matrices of $\tilde{x}_{i_\kappa^r}(k-1)$ and $\tilde{x}_{i_\kappa^c}(k-1)$ in (19) are proposed to be minimized for designing the ith local estimator. Therefore, when the measurement matrix of the in-neighbor $\mathbf{S}_{i_\kappa^c}$ has full column rank (i.e., $C_{i_\kappa^c}$, $i_\kappa^c \in \overline{\Omega}_{c_i}$), let the coefficient matrix of $\tilde{x}_{i_\kappa^c}(k-1)$ be equal to 0 at each instant, then the following equation holds:

$$K_{C_i}(k) A_{i,i_\kappa^c} = K_{i,i_\kappa^c}(k) C_{i_\kappa^c}. \tag{20}$$

Therefore, the decoupling gain $K_{i,i_\kappa^c}(k)$ in (14) is obtained from (20). However, when the measurement matrix of the in-neighbor $\mathbf{S}_{i_\kappa^r}$ does not have full column rank (i.e., $C_{i_\kappa^r}$, $i_\kappa^r \in \overline{\Omega}_{r_i}$), the matrix $C_{i_\kappa^c}^{\mathrm{T}} C_{i_\kappa^c}$ will become irreversible and, thus, (20) does not hold. To solve this problem, define the coefficient matrix of $\tilde{x}_{i_\kappa^r}(k-1)$ as

$$H_{i,i_\kappa^r}(k) = K_{C_i}(k) A_{i,i_\kappa^r} - K_{i,i_\kappa^r}(k) C_{i_\kappa^r} \tag{21}$$

and then the size of $H_{i,i_\kappa^r}(k)$ is depicted by introducing the Frobenius norm [21]

$$\big\|H_{i,i_\kappa^r}(k)\big\|_F \triangleq \sqrt{\mathrm{Tr}\Big(H_{i,i_\kappa^r}^{\mathrm{T}}(k) H_{i,i_\kappa^r}(k)\Big)}. \tag{22}$$

To minimize the impact of $\tilde{x}_{i_\kappa^r}(k-1)$, let $([\partial(\|H_{i,i_\kappa^r}(k)\|_F^2)]/[\partial K_{i,i_\kappa^r}(k)]) = 0$, then the following equality holds:

$$K_{C_i}(k) A_{i,i_\kappa^r} C_{i_\kappa^r}^{\mathrm{T}} - K_{i,i_\kappa^r}(k)\Big[C_{i_\kappa^r} C_{i_\kappa^r}^{\mathrm{T}}\Big] = 0. \tag{23}$$

As a consequence, the calculation process of $K_{i,i_\kappa^r}(k)$ in (14) is obtained in the sense of minimum $\|H_{i,i_\kappa^r}(k)\|_F$.

After $K_{i,i_\kappa^c}(k)$ and $K_{i,i_\kappa^r}(k)$ are determined by (14), the local estimation error system (19) reduces to

$$\begin{aligned} \tilde{x}_i(k) = {} & K_{C_i}(k) A_i \tilde{x}_i(k-1) - K_i(k) v_i(k) \\ & + K_{C_i}(k)\Gamma_i w_i(k-1) \\ & + \sum_{i_\kappa^r \in \overline{\Omega}_{r_i}} \Big\{K_{C_i}(k) A_{i,i_\kappa^r}\big(I - C_{M,i_\kappa^r} C_{i_\kappa^r}\big)\tilde{x}_{i_\kappa^r}(k-1)\Big\} \\ & - \sum_{i_\kappa^r \in \overline{\Omega}_{r_i}} K_{i,i_\kappa^r}(k) v_{i_\kappa^r}(k-1) - \sum_{i_\kappa^c \in \overline{\Omega}_{c_i}} K_{i,i_\kappa^c}(k) v_{i_\kappa^c}(k-1). \end{aligned} \tag{24}$$

It follows from (24) that:

$$\begin{aligned} \tilde{x}_i(k) = {} & K_{C_i}(k)\chi_i(k) - K_{C_i}(k)\xi_i(k) \\ & + K_{C_i}(k)\Gamma_i w_i(k-1) - K_i(k) v_i(k) \end{aligned} \tag{25}$$

where $\chi_i(k)$ and $\xi_i(k)$ are defined in (12). According to (7), one has the following orthogonal properties:

$$\begin{cases} w_i(k-1) \perp \{v_i(k), \chi_i(k), \xi_i(k)\} \\ v_i(k) \perp \{\chi_i(k), \xi_i(k)\}. \end{cases} \tag{26}$$

In this case, the estimation error covariance $P_{i,i}(k)$ is derived from (25) as

$$\begin{aligned} P_{i,i}(k) = {} & K_i(k) Q_{v_i} K_i^{\mathrm{T}}(k) + K_{C_i}(k) \\ & \times \Big[\Lambda_{i,i}(k) + \Upsilon_{i,i} - \Phi_{i,i}(k) - \Phi_{i,i}^{\mathrm{T}}(k) + \hat{Q}_{w_i}\Big] K_{C_i}^{\mathrm{T}}(k) \end{aligned} \tag{27}$$

where $\Upsilon_{i,i}$, $\Phi_{i,i}(k)$, and $\hat{Q}_{w_i}$ are defined in (12). By using the derivation formula of matrix trace, it can be deduced from (27) that

$$\begin{aligned} \frac{\partial \mathrm{Tr}\big\{P_{i,i}(k)\big\}}{\partial K_i(k)} = {} & 2K_i(k) Q_{v_i} + 2K_i(k) C_i \\ & \times \Big[\Lambda_{i,i}(k) + \Upsilon_{i,i} \\ & \qquad - \Phi_{i,i}(k) - \Phi_{i,i}^{\mathrm{T}}(k) + \hat{Q}_{w_i}\Big] C_i^{\mathrm{T}} \\ & - 2\Big[\Lambda_{i,i}(k) + \Upsilon_{i,i} - \Phi_{i,i}(k) \\ & \qquad - \Phi_{i,i}^{\mathrm{T}}(k) + \hat{Q}_{w_i}\Big] C_i^{\mathrm{T}}. \end{aligned} \tag{28}$$

Algorithm 1 For Given $\hat{x}_i(0)$ $(i = 1, ..., L)$ and $P_{i,j}(0)$ $(i = 1, ..., L;\ j = 1, ..., L)$

1: **for** i: = 1 **to** L **do**
2: Subsystem $\mathbf{S}_i$ gathers local measurement $y_i(k)$ and in-neighbors' measurements $y_{i_\kappa^\rho}(k-1)$ $(i_\kappa^\rho \in \overline{\Omega}_i)$ and estimates $\hat{x}_{i_\kappa^\rho}(k-1)$ $(i_\kappa^\rho \in \overline{\Omega}_i)$;
3: Calculate $\Lambda_{i,i}(k)$ by (16);
4: Calculate $K_i(k)$ by (15);
5: Calculate $K_{i,i_\kappa^\rho}(k)$ $(i_\kappa^\rho \in \overline{\Omega}_i)$ by (13) and (14);
6: Calculate $\hat{x}_i(k)$ by (8);
7: Calculate $\Lambda_{i,j}(k)$ by (18) and $P_{i,j}(k)$ by (17) offline;
8: Subsystem $\mathbf{S}_i$ sends local estimate $\hat{x}_i(k)$ and local measurement $y_i(k)$ to its out-neighbors $\mathbf{S}_j$ $(j \in \underline{\Omega}_i)$;
9: **end for**
10: Return to Step 1 and implement Steps 1-9 for calculating $\hat{x}_i(k+1)$ $(i = 1, ..., L)$;

Letting $([\partial \mathrm{Tr}\{P_{i,i}(k)\}]/[\partial K_i(k)]) = 0$, the optimal estimator gain $K_i(k)$ in (15) can be obtained in the minimum mean-squared error sense. Then, by substituting the optimal gain $K_i(k)$ into (27), the minimal estimation error covariance $P_{i,i}(k)$ can be derived as

$$P_{i,i}(k) = K_{C_i}(k)\Big[\Lambda_{i,i}(k) + \Upsilon_{i,i} - \Phi_{i,i}(k) - \Phi_{i,i}^{\mathrm{T}}(k) + \hat{Q}_{w_i}\Big]. \tag{29}$$

On the other hand, the following orthogonality properties hold for $\mathbf{S}_i$ and $\mathbf{S}_j (j \neq i)$:

$$\begin{cases} w_i(k-1) \perp \big\{w_j(k-1), v_j(k), \xi_j(k), \chi_j(k)\big\} \\ v_i(k) \perp \big\{v_j(k), \xi_j(k), \chi_j(k)\big\}. \end{cases} \tag{30}$$

Thus, the estimation error cross-covariance $P_{i,j}(k)$ $(i \neq j)$ in (17) can be obtained. Moreover, (16) and (18) are derived from the definition of $\chi_i(k)$. This completes the proof.

From Theorem 1, the computation procedures for the DKF $\hat{x}_i(k)$ can be summarized by Algorithm 1.

Remark 3: Notice that the designed Kalman filter is distributed. The estimation of local state in (8) is calculated only from in-neighbors' state estimates and measurements. Moreover, the designed optimal estimator gain $K_i(k)$ and the decoupling gain $K_{i,i_\kappa^\rho}(k)$ $(i_\kappa^\rho \in \overline{\Omega}_i)$ in Theorem 1 are independent of measurements and, thus, can be locally calculated in advance. Therefore, the proposed estimator for IDSs is distributed and can achieve communication, computation, and storage in the local sense. Compared with centralized estimators, the proposed distributed estimator can significantly reduce the communication and computational burden for subsystems, and is also more robust to the variation of system structure and communication structure.

Remark 4: When compared with centralized Kalman filers, the proposed estimator can only provide the suboptimal estimation performance due to its subsystem-level design process. However, the estimator gain $K_i(k)$ is designed in the minimum mean-squared error sense to achieve locally optimal estimation performance, and the decoupling gain $K_{i,i_\kappa^\rho}(k)$ $(i_\kappa^\rho \in \overline{\Omega}_i)$ is designed by minimizing the effect of interconnected estimation errors. Therefore, the proposed distributed estimator is easier to achieve stability for the local estimation error system and can provide locally optimal estimation performance.

B. Stability Analysis

The aim of this section is to provide stability conditions for the proposed DKF. Due to the decoupling strategy, the proposed estimator is not a standard Kalman filter and it is difficult to derive the boundedness of estimation error covariance directly. Therefore, we try to ensure the boundedness of the estimation error with the following conditions.

Assumption 3 (Truncated Gaussian Noise): The system noise $w_i(k)$ and measurement noise $v_i(k)$ are assumed to be truncated Gaussian distributed.

Assumption 4 (Directed Acyclic Relationship Graph): The estimation error relationship graph $\mathcal{G}_e$ after decoupling is a directed acyclic graph.

Remark 5: In practical applications, the Gaussian noises (i.e., voltage noises, current noises, or magnetic noises) are caused by finite energy interference, which means their values cannot be unbounded. Therefore, it is reasonable to assume the addressed system is under bounded Gaussian noises from a practical point of view. One of typical bounded Gaussian noises follows truncated Gaussian distribution, and it is commonly used in nonlinear filtering problems [26]. In this case, Assumption 3 is a reasonable compromise for analyzing the stability of the proposed estimator.

Remark 6: According to the decoupling strategy in the last section, the effect of the interconnected estimation error is minimized, which dramatically simplifies the estimation error relationship graph $\mathcal{G}_e$ with $e_{i,i_\kappa^c} = 0$ $(i_\kappa^c \in \overline{\Omega}_{c_i})$. Therefore, the directed acyclic graph assumption for the estimation error system $\tilde{\mathbf{S}}$ can cover a large part of the actual situation.

Proposition 1: The proposed DKF will be bounded, that is

$$\mathrm{Tr}\big\{P_{i,i}(k)\big\} < p_i \tag{31}$$

if there exists another estimator $\hat{x}_i^b(k)$ in the same form as (8), its decoupling gain $K_{i,i_\kappa^\rho}^b(k)$ is designed in the same form as (13) and (14) and its estimator gain $K_i^b(k)$ is designed differently such that the estimation error $\tilde{x}_i^b(k)$ is bounded, that is

$$\Big[\tilde{x}_i^b(k)\Big]^{\mathrm{T}}\tilde{x}_i^b(k) < \sigma_i \tag{32}$$

where p_i and σ_i are finite real numbers.

Proof: Since the estimator gain $K_i(k)$ in Theorem 1 is designed in the minimum mean-squared error sense, the following inequality must hold:

$$\begin{aligned} \mathrm{Tr}\big\{P_{i,i}(k)\big\} &= \mathrm{E}\big\{\mathrm{Tr}\big\{\tilde{x}_i(k)\tilde{x}_i^{\mathrm{T}}(k)\big\}\big\} \\ &= \mathrm{E}\big\{\tilde{x}_i^{\mathrm{T}}(k)\tilde{x}_i(k)\big\} \\ &\leq \mathrm{E}\bigg\{\Big[\tilde{x}_i^b(k)\Big]^{\mathrm{T}}\Big[\tilde{x}_i^b(k)\Big]\bigg\}. \end{aligned} \tag{33}$$

If (32) holds, one has that

$$\mathrm{Tr}\big\{P_{i,i}(k)\big\} \leq \mathrm{E}\bigg\{\Big[\tilde{x}_i^b(k)\Big]^{\mathrm{T}}\Big[\tilde{x}_i^b(k)\Big]\bigg\} < \sigma_i. \tag{34}$$

Then, taking $p_i = \sigma_i + p_0$ with p_0 being a non-negative real number, the inequality in (31) can be obtained. This completes the proof. ■

Before deriving the stability results of the proposed estimator, the following matrices are defined and an important Lyapunov stability criterion for the linear time-varying system is introduced:

$$\begin{cases} K_i^g(k) = \left[I - K_i^b(k)C_i\right]A_i \\ K_i^w(k) = \left[\left(I - K_i^b(k)C_i\right)\Gamma_i \quad -K_i^b(k)\right] \\ K_i^v(k) = \begin{bmatrix} K_{i,i_1^\rho}^b(k) & K_{i,i_2^\rho}^b(k) & \cdots & K_{i,i_{\theta_i}^\rho}^b(k) \end{bmatrix} \\ K_i^\zeta(k) = \Big[\left(I - K_i^b(k)C_i\right)A_{i,i_1^r}\left(I - C_{M,i_1^r}C_{i_1^r}\right) \\ \qquad \cdots \ \left(I - K_i^b(k)C_i\right)A_{i,i_{\theta_{r_i}}^r}\left(I - C_{M,i_{\theta_{r_i}}^r}C_{i_{\theta_{r_i}}^r}\right)\Big]. \end{cases} \tag{35}$$

■

Lemma 2 [27, Th. 23-3]: The linear time-varying system $x(k+1) = A(k)x(k)$ is uniformly exponentially stable if and only if there exists a matrix sequence $\mathcal{P}(k)$ that for all k is symmetric and such that

$$\begin{cases} \eta I \leq \mathcal{P}(k) \leq \rho I \\ A^{\mathrm{T}}(k)\mathcal{P}(k+1)A(k) - \mathcal{P}(k) \leq -\nu I \end{cases} \tag{36}$$

where η, ρ, and ν are finite positive constants.

Theorem 2: Under Assumptions 3 and 4, the proposed DKF is bounded if the following linear matrix inequality (LMI) optimization problem is feasible:

$$\min_{\mathcal{P}_i, \mathcal{W}_i, \hat{K}_i, \eta_i>0, \rho_i>0, \nu_i>0, t_i>0} t_i$$
$$\begin{cases} \begin{bmatrix} \mathcal{P}_i - \mathcal{W}_i - \mathcal{W}_i^{\mathrm{T}} & \mathcal{W}_iA_i - \hat{K}_iC_iA_i \\ * & -\mathcal{P}_i + \nu_i I \end{bmatrix} < -t_i I \\ \eta_i I \leq \mathcal{P}_i \leq \rho_i I \end{cases} \tag{37}$$

where η_i, ρ_i, and ν_i are finite positive constants.

Proof: The estimation error system of $\hat{x}_i^b(k)$ can be written as

$$\tilde{\mathbf{S}}_i^b : \begin{cases} \tilde{x}_i^b(k) = \left[I - K_i^b(k)C_i\right]A_i\tilde{x}_i^b(k-1) \\ \qquad + \left[I - K_i^b(k)C_i\right]\Gamma_i w_i(k-1) - K_i^b(k)v_i(k) \\ \qquad + \sum_{i_\kappa^r \in \overline{\Omega}_{r_i}} \Big\{\left(I - K_i^b(k)C_i\right)A_{i,i_\kappa^r}\left(I - C_{M,i_\kappa^r}C_{i_\kappa^r}\right) \\ \qquad\qquad \tilde{x}_{i_\kappa^r}^b(k-1)\Big\} \\ \qquad - \sum_{i_\kappa^r \in \overline{\Omega}_{r_i}} K_{i,i_\kappa^r}^b v_{i_\kappa^r}(k-1) \\ \qquad - \sum_{i_\kappa^c \in \overline{\Omega}_{c_i}} K_{i,i_\kappa^c}^b v_{i_\kappa^c}(k-1) \\ i = 1, 2, \ldots, L. \end{cases} \tag{38}$$

According to Lemma 2, the estimation error system $\tilde{\mathbf{S}}_i^b$ is uniformly exponentially stable if and only if there exists a matrix sequence $\mathcal{P}_i(k)$ that for all k is symmetric and such that

$$\begin{cases} \left(K_i^g(k)\right)^{\mathrm{T}}\mathcal{P}_i(k+1)K_i^g(k) - \mathcal{P}_i(k) \leq -\nu_i I \\ \eta_i I \leq \mathcal{P}_i(k) \leq \rho_i I \end{cases} \tag{39}$$

By Schur complement lemma [28], the first inequality in (39) can be converted as

$$\begin{bmatrix} -\mathcal{P}_i(k+1) & \mathcal{P}_i(k+1)\left[(I - K_i^b(k)C_i)A_i\right] \\ * & -\mathcal{P}_i(k) + \nu_i I \end{bmatrix} < 0. \tag{40}$$

According to [29, Th. 1], inequality (40) can be simplified as a matrix inequality

$$\begin{bmatrix} \mathcal{P}_i(k+1) - \mathcal{W}_i(k) - \mathcal{W}_i^{\mathrm{T}}(k) & \mathcal{W}_i(k)A_i - \hat{K}_i(k)C_iA_i \\ * & -\mathcal{P}_i(k) + \nu_i I \end{bmatrix} < 0 \tag{41}$$

where $\mathcal{W}_i(k)$ and $\hat{K}_i(k)$ are introduced matrix sequences. To make the verification of the inequality condition feasible, we choose the matrix sequence $\mathcal{P}_i(k) = \mathcal{P}_i$, then the system $\tilde{\mathbf{S}}_i^b$ is uniformly exponentially stable if there exists a symmetric matrix $\mathcal{P}_i$ and matrix sequences $\hat{K}_i(k)$ and $\mathcal{W}_i(k)$ such that $\eta_i I \leq \mathcal{P}_i \leq \rho_i I$ and

$$\begin{bmatrix} \mathcal{P}_i - \mathcal{W}_i(k) - \mathcal{W}_i^{\mathrm{T}}(k) & \mathcal{W}_i(k)A_i - \hat{K}_i(k)C_iA_i \\ * & -\mathcal{P}_i + \nu_i I \end{bmatrix} < 0. \tag{42}$$

The existence of such inequality (42) can be verified by solving the following LMI optimization problem:

$$\min_{\mathcal{P}_i, \mathcal{W}_i(k), \hat{K}_i(k), \eta_i>0, \rho_i>0, \nu_i>0, t_i>0} t_i$$
$$\begin{cases} \begin{bmatrix} \mathcal{P}_i - \mathcal{W}_i(k) - \mathcal{W}_i^{\mathrm{T}}(k) & \mathcal{W}_iA_i - \hat{K}_i(k)C_iA_i \\ * & -\mathcal{P}_i + \nu_i I \end{bmatrix} < -t_i I \\ \eta_i I \leq \mathcal{P}_i \leq \rho_i I. \end{cases} \tag{43}$$

Note that all constant matrices in (43) are time invariant, and thus, the solutions of this convex optimization problem are identical at each instant. Therefore, it is proposed to solve this problem once, then the LMI optimization problem is reduced to (37).

It is known from [30] that a uniformly exponentially stable time-varying system can keep a bounded state under various bounded disturbances. Therefore, the estimation error system is rewritten as

$$\begin{aligned} \tilde{x}_i^b(k) = {} & K_i^g(k)\tilde{x}_i^b(k-1) + K_i^w(k)\bar{w}_i(k) \\ & - K_i^v(k)\bar{v}_i(k-1) + K_i^\zeta(k)\zeta_i(k-1) \end{aligned} \tag{44}$$

where

$$\begin{cases} \bar{w}_i(k) = \mathrm{col}\{w_i(k-1),\ v_i(k)\} \\ \bar{v}_i(k-1) = \mathrm{col}\left\{v_{i_1^\rho}(k-1), \ldots, v_{i_{\theta_i}^\rho}(k-1)\right\} \\ \zeta_i(k-1) = \mathrm{col}\left\{\tilde{x}_{i_1^r}^b(k-1), \ldots, \tilde{x}_{i_{\theta_{r_i}}^r}^b(k-1)\right\}. \end{cases} \tag{45}$$

From (39), one can also infer the boundedness of $K_i^g(k)$

$$\left\|K_i^g(k)\right\|_2 \leq \sqrt{\frac{\rho_i - \nu_i}{\eta_i}}. \tag{46}$$

Then, we have the boundedness of matrices $K_i^b(k)$, $K_i^w(k)$, $K_i^v(k)$ and $K_i^\zeta(k)$. Under Assumption 3, the order of each subsystem can be rearranged by Lemma 1 to obtain acyclic ordering. This acyclic ordering ensures that $i_\kappa^r (\in \overline{\Omega}_{r_i}) < i$ or $\overline{\Omega}_{r_i} = \emptyset$ for $i = 1, 2, \ldots, L$, and the estimation error systems after reordering can be concretely described as follows.

1) When $i = 1$, it can be derived that $\overline{\Omega}_{r_1} = \emptyset$. Then, the first local estimation error system $\tilde{\mathbf{S}}_1$ is reduced to

$$\begin{aligned} \tilde{x}_1^b(k) = {} & K_1^g(k)\tilde{x}_1^b(k-1) + K_1^w(k)\bar{w}_1(k) \\ & - K_1^v(k)\bar{v}_1(k-1). \end{aligned} \tag{47}$$

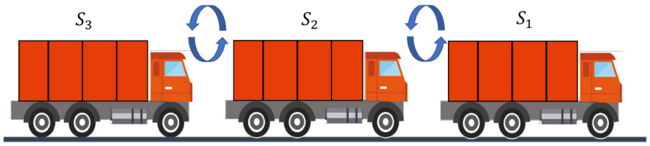


Fig. 2. Platoon of three HDVs with couplings caused by aerodynamics, where each vehicle is modeled as a subsystem S_i ($i \in \{1, 2, 3\}$).

Obviously, the feasibility of (37) can ensure the boundedness of $\|\tilde{x}_1(k)\|_2$ at each instant under Assumption 4.

2) When $i = 2$, it follows that $\overline{\Omega}_{r_2} = \emptyset$ or $\overline{\Omega}_{r_2} = \{1\}$. For the case $\overline{\Omega}_{r_2} = \emptyset$, the structure of local estimation error system $\tilde{\mathbf{S}}_2$ is the same as 1) and, thus, $\tilde{x}_2^b(k)$ is bounded at each instant. For the case $\overline{\Omega}_{r_2} = \{1\}$, one has

$$\tilde{x}_2^b(k) = K_2^g(k)\tilde{x}_2^b(k-1) + K_2^w(k)\bar{w}_2(k) - K_2^v(k)\bar{v}_2(k-1) + K_2^\zeta(k)\zeta_2(k-1). \quad (48)$$

Since the boundedness of $\tilde{x}_1^b(k)$ has been proved at each instant, it can also be concluded that the feasibility of (37) can ensure the boundedness of $\tilde{x}_2^b(k)$ at each instant for $\overline{\Omega}_{r_2} = \{1\}$ under Assumption 4.

3) When $i = 3, \dots, L$, the structure of local estimation error system $\tilde{\mathbf{S}}_i$ will be the same as (47) or (48). Then, similar to the derivation process in 1) and 2), we can derive, in turn, that the feasibility of (37) can ensure the boundedness of $\|\tilde{x}_i^b(k)\|_2$ ($i = 3, \dots, L$) at each instant.

The above analysis shows that the feasibility of (37) ensures that another estimator $\hat{x}_i^b(k)$ is bounded. By Proposition 1, one has that the proposed DKF is bounded. This completes the proof. ■

Remark 7: The LMI criterion (37) can be verified by searching for a feasible solution using the LMI Toolbox in MATLAB after the structure of the overall system is determined. Meanwhile, the finding of the feasible solution for (37) is independent of time and, thus, the stability of the designed estimator with time-varying gains can be guaranteed in advance.

IV. Simulation Examples

Consider a platoon of three heavy duty vehicles (HDVs) with physical couplings caused by aerodynamics (see Fig. 2). The dynamics of a single HDV are described by [31]

$$\begin{cases} m\dot{v} = F_{\text{engine}} - F_{\text{brake}} - F_{\text{roll}}(\alpha) - F_{\text{airdrag}}(v) \\ \qquad - F_{\text{gravity}}(\alpha) \\ \quad = k_u u - k_b F_{\text{brake}} - k_d v^2 - k_{f_r}\cos\alpha - k_g \sin\alpha \\ \dot{s} = v \end{cases} \quad (49)$$

where the notation v is the vehicle velocity, m denotes the accelerated mass, and u represents the net engine torque. k_u (k_{ui}), k_b, k_d, k_{f_r}, and k_g denote coefficients for engine, brake, air drag, road friction, and gravitation, respectively. When several vehicles are operating in a platoon formation, the coupling caused by aerodynamics is introduced to each vehicle subsystem. Concretely, the air drag coefficient k_d is modified to change with the relative distance from the preceding vehicle and the relative distance from follower vehicle. In this case, by linearizing and applying a one-step forward discretization, the discrete-time model for a platoon of three HDVs is described by

$$\begin{cases} x_1(k+1) = A_1x_1(k) + A_{12}x_2(k) + B_1u_1(k) + w_1(k) \\ x_2(k+1) = A_2x_2(k) + A_{21}x_1(k) + A_{23}x_3(k) \\ \qquad\qquad + B_2u_2(k) + w_2(k) \\ x_3(k+1) = A_3x_3(k) + A_{32}x_2(k) + B_3u_3(k) + w_3(k) \end{cases} \quad (50)$$

where $x_1 \triangleq \text{col}\{\mathrm{v}_0, \mathrm{v}_1\}$, $x_2 \triangleq \text{col}\{d_{12}, v_2\}$ and $x_3 \triangleq \text{col}\{d_{23}, v_3\}$, respectively, represent the state of each subsystem, v_0 is the reference speed of the formation vehicles, v_i ($i = 1, 2, 3$) is the speed of individual vehicle, and d_{12}, d_{23} are the relative distances among the vehicles. At the same time, u_i denotes the control signal for each subsystem, while $w_i(k)$ is the subsystem noise, and

$$\begin{cases} A_1 = \begin{bmatrix} 1 & 0 \\ 0 & \Theta_1 \end{bmatrix}, \ A_{12} = \begin{bmatrix} 0 & 0 \\ \gamma_2 & 0 \end{bmatrix} \\ A_2 = \begin{bmatrix} 1 & -1 \\ \delta_2 & \Theta_2 \end{bmatrix}, \ A_{21} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \ A_{23} = \begin{bmatrix} 0 & 0 \\ \gamma_3 & 0 \end{bmatrix} \\ A_3 = \begin{bmatrix} 1 & -1 \\ \delta_3 & \Theta_3 \end{bmatrix}, \ A_{32} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \\ B_1 = \begin{bmatrix} 0 \\ k_{u1} \end{bmatrix}, \ B_2 = \begin{bmatrix} 0 \\ k_{u2} \end{bmatrix}, \ B_3 = \begin{bmatrix} 0 \\ k_{u3} \end{bmatrix} \end{cases} \quad (51)$$

with the velocity control coefficient k_{ui} and velocity attenuation coefficients Θ_i, γ_i, and δ_i. The measurement for the interconnected HDV system is given by

$$\begin{cases} y_1(k) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x_1(k) + v_1(k) \\ y_2(k) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x_2(k) + v_2(k) \\ y_3(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} x_3(k) + v_3(k) \end{cases} \quad (52)$$

where $y_1, y_2 \in \mathbb{R}^2$, and $y_3 \in \mathbb{R}$ are the measurements of subsystems, while $v_i(k)$ is the measurement noise of each subsystem. By Assumption 1, the communication with minimum information exchange is taken into account, where it is assumed that there exist the communication links between the first vehicle and the second vehicle, and between the second vehicle and the third vehicle.

According to the parameters used in [32], when the mass of each identical vehicle is set as $m = 40000$ kg, we choose $\Theta_i = 0.9999$, $\gamma_i = \delta_i = 0.1476$, $k_{ui} = 1.48$ for the interconnected HDV system, while the system noises $w_i(k)$ and $v_i(k)$ are set as Gaussian white noises with covariances $Q_{w_1} = \text{diag}\{0, 0.001\}, Q_{w_2} = \text{diag}\{0.001, 0.001\}, Q_{w_3} = \text{diag}\{0.001, 0.001\}$ and $Q_{v1} = \text{diag}\{0, 0.01\}$, $Q_{v2} = \text{diag}\{0.01. \ 0.01\}$, and $Q_{v3} = \text{diag}\{0.01\}$, respectively. Meanwhile, the centralized LQG control scheme is used to minimize the overall fuel consumption of the formation vehicles and keep safe relative distances, as well as maintaining the lead vehicle at a set reference velocity. With the optimization objective and desired relative distance defined in [31], the

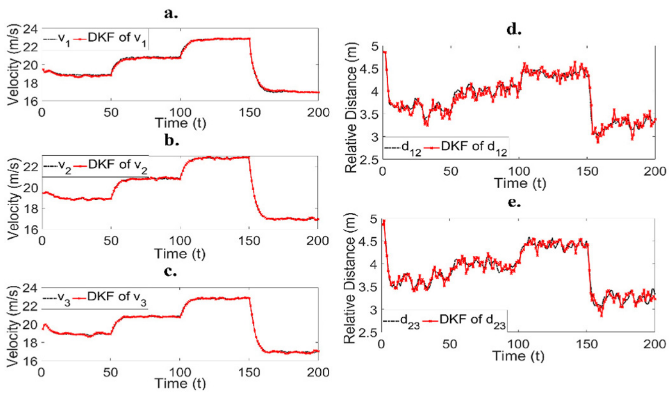


Fig. 3. (a)–(c) Velocity of each HDV and its estimated values. (d) Relative distance between the first two vehicles and its estimated value. (e) Relative distance between the last two vehicles and its estimated value.

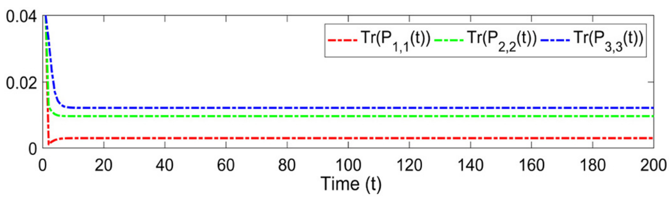


Fig. 4. Estimation performance of the DKF designed by Theorem 1.

control input $u(k) \triangleq \mathrm{col}\{u_1(k), u_2(k), u_3(k)\}$ is calculated as

$$u(k) = \begin{bmatrix} 0.19 & -0.52 & -0.18 & 0.23 & 0.01 & 0.11 \\ 0.14 & 0.21 & 0.05 & -0.59 & -0.14 & 0.22 \\ 0.09 & 0.16 & 0.11 & 0.17 & 0.09 & -0.49 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}. \tag{53}$$

The proposed distributed Kalman filtering method is successfully applied to the interconnected HDV system. To demonstrate the effectiveness of the theoretical results, the velocity of each HDV and its estimated value are plotted in Fig. 3(a)–(c), and the relative distances among HDVs and their estimated values are plotted in Fig. 3(d)–(e). It can be seen from this figure that the designed DKF can track these state variables well based on the information of local measurements and in-neighbors' measurements and estimates. Besides, the performance of the DKF that was evaluated by the trace of estimation error covariance of each subsystem is plotted in Fig. 4. The result shows that the estimation error in all subsystems, though with unobservable states and coupling terms, can converge to a satisfactory level within a few running steps. Meanwhile, due to the random noises, the estimation performance of the proposed DKF is accessed by mean-square errors (MSEs) that are computed by the Monte Carlo method with an average of 100 runs. The MSEs of the proposed DKF are plotted in Fig. 5, compared with the DKF in the minimum mean-squared error sense for sequential systems [17]. The result shows that the estimation performance of the proposed DKF is much better than the DKF in [17] when the system structure is not sequential. On the other hand, the decoupling and reordering of the interconnected HDV system can be seen

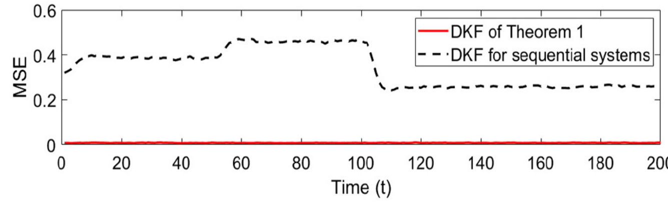


Fig. 5. MSE performance comparison of the proposed DKF and the DKF for sequential systems in [17].

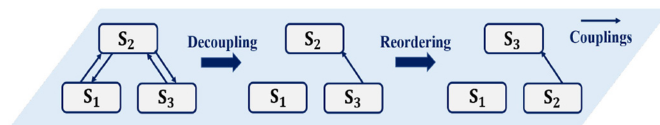


Fig. 6. Decoupling and reordering of the interconnected HDV system.

from Fig. 6. By our decoupling strategy, the adjacent matrix of the estimation error relationship graph $\mathcal{A}_e$ is reduced from $\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ to $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$. By reordering subsystems into an acyclic ordering (i.e., exchange the orders of the second and third subsystems), this adjacent matrix can be further transformed as an upper triangular matrix $\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$ and, thus, the stability analysis in [1)–3)] in the proof of Theorem 2 can be conducted. Moreover, the LMI feasible problem (37) for each subsystem is solved in advance, then the stability of DKF is verified by Theorem 2. Moreover, the effectiveness of the proposed stability conditions can be shown from the observation in Fig. 5 that the MSEs of the designed DKF are bounded.

V. Conclusion

In this article, the distributed Kalman filtering problem has been investigated for large-scale interconnected systems. A Kalman-like recursive distributed estimator was proposed based on the information of local measurements and in-neighbors' measurements and estimates, together with the decoupling strategy that can minimize the impact of interconnected terms for the estimation error system. When considering Gaussian white noises with known covariances, a locally optimal estimator gain was designed in the linear minimum mean-squared error sense, and stability conditions were derived such that the proposed DKF is bounded. Finally, an interconnected HDV system was given to show the effectiveness and advantages of the proposed method.

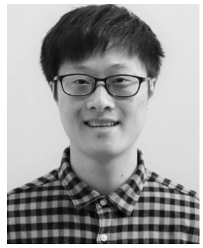
References

[1] D. D. Šiljak, *Large-Scale Dynamic Systems: Stability and Structure*, vol. 2. Amsterdam, The Netherlands: North Holland, 1978.

[2] G. Antonelli, "Interconnected dynamic systems: An overview on distributed control," *IEEE Control Syst. Mag.*, vol. 33, no. 1, pp. 76–88, Feb. 2013.

[3] H. Cai and G. Hu, "Distributed tracking control of an interconnected leader–follower multiagent system," *IEEE Trans. Autom. Control*, vol. 62, no. 7, pp. 3494–3501, Jul. 2017.

[4] W. Li, Y. Jia, and J. Du, “State estimation for stochastic complex networks with switching topology,” *IEEE Trans. Autom. Control*, vol. 62, no. 12, pp. 6377–6384, Dec. 2017.
[5] V. Kekatos and G. B. Giannakis, “Distributed robust power system state estimation,” *IEEE Trans. Power Syst.*, vol. 28, no. 2, pp. 1617–1626, May 2013.
[6] Y. Li, S. Tong, and T. Li, “Adaptive fuzzy output feedback dynamic surface control of interconnected nonlinear pure-feedback systems,” *IEEE Trans. Cybern.*, vol. 45, no. 1, pp. 138–149, Jan. 2015.
[7] B. Khaleghi, A. Khamis, F. O. Karray, and S. N. Razavi, “Multisensor data fusion: A review of the state-of-the-art,” *Inf. Fusion*, vol. 14, no. 1, pp. 28–44, 2013.
[8] E. Yeung, J. Gonçalves, H. Sandberg, and S. Warnick, “The meaning of structure in interconnected dynamic systems,” 2011. [Online]. Available: arXiv:1108.2755.
[9] R. Olfati-Saber, “Distributed kalman filtering for sensor networks,” in *Proc. 46th IEEE Conf. Decis. Control*, 2007, pp. 5492–5498.
[10] Y. Xu, R. Lu, P. Shi, H. Li, and S. Xie, “Finite-time distributed state estimation over sensor networks with round-robin protocol and fading channels,” *IEEE Trans. Cybern.*, vol. 48, no. 1, pp. 336–345, Jan. 2018.
[11] H. Yang, H. Li, Y. Xia, and L. Li, “Distributed kalman filtering over sensor networks with transmission delays,” *IEEE Trans. Cybern.*, early access, Apr. 3, 2020, doi: 10.1109/TCYB.2020.2980582.
[12] C. Sanders, E. Tacker, T. Linton, and R. Ling, “Specific structures for large-scale state estimation algorithms having information exchange,” *IEEE Trans. Autom. Control*, vol. 23, no. 2, pp. 255–261, Apr. 1978.
[13] C. Sanders, E. Tacker, and T. Linton, “A new class of decentralized filters for interconnected systems,” *IEEE Trans. Autom. Control*, vol. 19, no. 3, pp. 259–262, Jun. 1974.
[14] 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.
[15] A. Haber and M. Verhaegen, “Moving horizon estimation for large-scale interconnected systems,” *IEEE Trans. Autom. Control*, vol. 58, no. 11, pp. 2834–2847, Nov. 2013.
[16] U. A. Khan and J. M. Moura, “Distributing the kalman filter for large-scale systems,” *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4919–4935, Oct. 2008.
[17] B. Chen, G. Hu, D. W. Ho, and L. Yu, “Distributed kalman filtering for time-varying discrete sequential systems,” *Automatica*, vol. 99, pp. 228–236, Jan. 2019.
[18] J. Sijs, U. Hanebeck, and B. Noack, “An empirical method to fuse partially overlapping state vectors for distributed state estimation,” in *Proc. Eur. Control Conf. (ECC)*, 2013, pp. 1615–1620.
[19] J. Sijs, M. Lazar, and P. P. J. V. D. Bosch, “State fusion with unknown correlation: Ellipsoidal intersection,” in *Proc. Amer. Control Conf.*, 2010, pp. 3992–3997.
[20] T. Cline, “Near-optimal state estimation for interconnected systems,” *IEEE Trans. Autom. Control*, vol. 20, no. 3, pp. 348–351, Jun. 1975.
[21] B. Chen, G. Hu, D. W. Ho, and L. Yu, “Distributed estimation for discrete-time interconnected systems,” in *Proc. Chin. Control Conf. (CCC)*, 2019, pp. 3708–3714.
[22] S. Riverso, M. Farina, and G. Ferrari-Trecate, “Plug-and-play decentralized model predictive control for linear systems,” *IEEE Trans. Autom. Control*, vol. 58, no. 10, pp. 2608–2614, Oct. 2013.
[23] S. Riverso, M. Farina, and G. Ferrari-Trecate, “Plug-and-play model predictive control based on robust control invariant sets,” *Automatica*, vol. 50, no. 8, pp. 2179–2186, 2014.
[24] J. Hu, Z. Wang, S. Liu, and H. Gao, “A variance-constrained approach to recursive state estimation for time-varying complex networks with missing measurements,” *Automatica*, vol. 64, pp. 155–162, Feb. 2016.
[25] F. Harary, R. Z. Norman, and D. Cartwright, *Structural Models: An Introduction to the Theory of Directed Graphs*. New York, NY, USA: Wiley, 1965.
[26] Á. F. Garcia-Fernandez, M. R. Morelande, and J. Grajal, “Truncated unscented kalman filtering,” *IEEE Trans. Signal Process.*, vol. 60, no. 7, pp. 3372–3386, Jul. 2012.
[27] W. J. Rugh, *Linear System Theory*. Hoboken, NJ, USA: Prentice-Hall, 1996.
[28] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*. Philadelphia, PA, USA: SIAM, 1994.
[29] M. C. De Oliveira, J. Bernussou, and J. C. Geromel, “A new discrete-time robust stability condition,” *Syst. Control Lett.*, vol. 37, no. 4, pp. 261–265, 1999.
[30] B. D. Anderson, “Internal and external stability of linear time-varying systems,” *SIAM J. Control Optim.*, vol. 20, no. 3, pp. 408–413, 1982.
[31] H. R. Feyzmahdavian, A. Alam, and A. Gattami, “Optimal distributed controller design with communication delays: Application to vehicle formations,” in *Proc. IEEE 51st IEEE Conf. Decis. Control (CDC)*, 2012, pp. 2232–2237.
[32] Y. Wang, J. Xiong, and D. W. C. Ho, “Decentralized control scheme for large-scale systems defined over a graph in presence of communication delays and random missing measurements,” *Automatica*, vol. 98, pp. 190–200, Dec. 2018.



Yuchen Zhang (Graduate Student Member, IEEE) received the B.E. degree in automation from the Zhejiang University of Technology, Hangzhou, China, in 2017, where he is currently pursuing the Ph.D. degree in control science and engineering.

His current research interests include distributed estimation of interconnected systems and fusion estimation of binary sensor systems.

Bo Chen (Member, IEEE) received the B.S. degree in information and computing science from the Jiangxi University of Science and Technology, Ganzhou, China, in 2008, and the Ph.D. degree in control theory and control engineering from the Zhejiang University of Technology, Hangzhou, China, in 2014.

He is currently a Professor with the Department of Automation, Zhejiang University of Technology. He was a Research Fellow with the School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore, from 2014 to 2015 and from 2017 to 2018. He was also a Postdoctoral Research Fellow with the Department of Mathematics, City University of Hong Kong, Hong Kong, from 2015 to 2017. His current research interests include information fusion, distributed estimation and control, cyber-physical systems security, and networked fusion systems.

Prof. Chen was a recipient of the Outstanding Thesis Award of the Chinese Association of Automation in 2015.

Li Yu (Member, IEEE) received the B.S. degree in control theory from Nankai University, Tianjin, China, in 1982, and the M.S. and Ph.D. degrees from Zhejiang University, Hangzhou, China, in 1988 and 1999, respectively.

He is currently a Professor with the College of Information Engineering, Zhejiang University of Technology, Hangzhou. He has authored or coauthored three books and over 200 journal papers. His current research interests include cyber-physical systems security, networked control systems, motion control, and information fusion.

Daniel W. C. Ho (Fellow, IEEE) received the B.S., M.S., and Ph.D. degrees in mathematics from the University of Salford, Greater Manchester, U.K., in 1980, 1982, and 1986, respectively.

From 1985 to 1988, he was a Research Fellow with the Industrial Control Unit, University of Strathclyde, Glasgow, U.K. In 1989, he joined the City University of Hong Kong, Hong Kong, where he is currently a Chair Professor of applied mathematics and an Associate Dean with the College of Science. He has over 250 publications in scientific journals. His current research interests include control and estimation theory, complex dynamical distributed networks, multiagent networks, and stochastic systems.

Dr. Ho was a recipient of the Chang Jiang Chair Professor Awarded by the Ministry of Education, China, in 2012, and the ISI Highly Cited Researchers Award in Engineering by Clarivate Analytics, from 2014 to 2020. He has been on the Editorial Board of a number of journals, including the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, *IET Control Theory and Applications*, the *Journal of the Franklin Institute*, and the *Asian Journal of Control*.