Automatica 161 (2024) 111421

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Split-spectrum based distributed state estimation for linear systems☆

Lili Wang [a,*], Ji Liu [b], Brian D.O. Anderson [c], A. Stephen Morse [d]

[a] College of Engineering, Purdue University, United States of America
[b] Department of Electrical and Computer Engineering, Stony Brook University, United States of America
[c] School of Engineering, Australian National University, Australia
[d] Department of Electrical Engineering, Yale University, United States of America

ARTICLE INFO

Article history:
Received 1 October 2022
Received in revised form 24 October 2023
Accepted 27 October 2023
Available online 14 December 2023

Keywords:
Distributed estimation
Multi-agent systems
Cooperative control
Linear systems

ABSTRACT

This paper studies a distributed state estimation problem for both continuous- and discrete-time linear systems. A simply structured distributed estimator {comprising interconnected local estimators} is first described for estimating the state of a continuous and multi-channel linear system whose sensed outputs are distributed across a fixed multi-agent network. The estimator is then extended to non-stationary networks whose graphs switch according to a switching signal. The estimator is guaranteed to solve the problem, provided a network-widely shared high gain condition achieving a form of spectrum separation is satisfied. As an alternative to sharing a common gain across the network, a fully distributed version of the estimator is also studied in which each agent adaptively adjusts a local gain, though the practicality of this approach is subject to a robustness issue common to adaptive control. A discrete-time version of the distributed state estimation problem is also studied, and a corresponding estimator based again on spectrum separation, but not high gain, is proposed for time-varying networks. For each scenario, it is explained how to construct the estimator so that the state estimation errors in the local estimators all converge to zero exponentially fast at a fixed but arbitrarily chosen rate, provided the network's graph is strongly connected for all time. The proposed estimators are inherently resilient to abrupt changes in the number of agents and communication links in the inter-agent communication graph upon which the algorithms depend, provided the network is redundantly strongly connected and redundantly jointly observable.

1. Introduction

With the growing interest in sensor networks and multi-agent systems, the problem of estimating the state of a dynamical system whose measured outputs are distributed across a network has been under study in one form or another for a number of years (Dorfler, Pasqualetti, & Bullo, 2013; Han, Trentelman, Wang, & Shen, 2018, 2019; Khan & Jadbabaie, 2011; Kim, Shim, & Cho, 2016; Li, Phillips, & Sanfelice, 2018; Mitra & Sundaram, 2018; Olfati-Reza, 2009; Olfati-Reza & Shamma, 2005; Park & Martins, 2017; Wang, Liu, & Morse, 2019; Wang & Morse, 2018; Wang, Morse, Fullmer, & Liu, 2017; Xiao, Boyd, & Lall, 2005).

Depending on the nature of the system to be estimated, the distributed estimation problem has continuous- and discrete-time versions. In its simplest form, the continuous-time version of the distributed state estimation problem starts with a network of $m > 1$ agents labeled $1, 2, \ldots, m$ which are able to receive information from their neighbors. Neighbor relations are characterized by a directed graph $\mathbb{N}$, which may or may not depend on time, whose vertices correspond to agents and whose arcs depict neighbor relations. Each agent i senses a signal $y_i \in \mathbb{R}^{s_i}$, $i \in \mathbf{m} \triangleq \{1, 2, \ldots, m\}$ generated by a continuous-time system of the form $\dot{x} = Ax$, $y_i(t) = C_i x$, $i \in \mathbf{m}$ and $x \in \mathbb{R}^n$. It is typically assumed that $\mathbb{N}$ is strongly connected and that the system is jointly observable. It is invariably assumed that each agent receives certain real-time signals from its neighbors although what is received can vary from one problem formulation to the next. In all formulations, the goal is to devise local estimators, one for each agent, whose outputs are all asymptotically correct estimates of x. The local estimator dynamics for agent i are typically assumed to depend only on the pair (C_i, A) and certain properties of $\mathbb{N}$. The problem is basically the same in discrete time, except that rather than the

☆ This work was supported in part by the Air Force Office of Scientific Research, United States under award number FA9550-23-1-0175 and by the National Science Foundation, United States under Grant No. 2230101. The material in this paper was partially presented at the 58th IEEE Conference on Decision and Control, December 11–13, 2019, Nice, France. This paper was recommended for publication in revised form by Associate Editor Luca Schenato under the direction of Editor Christos G. Cassandras.

* Corresponding author.
E-mail addresses: wang6127@purdue.edu (L. Wang), ji.liu@stonybrook.edu (J. Liu), brian.anderson@anu.edu.au (B.D.O. Anderson), as.morse@yale.edu (A.S. Morse).

https://doi.org/10.1016/j.automatica.2023.111421
0005-1098/© 2023 Published by Elsevier Ltd.

continuous-time model just described, the discrete-time model $x(\tau+1)=Ax(\tau),\ y_i(\tau)=C_ix(\tau),\ i\in\mathbf{m},\ x\in\mathbb{R}^n$ is considered instead. More precise problem formulations will be given later.

1.1. Background

The study of distributed state estimation for linear systems can be dated back to the so-called distributed Kalman filter problem (Olfati-Saber, 2005), which involves system and measurement noise in the problem formulation and has been widely studied for years (Khan & Jadbabaie, 2011; Olfati-Saber, 2007). Most available Kalman filter based approaches (Khan & Jadbabaie, 2011; Olfati-Reza, 2009; Olfati-Reza & Shamma, 2005; Olfati-Saber, 2005, 2007) require the agents to both share "signal information", which can be measurements or local state estimates, and fuse certain "structural information", which forms the covariance or information matrix of the nominal centralized Kalman filter for each agent. For the problem just described, the existing literature based on only "signal information" sharing can be classified into two categories, namely continuous- and discrete-time estimators/observers, except for our earlier work of Wang et al. (2017), Wang, Morse, and Liu (2022) in which a hybrid observer was proposed for a continuous-time linear system.

Continuous-time distributed estimators have recently appeared in Han et al. (2018, 2019), Kim et al. (2016), Wang, Liu, and Morse (2019), Wang and Morse (2017, 2018). By recasting and then solving the distributed estimation problem as a classical decentralized control problem, the resulting estimator becomes capable of estimating the state at a pre-assigned exponentially fast rate, assuming $\mathbb{N}$ is a constant strongly connected graph (Wang & Morse, 2018). The work of Kim et al. (2016) seeks to propose a distributed estimator for a continuous-time system at the expense of certain design flexibility. This is done, in essence, by exploiting the A-invariance of the unobservable spaces of the pairs (C_i, A); this in turn enables one to "split" the local estimators into two parts, one based on conventional spectrum assignment techniques applied to the observable part of the state at each local estimator and the other based on consensus among the unobserved parts of the state at each local estimator. The two parts are interacting but the use of a high gain serves to simplify the stability issue because of a split in the spectrum arising from the design of the estimator. The idea has been further developed in Han et al. (2018, 2019), Kim, Lee, and Shim (2020). Specifically, these latter references start to move beyond a restriction in Kim et al. (2016) permitting only constant, undirected, connected neighbor graphs to be addressed. The work of Han et al. (2018, 2019) extends the result of Kim et al. (2016) to the case when the neighbor graph is constant, directed, strongly connected, while requiring that one chooses gains to ensure that certain LMIs hold which are difficult to grasp intuitively. In Lee and Shim (2020), motivated by a distributed least squares problem, a modified algorithm is proposed to deal with measurement noise constant, undirected, connected neighbor graphs. A distributed adaptive algorithm has recently been proposed in Kim et al. (2020) which allows agents to join or leave the network over time, provided the resulting agent network always remains jointly connected and joint detectable. An evident disadvantage of all these existing continuous-time distributed estimators is that they require a somewhat complicated gain computation procedure, and partially because of this, do not, at least not directly, admit discrete-time counterparts.

Discrete-time distributed estimators have been recently studied in Açikmeşe, Mandić, and Speyer (2014), Doostmohammadian and Khan (2013), Khan and Jadbabaie (2014), Mitra, Richards, Bagchi, and Sundaram (2019), Mitra and Sundaram (2018), Park and Martins (2012a, 2012b, 2017), Rego et al. (2021), Ugrinovskii (2013). Notable among them is the paper (Park & Martins, 2017). Published prior to the appearance of the early continuous time paper (Wang & Morse, 2018) applying to the same class of distributed systems, Park and Martins (2017) solves the discrete-time distributed estimation problem for jointly observable, linear systems with constant, directed, strongly connected neighbor graphs. It builds on the idea of recasting the estimation problem as a classical decentralized control problem. Although the observer is limited to discrete-time systems, it has been proved possible to make use of the ideas in Park and Martins (2017) to obtain, as noted earlier, a distributed observer for continuous-time systems (Wang & Morse, 2018), but still for constant neighbor graphs. There are however other discrete-time distributed observers/estimators which do not admit continuous-time extensions, illustrating that passage between discrete-time and continuous-time thinking may be harder than intuition initially suggests for distributed estimation problems. By expanding on earlier work in Mitra and Sundaram (2018), the papers (Mitra et al., 2019; Mitra, Richards, Bagchi, & Sundaram, 2022) provide a procedure for constructing a centralized designed distributed observer for time-varying neighbor graphs. It requires the sharing of an index that records the age of the information across the network, and the agents are designed to act in a sequential manner to do state estimation. The resulting algorithm, which is tailored exclusively to discrete-time systems, requires a network-wide initialization step that serves to sort the agents in a specific order. Thereby it can deal with state estimation under assumptions that are weaker than strong connectivity.

Different approaches to the distributed state estimation problem are summarized in Table 1. It turns out that the current paper is the first paper that can deal with both continuous-time and discrete-time systems while ensuring exponential convergence under time-varying neighbor graphs.

The contribution of this paper rests on the following three distinguishing features, differentiating it and highlighting it as a development of earlier work:

- The paper describes a simply structured, unified approach to the distributed state estimation problem and to design and analyze the corresponding distributed estimators for both continuous- and discrete-time linear systems with possibly time-varying graphs. It is termed the "split-spectrum" approach because it "splits" the system spectrum into disjoint subsets corresponding to observable and unobservable subspaces. In continuous time, this is achieved by a high gain mechanism, but in discrete-time by a different mechanism, viz. the adoption of two integrally related sampling rates. Though the mechanisms are instrumentally different, their purpose is fundamentally the same. It is termed 'unified' because the approach is shown to work for both continuous- and discrete-time linear systems over both constant and time-varying neighbor graphs.
- A fully distributed version of the estimator is separately studied where each agent can adaptively adjust a local gain, with simpler gain computation procedure and analysis compared with Kim et al. (2020).
- Exponential convergence of the error dynamics is ensured with an arbitrarily assigned convergence rate. A great advantage of our methodology is that the designs and algorithms developed under a noiseless assumption are then necessarily tolerant of some level of noise, simply because we take care to ensure an exponential convergence.

It is assumed in this paper that the neighbor graph of the network is always strongly connected. From the perspective of the real world, requiring the underlying network to be strongly connected "at every time step" is an assumption that will occur

Table 1
Comparison of different approaches to design distributed estimators.

Nature of approach	Reference	Continuous-time systems	Discrete-time systems	Exponential convergence	Time-varying graphs
Kalman Filter Based Approach	Khan and Jadbabaie (2011), Olfati-Reza (2009), Olfati-Saber (2005, 2007)	✗	✓	✗	✗
	Olfati-Reza and Shamma (2005)	✓	✗	✗	✗
Observability Decomposition Based Approach	Mitra and Sundaram (2018)	✗	✓	✓	✗
	Mitra et al. (2019, 2022)	✗	✓	✓	✓
Decentralized Control Based Approach	Park and Martins (2012a, 2012b, 2017)	✗	✓	✓	✗
	Wang and Morse (2017, 2018)	✓	✗	✓	✗
Split-Spectrum Based Approach	Han et al. (2018, 2019), Kim et al. (2016), Wang, Liu, and Morse (2019)	✓	✗	✓	✗
	Wang, Liu, Morse, and Anderson (2019)	✗	✓	✓	✗
	This work	✓	✓	✓	✓

in a great many (though obviously not all) cases, and as such, is deserving of a separate study in its own right. The extension to more general time-varying graphs is one future direction. It may not be conceptually difficult, however intricate the details may be.

The paper first describes the split-spectrum based distributed estimator for the case when the system dynamics are continuous with a stationary network in Section 2.1, and with associated background analysis is given in Section 2.2. The estimator is then extended to deal with non-stationary networks whose neighbor graphs switch according to a switching signal with a fixed dwell time or a variable dwell time with prescribed average, the ideas being detailed in Section 2.3. In the case when the interconnection among the agents can always be modeled using doubly stochastic matrices {e.g., undirected graphs with the Metropolis weights (Xiao et al., 2005)}, it is shown in Section 2.3 that the estimator functions correctly even if the neighbor graph switches arbitrarily, provided the graph is always strongly connected. The estimators mentioned above all rely on the existence of a sufficiently large, network-widely shared gain. A fully distributed version of the estimator is then studied in Section 2.4 where each agent can adaptively adjust a local gain. The adaptive estimator is subject to a robustness issue common to adaptive control. The proposed estimators, except for the adaptive one, are inherently resilient to abrupt changes in the number of agents and communication links in the inter-agent communication graph upon which the algorithms depend, an issue which is discussed in Section 2.5. Then the split-spectrum based estimator design is extended to the case when the system dynamics is discrete in Section 3 for both constant and time-varying neighbor graphs. Simulation validation is provided in Section 4.

The material in this paper was partially presented in Wang, Liu, and Morse (2019), Wang, Liu, Morse, and Anderson (2019), but this paper presents a more comprehensive treatment of the work. Specifically, the paper crafts continuous-time distributed estimators for two types of non-stationary networks in Section 2.3 and a fully distributed adaptive estimator in Section 2.4, which were not included in Wang, Liu, and Morse (2019), Wang, Liu, Morse, and Anderson (2019).

2. Continuous-time distributed estimator

We are interested in a network of $m > 0$ {possibly mobile} autonomous agents labeled $1, 2, \ldots, m$ which are able to receive information from their "neighbors", where by a *neighbor* of agent i is meant any other agent within agent i's reception range. We write $\mathcal{N}_i(t)$ for the labels of agent i's neighbors at time $t \in [0, \infty)$ and always take agent i to be a neighbor of itself. Neighbor relations at time t are characterized by a directed graph $\mathbb{N}(t)$ with m vertices and a set of arcs defined so that there is an arc in $\mathbb{N}(t)$ from vertex j to vertex i whenever agent j is a neighbor of agent i at time t. Since each agent i is always a neighbor of itself, $\mathbb{N}(t)$ has a self-arc at each of its vertices. Each agent i can sense a continuous-time signal $y_i \in \mathbb{R}^{s_i}$, $i \in \mathbf{m} \stackrel{\Delta}{=} \{1, 2, \ldots, m\}$, where

$$y_i = C_i x, \quad i \in \mathbf{m} \tag{1}$$

$$\dot{x} = Ax \tag{2}$$

and $x \in \mathbb{R}^n$. We assume throughout that $C_i \neq 0$, $i \in \mathbf{m}$, and that the system defined by (1) and (2) is *jointly observable*; i.e., with $C = \begin{bmatrix} C_1' & C_2' & \cdots & C_m' \end{bmatrix}'$, the matrix pair (C, A) is observable. Joint observability is equivalent to the requirement that $\bigcap_{i\in\mathbf{m}} \mathcal{V}_i = 0$, where $\mathcal{V}_i$ is the *unobservable space* of (C_i, A); i.e. $\mathcal{V}_i = \ker \begin{bmatrix} C_i' & (C_iA)' & \cdots & (C_iA^{n-1})' \end{bmatrix}'$. As is well known, $\mathcal{V}_i$ is the largest A-invariant subspace contained in the kernel of C_i. Generalizing the results that follow to the case when (C, A) is only detectable is quite straightforward and can be accomplished using well-known ideas. However, the commonly made assumption that each pair (C_i, A), $i \in \mathbf{m}$, is observable, or even just detectable, is very restrictive, grossly simplifies the problem and is unnecessary. The assumption $C_i \neq 0$ is not necessary provided the more relaxed problem is properly formulated. The assumption is made for the sake of simplicity. The problem of interest is to construct a suitably defined family of linear estimators in such a way so that no matter what the estimators' initial states are, each agent obtains an asymptotically correct estimate x_i of x in the sense that the estimation error $x_i(t) - x(t)$ converges to zero as fast as $\exp(-\lambda t)$ does, where λ is an arbitrarily chosen but fixed positive number.

This section proposes the estimator first, and then analyzes the estimator beginning with the condition that the neighbor graph $\mathbb{N}(t)$ is constant. A time-varying neighbor graph $\mathbb{N}(t)$ is then considered in which changes occur according to a switching signal. Later, a fully distributed algorithm based on use of multiplicative adaptive gain control is developed.

2.1. The estimator

The estimator to be considered consists of m local or private estimators of the form for each $i \in \mathbf{m}$,

$$\dot{x}_i = (A + K_iC_i)x_i - K_iy_i - gP_i\left(x_i - \frac{1}{m_i(t)}\sum_{j\in\mathcal{N}_i(t)} x_j\right) \tag{3}$$

where $m_i(t)$ is the number of labels in $\mathcal{N}_i(t)$, g is a suitably defined positive gain common to all local estimators, each K_i is a suitably defined matrix, and for each $i \in \mathbf{m}$, P_i is the orthogonal projection on the unobservable space of (C_i, A). The term $(A + K_iC_i)x_i - K_iy_i$ is designed to enable each agent i to be able to recover the observable part of the state by itself, and the term $-gP_i(x_i - \frac{1}{m_i(t)}\sum_{j\in\mathcal{N}_i(t)} x_j)$ is for the purpose of utilizing information from its neighbors to recover the unobservable part of the state. It will be shown that with this design, the spectrum of the system matrix can be split into two subsets. Details on how to design the parameters K_i and g will be provided in the following.

2.1.1. The error dynamics

This subsection provides the error dynamics of the proposed estimator. It will be shown that the spectrum of the error system matrix can be split into two subsets based on the observability of each agent with K_i defined properly.

First note from (3), that the state estimation error $e_i = x_i - x$ satisfies

$$\dot{e}_i = (A + K_iC_i)e_i - gP_i\left(e_i - \frac{1}{m_i(t)}\sum_{j\in\mathcal{N}_i(t)} e_j\right) \tag{4}$$

Consequently the overall error vector $e = \begin{bmatrix} e_1' & \dots & e_m' \end{bmatrix}'$ satisfies

$$\dot{e} = (\bar{A} - gP(I_{mn} - \bar{S}(t)))e \tag{5}$$

where $\bar{A} =$ block diag $\{A + K_1C_1, A + K_2C_2, \dots, A + K_mC_m\}$, $P =$ block diag $\{P_1, P_2, \dots, P_m\}$, $\bar{S}(t) = S(t) \otimes I_n$ with $S(t) = D_{\mathbb{N}(t)}^{-1}\mathcal{A}_{\mathbb{N}(t)}'$. Here I_k is the $k \times k$ identity matrix, and $\mathcal{A}_{\mathbb{N}(t)}$ is the adjacency matrix of $\mathbb{N}(t)$ and $D_{\mathbb{N}(t)}$ is the diagonal matrix whose ith diagonal entry is the in-degree of $\mathbb{N}(t)$'s ith vertex. Note that $\mathbb{N}(t)$ is the graph[1] of $S'(t)$ and that the diagonal entries of $S'(t)$ are all positive because each agent is a neighbor of itself. The matrix $S(t)$ is evidently a stochastic matrix.

Proposition 1. *The spectrum of the error system matrix $\bar{A} - gP(I_{mn} - \bar{S})$ splits into two subsets. One subset contains the union of certain subsets of the eigenvalues associated with the ith local estimator, $i \in \mathbf{m}$, these being able to be arbitrarily positioned by choice of the K_i and independent of g. The second subset is independent of the choice of the K_i, and depends, though not to the extent of being able to be arbitrarily positioned, on g.*

We remark that in the proof below, we will explain how to choose the K_i to ensure that the associated set of eigenvalues has degree of stability λ (ensuring an estimation error decay at least as fast as $\exp(-\lambda t)$), while subsequently we will explain how to choose g to ensure stability of the remaining part of the spectrum with the same minimum exponential decay rate.

Proof of Proposition 1. The definitions of K_i and g begin with the specification of a desired convergence rate bound $\lambda > 0$. To begin with, each matrix K_i is defined as follows. For each fixed $i \in \mathbf{m}$, write Q_i for any full rank matrix whose kernel is the unobservable space of (C_i, A) and let $\bar{C}_i$ and $\bar{A}_i$ be the unique solutions to $\bar{C}_iQ_i = C_i$ and $Q_iA = \bar{A}_iQ_i$ respectively. Then the matrix pair $(\bar{C}_i, \bar{A}_i)$ is observable. A matrix $\bar{K}_i$ can be chosen to ensure that the convergence of $\exp\{(\bar{A}_i + \bar{K}_i\bar{C}_i)t\}$ to zero is as fast as the convergence of $\exp(-\lambda t)$ to zero is. There are several well-documented ways to do this {e.g, spectrum assignment algorithms or Riccati equation solvers}, since each pair $(\bar{C}_i, \bar{A}_i)$ is observable. Having chosen such $\bar{K}_i$, K_i is then chosen to be $K_i = Q_i^{-1}\bar{K}_i$ where Q_i^{-1} is a right inverse for Q_i. The definition implies that

$$Q_i(A + K_iC_i) = (\bar{A}_i + \bar{K}_i\bar{C}_i)Q_i \tag{6}$$

and that $(A + K_iC_i)\mathcal{V}_i \subset \mathcal{V}_i$. The latter, in turn, implies that there is a unique matrix A_i which satisfies

$$(A + K_iC_i)V_i = V_iA_i \tag{7}$$

where V_i is a basis matrix[2] for $\mathcal{V}_i$.

Next we show what defining the K_i in this way accomplishes. Note that the subspace $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_m$ is $\bar{A}$ - invariant because $(A + K_iC_i)\mathcal{V}_i \subset \mathcal{V}_i$, $i \in \mathbf{m}$. Next, let $Q =$ block diag $\{Q_1, Q_2, \dots, Q_m\}$ and $V =$ block diag $\{V_1, V_2, \dots, V_m\}$ where V_i is a matrix whose columns form an orthonormal basis for $\mathcal{V}_i$. Then Q is a full rank matrix whose kernel is $\mathcal{V}$ and V is a basis matrix for $\mathcal{V}$ whose columns form an orthonormal set. It follows that $P = VV'$, that $QP = 0$, and that

$$Q\bar{A} = \bar{A}_VQ \tag{8}$$

$$\bar{A}V = V\tilde{A} \tag{9}$$

where $\tilde{A} =$ block diag $\{A_1, A_2, \dots, A_m\}$ and

$$\bar{A}_V = \text{block diag } \{\bar{A}_1 + \bar{K}_1\bar{C}_1, \dots, \bar{A}_m + \bar{K}_m\bar{C}_m\}. \tag{10}$$

Let V^{-1} be any left inverse for V and let Q^{-1} be that right inverse for Q for which $V^{-1}Q^{-1} = 0$. Then

$$\bar{A} - gP(I_{mn} - \bar{S}(t)) = T\begin{bmatrix} \bar{A}_V & 0 \\ \hat{A}_V(t) & A_V(t) \end{bmatrix}T^{-1} \tag{11}$$

where $\hat{A}_V(t) = V^{-1}(\bar{A} - g(I_{mn} - \bar{S}(t)))Q^{-1}$ and $A_V(t) = \tilde{A} - gV'(I_{mn} - \bar{S}(t))V$. Here $T = \begin{bmatrix} Q^{-1} & V \end{bmatrix}$. It is easy to check that $T^{-1} = \begin{bmatrix} Q' & V \end{bmatrix}'$.

According to (11), the spectrum of $\bar{A} - gP(I_{mn} - \bar{S})$ is equivalent to the union of the spectrum of $\bar{A}_V$ and A_V. ■

Recall that the $\bar{K}_i$ have been already been chosen so that each matrix exponential $\exp\{(\bar{A}_i + \bar{K}_i\bar{C}_i)t\}$ converges to zero as fast as $\exp(-\lambda t)$ does. Because of this and the fact that $\hat{A}_V(t)$ is a bounded matrix, to ensure that for each fixed τ, the state transition matrix $\Phi(t, \tau)$ converges to zero as fast as $\exp(-\lambda t)$ does, it is enough to choose g so that the state transition matrix of $A_V(t)$ converges to zero at least as fast as $\exp(-\lambda t)$ does. The requisite condition on g is provided below for three different neighbor graph connectivity assumptions.

2.2. Constant neighbor graph

This subsection focuses on the case when the neighbor graph $\mathbb{N}(t)$ is a constant graph $\mathbb{N}$. The following result can be obtained.

Theorem 1. *For any given positive number λ, if the neighbor graph $\mathbb{N}$ is fixed and strongly connected, and the system defined by (1) and (2) is jointly observable, there are matrices K_i, $i \in \mathbf{m}$ such that for g sufficiently large, each estimation error $x_i(t) - x(t)$ of the distributed estimator defined by (3), converges to zero as $t \to \infty$ as fast as $\exp(-\lambda t)$ converges to zero.*

The proof of the theorem involves making use of the following result, with all proofs being contained in Appendix A. In particular, the proof of Proposition 2 makes use of the properties of strong connectivity of the neighbor graph and joint observability of the system.

Proposition 2. *$-V'(I_{mn} - \bar{S})V$ is a continuous-time stability matrix.*

[1] The *graph* of an $n \times n$ matrix M is that directed graph on n vertices possessing a directed arc from vertex i to vertex j if $m_{ij} \neq 0$ {p. 357, Horn and Johnson (1985)}.

[2] For simplicity, we assume that the columns of V_i constitute an orthonormal basis for $\mathcal{V}_i$ in which case $P_i = V_iV_i'$.

2.3. Switching neighbor graph

In the sequel the problem is studied under the assumption that $\mathbb{N}(t)$ changes according to a switching signal with a fixed dwell time or a variable dwell time with fixed average. To characterize the assumed time dependence of $\mathbb{N}(t)$, let $\mathcal{G} = \{\mathbb{G}_1, \mathbb{G}_2, \ldots, \mathbb{G}_{n_g}\}$ denote the set of all directed, strongly connected graphs on m vertices which have self-arcs at all vertices; here n_g is the number of graphs in $\mathcal{G}$. In some situations, the switching signals always have consecutive discontinuities separated by a value which is no less than a fixed positive real number τ_D. It is called a *dwell time* (Morse, 1993). In certain situations, the switching signals may occasionally have consecutive discontinuities separated by less than τ_D, but for which the average interval between consecutive discontinuities is no less than τ_D. This leads to the concept of an average dwell-time. With τ_D and δ fixed define $\mathcal{S}_{\text{avg}}$ for the set of all piecewise-constant switching signals $\sigma : [0, \infty) \to \{1, 2, \ldots, |\mathcal{G}|\}$ satisfying $\delta_\sigma(t_0, t) \leq \delta_0 + \frac{t - t_0}{\tau_D}$. Here $\delta_\sigma(t_0, t)$ denotes the number of discontinuities of σ in the open interval (t_0, t). The constant τ_D is called the average dwell-time and δ_0 the chatter bound (Hespahna & Morse, 1999). By the set of all time-varying neighbor graphs with average dwell-time τ_D is meant the set $\{\mathbb{G}_\sigma : \sigma \in \mathcal{S}_{\text{avg}}\}$. Note that switching according to a dwell time is a special case of switching according to an average dwell time. In the following, it is assumed that $\mathbb{N} \in \{\mathbb{G}_\sigma : \sigma \in \mathcal{S}_{\text{avg}}\}$.

The problem to which this subsection is addressed is this. For fixed averaged dwell-time τ_D and the chatter bound δ_0, devise a procedure for crafting m local estimators, one for each agent, so that for each neighbor graph $\mathbb{N} \in \{\mathbb{G}_\sigma : \sigma \in \mathcal{S}_{\text{avg}}\}$, all m state estimation errors converge to zero exponentially fast at a prescribed rate.

The estimator to be considered is still the same as the estimator described in (3), with the exception that g is chosen differently. The following result can be derived.

Theorem 2. *For any fixed positive numbers τ_D and λ, there exists a positive number g^* with the following property. For any value of $g \geq g^*$, any neighbor graph $\mathbb{N} \in \{\mathbb{G}_\sigma : \sigma \in \mathcal{S}_{\text{avg}}\}$, if the system defined by (1) and (2) is jointly observable, there are matrices K_i, $i \in \mathbf{m}$ such that, each state estimation error $e_i = x_i - x$, $i \in \mathbf{m}$ of the distributed estimator defined by (3) converges to zero as $t \to \infty$ as fast as $\exp(-\lambda t)$ does.*

The proof of Theorem 2 depends on the following lemma.

Lemma 1. *Let $M_1, M_2, \ldots, M_{|\mathcal{G}|}$ be a set of $n \times n$ exponentially stable real matrices associated with a set $\mathcal{G} = \{\mathbb{G}_1, \mathbb{G}_2, \ldots, \mathbb{G}_{|\mathcal{G}|}\}$ of directed strongly connected graphs with self-arcs at all vertices. Let σ denote the switching signal with average dwell time τ_D governing the selection of a graph from $\mathcal{G}$. Then for any $n \times n$ real matrix N and positive number λ there is a positive number g^*, depending on τ_D for which, for each $\sigma \in \mathcal{S}_{avg}$ and $g \geq g^*$, all solutions to*

$$\dot{x} = (N + gM_\sigma)x \tag{12}$$

converge to zero as fast as $\exp(-\lambda t)$ does.

The proofs of Lemma 1 and Theorem 2 can be found in Appendix A.

In the sequel, the problem is studied for a certain type of switching neighbor graphs. It turns out that if the stochastic matrices of undirected neighbor graphs are chosen to be doubly stochastic, there exist estimators which can deal with arbitrary switching signals, and the notion of dwell times ceases to be relevant. The estimator to be considered is again the same as (3), with the exception that g is chosen differently.

Theorem 3. *For any fixed positive number λ, there exists a positive number g^* with the following property. For any value of $g \geq g^*$, any time-varying neighbor graph $\mathbb{N}(t)$, if the system defined by (1) and (2) is jointly observable, the neighbor graph $\mathbb{N}(t)$ is undirected and connected, and the stochastic matrix $S(t)$ of graph $\mathbb{N}(t)$ is doubly stochastic, there are matrices K_i, $i \in \mathbf{m}$ such that each state estimation error $x_i(t) - x(t)$ of the distributed observer defined by (3), converges to zero as $t \to \infty$ as fast as $\exp(-\lambda t)$ converges to zero.*

The proof of Theorem 3 can be found in Appendix A.

Given that $\tilde{A}$ has eigenvalues which are a subset of those of A, it is seen that the effect of large g is to force the instantaneous value of the eigenvalues of $A_V(t)$ well to the left of those of $\tilde{A}$, and indeed the same for the Lyapunov exponent. This is a spectral separation idea — consensus dynamics within the estimator are faster than those of the original system.

2.4. Distributed estimator with adaptive gains

Obviously it may be disadvantageous to share a gain across the whole network, and here we aim to design a simple adaptive distributed estimator to get gains for each estimators in a distributed way. The estimator for each agent i still has the form (3) while each agent i has its own gain g_i which is obtained by

$$\dot{g}_i = \left| V_i' \sum_{j \in \mathcal{N}_i} \frac{1}{m_i}(x_j - x_i) \right|_2^2, \quad i \in \mathbf{m} \tag{13}$$

where $|\cdot|_2$ denotes the two norm of a matrix and $g_i(0)$ is nonnegative but otherwise arbitrary. Key questions arising are whether the g_i are bounded, and whether the matrices K_i can be chosen in the same way as previously. We have in fact with K_i chosen as in the proof of Proposition 1:

Theorem 4. *For any neighbor graph $\mathbb{N} \in \{\mathbb{G}_\sigma : \sigma \in \mathcal{S}_{avg}\}$, if the system defined by (1) and (2) is jointly observable, and the gain is defined by (13), there are matrices K_i, $i \in \mathbf{m}$ such that all the g_i are bounded, and each state estimation error $e_i = x_i - x$ of the distributed estimator defined by (3) asymptotically converges to zero as $t \to \infty$.*

The proof of Theorem 4 can be found in Appendix A.

As with any adaptive control algorithm, we must recognize that there are fundamental challenges that can arise in practice and have the potential to undermine the approach (Anderson & Gevers, 1998): these include the need to work with models of plants that may be very accurate but are virtually never exact; the inability to know, given an unknown plant, whether a desired control objective is practical or impractical, and the possibility of transient instability, or extremely large signals occurring before convergence. Thus, for this paper, our preference is to stick with a given g instead of using an adaptive algorithm.

2.5. Resilience

The concept of a passively resilient algorithm is proposed in Wang et al. (2022). By a passively resilient algorithm for a distributed process is meant an algorithm which, by exploiting built-in network and data redundancies, is able to continue to function correctly in the face of abrupt changes in the number of vertices and arcs in the inter-agent communication graph upon which the algorithm depends. All the proposed continuous-time distributed estimators, except for the adaptive one, are inherently resilient to these abrupt changes provided the network is redundantly strongly connected and redundantly jointly observable, with a careful gain picking before the algorithm starts. Details can be found in Section 5 of Wang et al. (2022). The same resilience property is also possessed by the discrete-time distributed estimators to be developed in the next section.

3. Discrete-time distributed estimator

In this section, a discrete time version of the distributed estimator problem is studied. The estimator which solves this problem in discrete time is described. Of central concern is the achieving of the split spectrum property, which underpins a stability guarantee for the estimator. High gain is not the answer: in continuous time, this produced a split spectrum (with some very fast modes). The discrete-time analogy rests on having part of the spectrum in the estimator achieved by gains K_i as in continuous-time and partly determined by the dynamics creating consensus among the components of local estimated states corresponding to unobserved components of the underlying system state, with the associated time scale significantly faster than the single time scale associated with the underlying system dynamics. The second time scale is made possible through the introduction of a faster sampling rate.

We are interested in the same time-varying network as used for the continuous-time linear system and which is characterized by the neighbor graph $\mathbb{N}(t)$. Each agent i can sense a discrete-time signal $y_i(\tau) \in \mathbb{R}^{s_i}$ at *event times* τT, $\tau = 0, 1, 2, \ldots$ where T is a positive constant; for $i \in \mathbf{m}$ and $\tau = 0, 1, 2, \ldots$

$$y_i(\tau) = C_i x(\tau), \qquad x(\tau+1) = Ax(\tau) \tag{14}$$

and $x \in \mathbb{R}^n$. We assume throughout that $\mathbb{N}(t)$ is strongly connected, and the system defined by (14) is *jointly observable*.

Each agent i is to estimate x using a dynamical system whose output $x_i(\tau) \in \mathbb{R}^n$ is to be an asymptotically correct estimate of $x(\tau)$ in the sense that the estimation error $x_i(\tau) - x(\tau)$ converges to zero as $\tau \to \infty$ as fast as λ^τ does, where λ is an arbitrarily chosen but fixed positive number[3] less than 1. To accomplish this it is assumed that the information agent i can receive from neighbor j at event time τT is $x_j(\tau)$. It is further assumed that agent i can also receive certain additional information from its neighbors at a finite number of times between each successive pair of event times; what this information is will be specified below.

3.1. The estimator

In this section it will be assumed that each agent's neighbors do not change between event times. In other words, for $i \in \mathbf{m}$,

$$\mathcal{N}_i(t) = \mathcal{N}_i(\tau T), \qquad t \in [\tau T, (\tau+1)T), \qquad \tau = 0, 1, 2, \ldots$$

With this assumption, the estimator to be considered consists of m private estimators, one for each agent. The estimator for agent i is of the form

$$x_i(\tau+1) = (A + K_i C_i)\bar{x}_i(\tau) - K_i y_i(\tau) \tag{15}$$

where $\bar{x}_i(\tau)$ is an "averaged state" computed recursively over q steps during the real time interval $[\tau T,\ (\tau+1)T)$ using the update equations

$$z_i(0, \tau) = x_i(\tau) \tag{16}$$

$$z_i(k, \tau) = (I - P_i) z_i(k-1, \tau) + \frac{1}{m_i(\tau)} P_i \sum_{j \in \mathcal{N}_i(\tau T)} z_j(k-1, \tau), \quad k \in \mathbf{q} \tag{17}$$

$$\bar{x}_i(t) = z_i(q, \tau) \tag{18}$$

Here $m_i(\tau)$ is the number of labels in $\mathcal{N}_i(\tau T)$, q is a suitably defined positive integer, further detail being given below, $\mathbf{q} = \{1, 2, \ldots, q\}$, and P_i is the orthogonal projection on the unobservable space of (C_i, A). Each matrix K_i is defined as explained in the next paragraph. Meanwhile, we note that the estimators incorporate two time scales. An agent's local estimator obtains data from its neighbors at a rate q times that at which it obtains measurement data from the underlying system.

As described in Section 2.1, for fixed $i \in \mathbf{m}$, write Q_i for any full rank matrix whose kernel is the unobservable space of (C_i, A), and let $\bar{C}_i$ and $\bar{A}_i$ be the unique solutions to $\bar{C}_i Q_i = C_i$ and $Q_i A = \bar{A}_i Q_i$ respectively. Let λ be a positive value bounded by 1. Then the matrix pair $(\bar{C}_i, \bar{A}_i)$ is observable. Thus by using a standard spectrum assignment algorithm, a matrix $\bar{K}_i$ can be chosen to ensure that the convergence of $(\bar{A}_i + \bar{K}_i \bar{C}_i)^\tau$ to zero as $\tau \to \infty$ is as fast as the convergence to zero of λ^τ. Having chosen such $\bar{K}_i$, K_i is then defined to be $K_i = Q_i^{-1} \bar{K}_i$ where Q_i^{-1} is a right inverse for Q_i. To explain what needs to be considered in choosing q, which is a rough analog of the gain g of the continuous-time solution, it is necessary to describe the structure of the "error model" of the overall estimator. This will be done next.

3.1.1. The error dynamics

For $i \in \mathbf{m}$, write $e_i(\tau)$ for the *state estimation error* $e_i(\tau) = x_i(\tau) - x(\tau)$. In view of (15),

$$e_i(\tau+1) = (A + K_i C_i)\bar{e}_i(\tau)$$

where $\bar{e}_i(\tau) = \bar{x}_i(\tau) - x(\tau)$. Moreover if $\epsilon_i(k, \tau) \stackrel{\Delta}{=} z_i(k, \tau) - x(\tau)$, $k \in \{0, 1, \ldots, q\}$ then for $k \in \mathbf{q}$,

$$\epsilon_i(0, \tau) = e_i(\tau)$$

$$\epsilon_i(k, \tau) = (I - P_i)\epsilon_i(k-1, \tau) + \frac{1}{m_i(\tau)} P_i \sum_{j \in \mathcal{N}_i(\tau T)} \epsilon_j(k-1, \tau)$$

$$\bar{e}_i(\tau) = \epsilon_i(q, \tau)$$

because of (16)–(18). It is possible to combine these m subsystems into a single system. Paralleling Section 2.1.1 let $e = \text{col}\{e_1, e_2, \ldots, e_m\}$, define $\bar{A} = \text{block diag}\{A + K_1 C_1, A + K_2 C_2, \ldots, A + K_m C_m\}$, $P = \text{block diag}\{P_1, P_2, \ldots, P_m\}$ and write $S(\tau)$ for the stochastic matrix $S(\tau) = D^{-1}_{\mathbb{N}(\tau T)} \mathcal{A}'_{\mathbb{N}(\tau T)}$ where $\mathcal{A}_{\mathbb{N}(\tau T)}$ is the adjacency matrix of $\mathbb{N}(\tau T)$ and $D_{\mathbb{N}(\tau T)}$ is the diagonal matrix whose ith diagonal entry is the in-degree of $\mathbb{N}(\tau T)$'s ith vertex. Let $\bar{e}(\tau) = \text{col}\{\bar{e}_1(\tau), \bar{e}_2(\tau), \ldots, \bar{e}_m(\tau)\}$ and $\epsilon(k, \tau) = \text{col}\{\epsilon_1(k, \tau), \epsilon_2(k, \tau), \ldots, \epsilon_m(k, \tau)\}$. Then $e(\tau+1) = \bar{A}\bar{e}(\tau)$ and

$$\epsilon(0, \tau) = e(\tau)$$

$$\epsilon(k, \tau) = (I_{mn} - P(I_{mn} - \bar{S}(\tau)))\epsilon(k-1, \tau), \quad k \in \mathbf{q}$$

$$\bar{e}(\tau) = \epsilon(q, \tau)$$

where $\bar{S}(\tau) = S(\tau) \otimes I_n$. Clearly $\bar{e}(\tau) = (I_{mn} - P(I_{mn} - \bar{S}(\tau)))^q e(\tau)$, so

$$e(\tau+1) = \bar{A}(I_{mn} - P(I_{mn} - \bar{S}(\tau)))^q e(\tau) \tag{19}$$

Our immediate aim is now to explain why for q sufficiently large, the time-varying matrix $\bar{A}(I_{mn} - P(I_{mn} - \bar{S}(\tau)))^q$ appearing in (19) is a discrete-time stability matrix for which the product

$$\Phi(\tau) = \prod_{s=1}^{\tau} \bar{A}(I_{mn} - P(I_{mn} - \bar{S}(s)))^q \tag{20}$$

converges to zero as $\tau \to \infty$ as fast as λ^τ does.

To begin with, we explore the property of matrix $\bar{A}(I_{mn} - P(I_{mn} - \bar{S}(\tau)))^q$.

Proposition 3. *The spectrum of the error system matrix $\bar{A}(I_{mn} - P(I_{mn} - \bar{S}(\tau)))^q$ splits into two subsets. One subset contains the union of certain subsets of the eigenvalues associated with the ith local estimator, $i \in \mathbf{m}$, these being able to be arbitrarily positioned by choice of the K_i and independent of q. The second subset is independent of the choice of the K_i and depends, though not to the extent of being able to be arbitrarily positioned, on q.*

[3] For the type of observer to be developed, finite-time convergence is not possible.

Proof of Proposition 3. As described in Section 2.1, note that the subspace $\mathcal{V} = \mathcal{V}_1 \oplus \mathcal{V}_2 \oplus \cdots \oplus \mathcal{V}_m$ is $\bar{A}$ - invariant because $(A+K_iC_i)\mathcal{V}_i \subset \mathcal{V}_i,\ i \in \mathbf{m}$. Next, let $Q = \text{block diag}\ \{Q_1, Q_2, \ldots, Q_m\}$ and $V = \text{block diag}\ \{V_1, V_2, \ldots, V_m\}$ It follows that $P = VV'$, (8) and (9). Also as before, $(A + K_iC_i)V_i = V_iA_i$. Moreover

$$Q(I_{mn} - P(I_{mn} - \bar{S}(\tau)))^q = Q \tag{21}$$

$$(I_{mn} - P(I_{mn} - \bar{S}(\tau)))^qV = V(V'\bar{S}(\tau)V)^q \tag{22}$$

Note that (21) holds because $QP = 0$. To understand why (22) is true, note first that $(I_{mn} - P(I_{mn} - \bar{S}(\tau)))V = V(I_{\bar{n}} - V'(I_{mn} - \bar{S}(\tau))V)$ because $P = VV'$; here $\bar{n} = \dim(\mathcal{V})$. But $I_{\bar{n}} - V'(I_{mn} - \bar{S}(\tau))V = V'\bar{S}(\tau)V$ because $V'V = I_{\bar{n}}$. Thus (22) holds for $q = 1$; it follows by induction that (22) holds for any positive integer q.

Using (8), (9), (21), and (22), one obtains the equations

$$Q\bar{A}(I_{mn} - P(I_{mn} - \bar{S}(\tau)))^q = \bar{A}_VQ \tag{23}$$

$$\bar{A}(I_{mn} - P(I_{mn} - \bar{S}(\tau)))^qV = VA_V(\tau) \tag{24}$$

where

$$A_V(\tau) = \tilde{A}(V'\bar{S}(\tau)V)^q \tag{25}$$

with $\tilde{A} =$ block diag $\{A_1, A_2, \ldots, A_m\}$. These equations imply that[4]

$$\bar{A}(I_{mn} - P(I_{mn} - \bar{S}(\tau)))^q = T\begin{bmatrix} \bar{A}_V & 0 \\ \hat{A}_V(\tau) & A_V(\tau) \end{bmatrix}T^{-1} \tag{26}$$

with $T = \begin{bmatrix} Q^{-1} & V \end{bmatrix}$ and $\hat{A}_V(\tau) = V^{-1}\bar{A}(I_{mn} - P(I_{mn} - \bar{S}(\tau)))^qQ^{-1}$.

According to (26), the spectrum of $\bar{A}(I_{mn} - P(I_{mn} - \bar{S}(\tau)))^q$ is equivalent to the union of the spectrum of A_V and $A_V(\tau)$. ■

3.2. Time-varying neighbor graph

The following result can be obtained when the neighbor graph is time-varying but strongly connected.

Theorem 5. *For any given λ with $|\lambda| < 1$, if the neighbor graph $\mathbb{N}(\tau)$ is strongly connected, and the system defined by (14) is jointly observable, there are matrices $K_i,\ i \in \mathbf{m}$ such that for sufficiently large q, each estimation error $x_i(\tau) - x(\tau)$ of the distributed estimator defined by (15)–(18), converges to zero as $\tau \to \infty$ as fast as λ^τ converges to zero.*

The proof of the theorem involves studying the transition matrix and making use of the following results, with all proofs being contained in Appendix B.

Lemma 2. *Let M be an $m \times m$ row stochastic matrix which has a strongly connected graph. There exists a diagonal matrix Π_M whose diagonal entries are positive for which the matrix $L_M = \Pi_M - M'\Pi_MM$ is positive semi-definite; moreover $L_M\mathbf{1} = 0$ where $\mathbf{1}$ is the m-vector of 1s. If, in addition, the diagonal entries of M are all positive, then the kernel of L_M is one-dimensional.*

Proposition 4. *For each fixed value of τ,*

$$(V'\bar{S}(\tau)V)'R(\tau)(V'\bar{S}(\tau)V) - R(\tau) < 0 \tag{27}$$

where $R(\tau)$ is the positive definite matrix, $R(\tau) = V'(\Pi_{S(\tau)} \otimes I_n)V$.

Remark 1. It should be noticed that the computation of certain gains (g for the continuous-time case, and q for the discrete-time case) requires a centralized design. Besides this, all other design steps are distributed. Even though the computation of certain gains requires going over all possible strongly connected directed graphs on m vertices, which is a computationally intensive step, a clear distinction needs to be drawn between the computations required for designing the algorithm, and those required to run it. In design, we can afford to do more computations.

[4] The notation $\hat{A}_V$ and A_V are different from the two defined in Section 2.

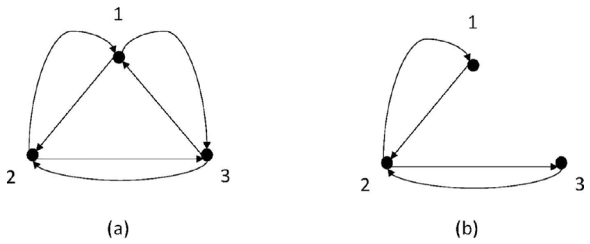


Fig. 1. The neighbor graph.

4. Simulations

This section provides simulations to illustrate the state estimation performance for both continuous and discrete time systems. The neighbor graph in some simulations will switch back and forth between Figs. 1(a) and 1(b). On occasion, it can serve to model a connection failure happening between agent 1 and agent 3 randomly.

4.1. Continuous dynamics

Consider the three channel, four-dimensional, continuous-time system described by the equations $\dot{x} = Ax,\ y_i = C_ix,\ i \in \{1,\ 2,\ 3\}$, where

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -2 & 0 \end{bmatrix}$$

and C_i is the ith unit row vector in $\mathbb{R}^{1\times4}$. Note that A is a matrix with eigenvalues at $\pm1j$, and $\pm1.4142j$. While the system is jointly observable, no single pair (C_i, A) is observable. The local observer convergence rate is designed to be at least with rate $\lambda = 1$. The first step is to design K_i as stated in Section 2.1. This is to control the spectrum of the matrix $\bar{A}_V$ as defined in (11).

For agent 1:

$$A_1 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},\quad Q_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix},\quad V_1 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}',$$

$$K_1 = \begin{bmatrix} -5 & -5 & 0 & 0 \end{bmatrix}'$$

For agent 2:

$$A_2 = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix},\quad Q_2 = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},\quad V_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}',$$

$$K_2 = \begin{bmatrix} 5 & -5 & 0 & 0 \end{bmatrix}'$$

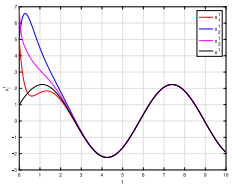
For agent 3:

$$A_3 = \begin{bmatrix} 0 & -2 \\ 1 & 0 \end{bmatrix},\quad Q_3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix},\quad V_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}',$$

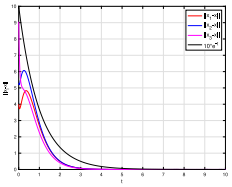
$$K_3 = \begin{bmatrix} 0 & 0 & -5 & -4 \end{bmatrix}'$$

Two cases are considered. First, suppose the neighbor graph $\mathbb{N}(t)$ is fixed as shown in Fig. 1(a). The system considered includes input white noise with zero mean, that is $\dot{x} = Ax + v$ where $E[v(t)] = 0, E[v(t)v'(s)] = 0.5^2\delta(t - s)$. With $g = 10$ obtained using (30), the real part of the right most eigenvalue of A_V is less than -1. With randomly chosen initial state values, traces of this simulation are shown in Fig. 2(a) where x_i^1 and x^1 denote the first components of x_i and x respectively. Moreover, the norm of the estimation error is plotted in Fig. 2(b) from which we can see that it is exponentially convergent with the approximate rate $\lambda = 1$.

Second, suppose the neighbor graph $\mathbb{N}(t)$ is time-varying and switching back and forth between Figs. 1(a) and 1(b) according



(a) Trajectory of the performance



(b) The trajectory of the norm of the estimation error

Fig. 2. The trajectory under time-varying neighbor graphs.

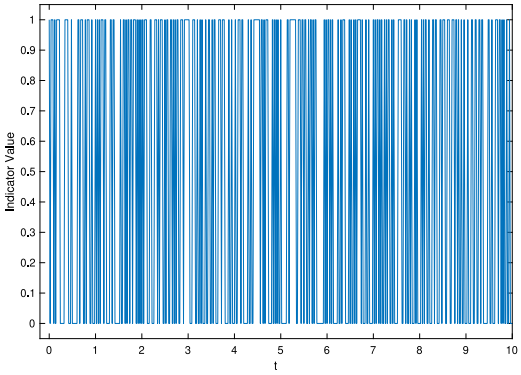
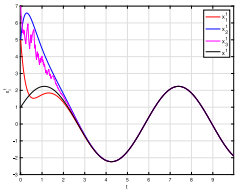
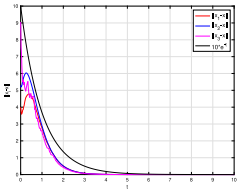


Fig. 3. The indicator function.



(a) The state value



(b) The estimation error

Fig. 4. The trajectory under time-varying neighbor graphs.

to the indicator function in Fig. 3. That is, when the function value is 1, the neighbor graph is Fig. 1(a), and when the function value is 0, the neighbor graph is Fig. 1(b). It is arranged that the average dwell time is $\tau_D = 0.0369$ for this simulation. With zero measurement noise the corresponding solution trajectories for x_i^3 and x^1 are shown in Fig. 4(a) and the norm of the estimation error is shown in Fig. 4(b).

4.2. Discrete dynamics

The following simulations are intended to illustrate how to pick the integer parameter q of the observer. Consider the three channel, four-dimensional, discrete-time system described by the equations $x(\tau+1) = Ax(\tau)$, $y_i = C_i x$, $i \in \{1, 2, 3\}$, where A and C_i are the same as described in Section 4.1. The observer convergence rate is designed to be $\lambda = 0.5$. The first step is also to design K_i as stated in Section 3.1. For each agent i, matrices A_i, Q_i, and V_i remain the same as stated in Section 4.1.

For agent 1: $K_1 = \begin{bmatrix}0.5 & 0.94 & 0 & 0\end{bmatrix}'$

For agent 2: $K_2 = \begin{bmatrix}-0.94 & 0.5 & 0 & 0\end{bmatrix}'$

For agent 3: $K_3 = \begin{bmatrix}0 & 0 & 0.5 & 1.94\end{bmatrix}'$

Consider the case when the neighbor graph $\mathbb{N}$ is constant as in Fig. 1(a). With randomly chosen initial state values and

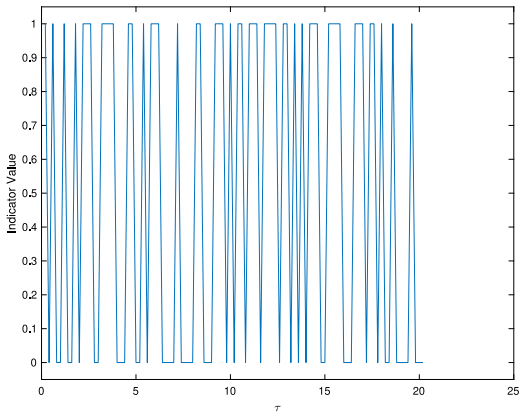
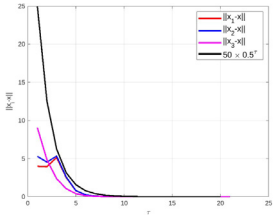
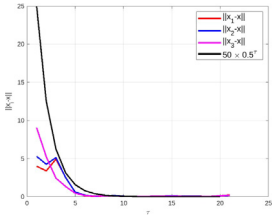


Fig. 5. The indicator function.



(a) Without noise



(b) With noise

Fig. 6. The trajectory of the norm of the estimation error for systems.

$q = 6$ obtained using (59), which leads to a spectral radius of $\tilde{A}(V'\bar{S}V)^6$ less than 0.5, the norm of the estimation error is plotted in Fig. 6(a) from which we can see that it is exponentially convergent with the rate $\lambda = 0.5$. The error traces are bounded by the curve $50 \times 0.5^{\tau}$.

Consider the case when the neighbor graph is switching between Fig. 1(a) and (b) according to Fig. 5. The system considered has input white noise v which satisfies $v \sim \mathcal{N}(0, 0.1^2)$, that is $x(\tau+1) = Ax(\tau) + v$. The norm of the estimation error is shown in Fig. 6(b) with the same value of q.

5. Concluding remarks

The distributed state estimation problem is studied when the neighbor graph is time-varying. It has been shown that, each agent can estimate the state exponentially fast with a pre-assigned convergence rate for both continuous-time and discrete-time systems.

The distributed state estimators developed in both Sections 2 and 3 rely on an especially useful observation about distributed estimator structure first noted in Kim et al. (2016) and subsequently exploited in Han et al. (2018) and Wang, Liu, and Morse (2019). That is to split the system spectrum into disjoint subsets corresponding to unobservable and observable subspaces. Just how much further this idea can be advanced remains to be seen. Generalization on the constraint of strong connectivity for the neighbor graph can be studied in future work. Certainly the synchronous switching upon which the local estimators in Section 3 depend can be relaxed by judicious application of the mixed matrix norm discussed here. This generalization will be addressed in future work.

Appendix A. Proofs for continuous-time distributed estimator

Proof of Proposition 2. Recall $\bar{S} = S \otimes I_n$. Since S is a stochastic matrix, S' must have a spectral radius of 1 and an eigenvalue at 1. Moreover, since $\mathbb{N}$ is the graph of S' and $\mathbb{N}$ is strongly connected by assumption, S' is irreducible {Theorem 6.2.24, Horn and Johnson (1985)}. Thus by the Perron–Frobenius Theorem there must be a positive vector π such that $S'\pi = \pi$. Without loss of generality, assume π is normalized so that the sum of its entries equals 1; i.e., π is a Perron vector (Horn & Johnson, 1985). Let Π be that diagonal matrix whose diagonal entries are the entries of π. Then $\Pi\mathbf{1} = \pi$ where $\mathbf{1}$ is the m-vector of all 1s. Let $L = 2\Pi - \Pi S - S'\Pi$. Clearly L is a symmetric matrix and $L\mathbf{1} = 0$.

We claim that the geometric multiplicity of L's eigenvalue at 0 is one. To establish this claim, note first that every nonzero entry of S' is a nonzero entry of $\Pi S + S'\Pi$ because S' is a nonnegative matrix and Π is a diagonal matrix whose diagonal entries are all positive. Moreover since $2(I_m - \Pi)$ is a nonnegative matrix, every nonzero entry of S' is also a nonzero entry of $2I_m - L = 2(I_m - \Pi) + \Pi S + S'\Pi$. Thus the graph of S' must be a spanning subgraph of the graph of $2I_m - L$ so the graph of $2I_m - L$ must be strongly connected. Therefore $2I_m - L$ must be irreducible. Note that the row sums of $2I_m - L$ all equal 2. Since $2I_m - L$ is nonnegative, its infinity norm is 2 so its spectral radius is no greater than 2 {Theorem 5.6.9, Horn and Johnson (1985)}. Moreover 2 is an eigenvalue of $2I_m - L$. Thus by the Perron–Frobenius Theorem, the geometric multiplicity of this eigenvalue is one. It follows that the geometric multiplicity of the eigenvalue of L at 0 is also one.

We claim that L is positive semi-definite. To establish this claim, note that L can also be written as $L = D - \widehat{A}$ where D is a diagonal matrix whose entries are the diagonal entries of L and $\widehat{A}$ is the nonnegative matrix $\widehat{A} = D - L$. As such, L is the generalized Laplacian (Godsil & Royle, 2001) of that simple undirected graph $\mathbb{G}$ whose adjacency matrix is the matrix that results when the nonzero entries a_{ij} in $\widehat{A}$ are replaced by ones. Since L can also be written as

$$L = \sum_{(i,j)\in\mathcal{E}} a_{ij}(u_i - u_j)(u_i - u_j)'$$

where u_i is the ith unit vector and $\mathcal{E}$ is the edge set of $\mathbb{G}$, L is positive semi-definite as claimed.

To proceed, set

$$H = \text{ block diag } \{\pi_1 I_{n_1}, \pi_2 I_{n_2}, \ldots, \pi_m I_{n_m}\} \tag{28}$$

where $n_i = \dim \mathcal{V}_i$ and note that $VH = (\Pi \otimes I_n)V$. Since $((S - I_m)' \otimes I_n)(\Pi \otimes I_n) = ((S - I_m)'\Pi) \otimes I_n$ it must be true that $(V'((S - I_m) \otimes I_n)V)'H = V'(((S - I_m)'\Pi) \otimes I_n)V$ and thus that

$$\begin{aligned} &H(V'(I_{mn} - \bar{S})V) + (V'(I_{mn} - \bar{S})V)'H \\ &= V'(L \otimes I_n)V \end{aligned} \tag{29}$$

Observe that this is a Lyapunov equation for the positive definite function $z'Hz$. Therefore to show that $-V'(I_{mn} - \bar{S})V$ is a stability matrix, it is enough to show that $V'(L \otimes I_n)V$ is positive definite.

Since L is positive semi-definite, so must be $L \otimes I_n$. Therefore $V'(L \otimes I_n)V$ is at least positive semi-definite. Suppose $z'V'(L \otimes I_n)Vz = 0$ where $z = \text{col}\{z_1, z_2, \ldots, z_m\}$ and $z_i \in \mathbb{R}^{\dim(\mathcal{V}_i)}$. To show that $V'(L \otimes I_n)V$ is positive definite, it is enough to show that $z = 0$. Since L's eigenvalue at 0 has multiplicity one, $\ker L = \text{span}\{\mathbf{1}\}$; therefore $\ker(L \otimes I_n) = \text{column span } \mathbf{1} \otimes I_n$. The hypothesis $z'V'(L \otimes I_n)Vz = 0$ implies that $(L \otimes I_n)Vz = 0$ so $Vz \in \ker(L \otimes I_n)$. Therefore $V_i z_i = V_j z_j,\ i, j \in \mathbf{m}$. But because of joint observability, $\bigcap_{i\in\mathbf{m}} \mathcal{V}_i = 0$ so $V_i z_i = 0,\ i \in \mathbf{m}$. Thus $z_i = 0,\ i \in \mathbf{m}$ so $z = 0$ implying that $V'(L \otimes I_n)V$ is positive definite. Therefore $-V'(I_{mn} - \bar{S})V$ is a continuous-time stability matrix as claimed. ■

Proof of Theorem 1. Recall that the state estimation error satisfied (4). The overall error dynamic is defined as (5). According to (11), the spectrum of $\bar{A} - gP(I_{mn} - \bar{S})$ is equivalent to the union of the spectrum of $\bar{A}_V$ and A_V. Since the spectrum of $\bar{A}_i + \bar{K}_i\bar{C}_i,\ i \in \mathbf{m}$, is assignable with $\bar{K}_i$, to show for g sufficiently large that $\bar{A} - gP(I_{mn} - \bar{S})$ is a continuous-time stability matrix with a prescribed convergence rate as large as λ, it is enough to show that for g sufficiently large, the matrix $A_V = \tilde{A} - gV'(I_{mn} - \bar{S})V$ is a continuous-time stability matrix with a prescribed convergence rate as large as λ.

To show that $\exp\{(\tilde{A} - gV'((I_m - S) \otimes I_n)V)t\}$ can be made to converge to zero as fast as $\exp(-\lambda t)$ does by choosing g sufficiently large, we exploit (29). Note in particular that

$$\begin{aligned} &H(\lambda I + A_V) + (\lambda I + A_V)'H \\ &= H(\lambda I + \tilde{A}) + (\lambda I + \tilde{A})'H - gV'(L \otimes I_n)V \end{aligned}$$

Since $V'(L \otimes I_n)V$ is positive definite, by picking g sufficiently large, $H(\lambda I + \tilde{A}) + (\lambda I + \tilde{A})'H - gV'(L \otimes I_n)V$ will be negative definite implying that $\lambda I + A_V$ is a stability matrix and thus that $\tilde{A} - gV'(I_{mn} - \bar{S})V$ is a stability matrix for which $\exp\{(\tilde{A} - gV'(I_{mn} - \bar{S})V)t\}$ converges to zero as fast as $\exp(-\lambda t)$ does. In other words, any value of g will have the desired property provided

$$g \geq \frac{\lambda_{\max}\left(H(\lambda I + \tilde{A}) + (\lambda I + \tilde{A})'H\right)}{\lambda_{\min}\left(V'(L \otimes I_n)V\right)} \tag{30}$$

where $\lambda_{\max}(.)$ and $\lambda_{\min}(.)$ are the largest eigenvalue and the smallest eigenvalue of a symmetric matrix respectively. ■

Proof of Lemma 1. [5] By hypothesis, each M_i is exponentially stable. Thus there are positive constants $c_i > 1$ and λ_i such that

$$\|\exp(M_i t)\| \leq c_i \exp(-\lambda_i t) \tag{31}$$

for any $i \in \{1, 2, \ldots, |\mathcal{G}|\}$. Here $\|\cdot\|$ is any given submultiplicative norm on $\mathbb{R}^{n\times n}$. Let

$$c = \max_{i\in\{1,2,\ldots,|\mathcal{G}|\}} c_i, \text{ and } \lambda^* = \min_{i\in\{1,2,\ldots,|\mathcal{G}|\}} \lambda_i.$$

Fix $\lambda > 0$ and let g be any gain satisfying

$$g \geq \frac{\tau_D(\lambda + \|N\|c) + \ln c}{\tau_D \lambda^*} \tag{32}$$

We claim that for any number τ, and any switching signal $\sigma \in \mathcal{S}_{\text{avg}}$, the transition matrix of gM_σ, namely $\Phi_\sigma(t, \tau)$, converges to zero as fast as $\exp(-\alpha t)$ does where

$$\alpha = \lambda + \|N\|c \tag{33}$$

To understand why this is so, by (31),

$$\|\Phi_\sigma(t, \tau)\| \leq c^{\delta_\sigma(\tau,t)} \exp(-g\lambda^*(t - \tau)) \tag{34}$$

where $\delta_\sigma(\tau, t)$ is the number of switching between (τ, t). By (32), $\exp(g\lambda^*) \geq c^{\frac{1}{\tau_D}} \exp(\alpha)$. From this and the fact that $\delta_\sigma(\tau, t) \leq \delta_0 + \frac{t-\tau}{\tau_D}$,

$$\begin{aligned} \|\Phi_\sigma(t, \tau)\| &\leq c^{\delta_0} c^{\frac{t-\tau}{\tau_D}} \exp(-g\lambda^*(t - \tau)) \\ &\leq c^{\delta_0 - \frac{\tau}{\tau_D}} \exp(-\alpha(t - \tau)). \end{aligned}$$

Thus, the claim is true.

In view of (12) and the variation of constants formula,

$$x(t) = \Phi_\sigma(t, 0)x(0) + \int_0^t \Phi_\sigma(t, \mu)Nx(\mu)d\mu \tag{35}$$

[5] The symbols used in this proof such as g, c and λ^* are generic and do not have the same meanings as the same symbols do when used elsewhere in this paper.

As $\|\Phi_\sigma(t,\tau)\| \le c\exp(-\alpha t)$ for all τ and $\sigma \in \mathcal{S}_{\text{avg}}$,

$$\|x(t)\| \le c\exp(-\alpha t)\|x(0)\| + \int_0^t c\exp(-\alpha(t-\mu))\|N\|\|x(\mu)\|d\mu$$

By multiplying by $\exp(\alpha t)$ on both sides, one obtains $\exp(\alpha t)\|x(t)\| \le c\|x(0)\| + \int_0^t \|N\|c\exp(\alpha\mu)\|x(\mu)\|d\mu$. From this and the Bellman–Gronwall Lemma there follows

$$\exp(\alpha t)\|x(t)\| \le c\|x(0)\|\exp(\int_0^t \|N\|cd\mu)$$

Since $\int_0^t \|N\|cd\mu = \|N\|ct$, it follows that $\exp(\alpha t)\|x(t)\| \le c\|x(0)\|\exp(\|N\|ct)$ and thus that

$$\|x(t)\| \le c\|x(0)\|\exp((\|N\|c-\alpha)t)$$

From this and (33) it follows that

$$\|x(t)\| \le c\|x(0)\|\exp(-\lambda t)$$

This completes the proof. ■

Proof of Theorem 2. Recall that $A_V(t) = \tilde{A} - gV'(I_{mn} - \bar{S}(t))V$. By Proposition 2, for any fixed time τ, $-V'(I_{mn} - \bar{S}(\tau))V$ is exponentially stable if the graph of $S(\tau)'$ is strongly connected. Note $\tilde{A}$ is fixed and bounded. According to Lemma 1, for each $\sigma \in \mathcal{S}_{\text{avg}}$ there is a positive number g, depending on τ_D so that the transition matrix of $A_V(t)$ converges to zero at least as fast as $\exp(-\lambda t)$ does. This is accomplished by choosing g sufficiently large. Based on the proof of Lemma 1, it is sufficient to pick g to satisfy

$$g \ge \frac{\ln c + (\lambda + \|\tilde{A}\|c)\tau_D}{\lambda^*\tau_D} \tag{36}$$

where c and λ^* are two positive numbers chosen so that for any fixed τ, $\|\exp\{-V'(I_{mn} - \bar{S}(\tau))Vt\}\| \le c\exp(-\lambda^* t)$, and $c > 1$. This completes the proof. ■

Proof of Theorem 3. Recall $\Phi_V(t,\tau)$ is the transition matrix of $A_V(t)$ for any $t \ge \tau \ge 0$. If we can show that there exist a constant c so that

$$\|\Phi_V(t,\tau)\| \le c\exp(-\lambda(t-\tau)), \quad \forall t \ge \tau \ge 0$$

the remaining proof is exactly the same as the proof of Theorem 1 which is omitted here.

It is left to show that $\|\Phi_V(t,\tau)\| \le c\exp(-\lambda(t-\tau))$ for all $t \ge \tau \ge 0$ by choosing g sufficiently large. We explore the matrix $A_V(t)$. Recall that $A_V(t) = \tilde{A} - gV'((I_m - S(t)) \otimes I_n)V$. In particular,

$$\begin{aligned}&(\lambda I + A_V(t)) + (\lambda I + A_V(t))'\\ &= (\lambda I + \tilde{A}) + (\lambda I + \tilde{A})'\\ &\quad -gV'((2I_m - S(t) - S'(t)) \otimes I_n)V\end{aligned}$$

Since each $S(t)$ is doubly stochastic, $2I_m - S(t) - S'(t)$ has row sum 0, all its off-diagonal entries are non-positive, and all its diagonal entries are positive. That is, this matrix can be seen as a generalized Laplacian matrix of a connected graph. By Proposition 2, for any t, $-V'((2I_m - S(t) - S'(t)) \otimes I_n)V = -V'(I_{mn} - \bar{S}(t))V - V'(I_{mn} - \bar{S}'(t))V$ is negative definite. Thus by picking g sufficiently large, $(\lambda I + A_V(t)) + (\lambda I + A_V(t))'$ will be negative definite for any time t.

Consider system

$$\dot{\bar{z}} = A_V(t)\bar{z}$$

Let $V = \bar{z}'\bar{z}$. Then

$$\dot{V} = \bar{z}'(A_V(t)' + A_V(t))\bar{z} \le -2\lambda\bar{z}'\bar{z}$$

Therefore, $\Phi_V(t,\tau)$ converges to zero as fast as $\exp(-\lambda(t-\tau))$ does, i.e.,

$$\|\Phi_V(t,\tau)\| \le c\exp(-\lambda(t-\tau)), \quad \forall t \ge \tau \ge 0$$

This completes the proof. ■

Proof of Theorem 4. Eq. (13) can be rewritten as

$$\dot{g}_i = |V_i'W_ie|_2^2, \quad i \in \mathbf{m} \tag{37}$$

where $W_i = \begin{bmatrix} W_{i1} & \dots & W_{im}\end{bmatrix} \in \mathbb{R}^{n\times nm}$. Here $W_{ij} \in \mathbb{R}^{n\times n}$ is $\frac{1}{m_i}I_n$ if $j \ne i$ and $j \in \mathcal{N}_i$, and W_{ij} is $-I_n$ if $j = i$. Otherwise W_{ij} is a 0 matrix. Let column$\{W_1, W_2, \dots, W_m\} = W$.

Different from (5), the error model turns to

$$\dot{e} = \bar{A} - G(t)P(I_{mn} - \bar{S}(t))e \tag{38}$$

where $G(t) =$ block diag $\{g_1(t)I_n, \dots, g_m(t)I_n\}$. Let $\begin{bmatrix}(Qe)' & z'\end{bmatrix}' = T^{-1}e$ where $T = \begin{bmatrix}Q^{-1} & V\end{bmatrix}$ as defined earlier. Here $z = [z_1', \dots, z_m']'$ iwth $z_i = V_i'e_i$.

Based on (11) and (38), the dynamic of z_i can be written in the following form

$$\dot{z}_i = \tilde{A}_iz_i - g_i(t)V_i'M_ie + \hat{A}_iQ_ie_i, \quad i \in \mathbf{m} \tag{39}$$

where $\tilde{A}_i = V_i'(A + K_iC_i)V_i$, and $\hat{A}_i = V_i'(A + K_iC_i)Q_i'$.

First, we want to show that all $g_i(t)$ are bounded. We prove this by contradiction. Without generality, suppose that g_i for $i \in \mathcal{V}_u = \{1, 2, \dots, m_1\}$ are unbounded, and g_i for $i \in \mathcal{V}_b = \{m_1 + 1, m_1+2, \dots, m\}$ are bounded where $\mathcal{V}_u \cap \mathcal{V}_b = 0$ and $\mathcal{V}_u \cup \mathcal{V}_b = \mathbf{m}$.

Let $R = R_1 + R_2 + R_3 + R_4$ where the individual R_i involve new positive parameters p, α_0, $\alpha_{m_1+1}, \dots, \alpha_m$.

$$R_1 = \frac{1}{2}\sum_{i=1}^{m_1}\pi_i\frac{p}{g_i(t)}|z_i|_2^2, \quad R_2 = \frac{1}{2}\sum_{i=m_1+1}^{m}\pi_i\frac{g_i(0)}{g_i(t)}|z_i|_2^2$$

$$R_3 = -\sum_{i=m_1+1}^{m}\alpha_ig_i(t), \text{ and } R_4 = -\alpha_0\int_0^t |Qe|_2^2dt.$$

The way to pick positive parameters p, α_0, and α_i for $i \in \mathcal{V}_b$ is specified as follows.

Picking $p \ge 1$:

Let $\mathcal{W}_1$ be a positive matrix chosen such that

$$z'\mathcal{W}_1z = \sum_{i=1}^{m}\pi_i|\tilde{A}_i|_2|z_i|_2^2 \tag{40}$$

According to (29),

$$F = \frac{1}{2}(HV'(\bar{S} - I_{mn})V + V'(\bar{S} - I_{mn})'VH) > 0 \tag{41}$$

Pick p so that $\mathcal{W}_2 = pF - \mathcal{W}_1 > 0$

Picking α_0, and α_i, $i \in \mathcal{V}_b$:

Using the Cauchy–Schwarz inequality, the following three inequalities can be derived. For $\beta_1 > 0$, $\beta_2 > 0$ and $\lambda_i > 0$ for $i \in \mathcal{V}_b$, all for the moment otherwise arbitrary, write $\hat{A} =$ block diag $\{\hat{A}_1, \dots, \hat{A}_m\}$

$$z'H\hat{A}Qe \le \frac{\beta_1}{2}|\hat{A}'Hz|_2^2 + \frac{1}{2\beta_1}|Qe|_2^2, \tag{42}$$

$$\begin{aligned}&-pz'HV'(\bar{S} - I_{mn})Q'Qe \le\\ &\frac{\beta_2}{2}|Q(\bar{S} - I_{mn})'VHz|_2^2 + \frac{1}{2\beta_2}|Qe|_2^2\end{aligned} \tag{43}$$

$$\begin{aligned}&-\pi_i(g_i(0) - p)z_i'V_i'W_ie \le\\ &\frac{\lambda_i}{2}|\pi_i(g_i(0) - p)z_i|_2^2 + \frac{1}{2\lambda_i}|V_i'W_ie|_2^2.\end{aligned} \tag{44}$$

Let $\mathcal{W}_3$ be a symmetric matrix chosen such that

$$z'\mathcal{W}_3 z = z'\mathcal{W}_2 z - \frac{\beta_1}{2}|\hat{A}'Hz|_2^2 - \frac{\beta_2}{2}|Q(\bar{S}-I_{mn})'VHz|_2^2 - \sum_{i=m_1+1}^{m}\frac{\lambda_i}{2}|\pi_i(g_i(0)-p)z_i|_2^2 \tag{45}$$

Pick positive β_1, β_2, and λ_i for $i \in \mathcal{V}_b$ to be small enough so that W_3 is positive definite. Pick α_0 and α_i, $i \in \mathcal{V}_b$ according to the following two equations:

$$\alpha_0 = \frac{1}{2\beta_1} + \frac{1}{2\beta_2} \quad \text{and} \quad \alpha_i = \frac{1}{2\lambda_i} \quad i \in \mathcal{V}_b \tag{46}$$

Now we consider the derivative of R, i.e., $\dot{R} = \dot{R}_1 + \dot{R}_2 + \dot{R}_3 + \dot{R}_4$.

$$\dot{R}_1 = -\frac{1}{2}\sum_{i=1}^{m_1}\pi_i\frac{p}{g_i^2}|z_i|_2^2|V_i'W_ie|_2^2 + \sum_{i=1}^{m_1}\pi_i\frac{p}{g_i}z_i'\tilde{A}_iz_i + \sum_{i=1}^{m_1}\pi_i\frac{p}{g_i}z_i'\hat{A}_iQ_ie_i - \sum_{i=1}^{m_1}\pi_ipz_i'V_i'W_ie$$

Since for $i \in \mathcal{V}_u$ each $g_i(t)$ is unbounded, there is a time T for $t \geq T$, $p \leq g_i(t)$ for all $i \in \mathcal{V}_u$. That is $\frac{p}{g_i(t)} \leq 1$. Hence

$$\dot{R}_1 \leq \sum_{i=1}^{m_1}\pi_iz_i'\tilde{A}_iz_i + \sum_{i=1}^{m_1}\pi_iz_i'\hat{A}_iQ_ie_i - \sum_{i=1}^{m_1}\pi_ipz_i'V_i'W_ie \tag{47}$$

According to (37), $g_i(t)$ is a non-decreasing function. As a result, $\frac{g_i(0)}{g_i(t)} \leq 1$. Similar to the derivation of the inequality for $\dot{R}_1$, it can be shown that

$$\dot{R}_2 \leq \sum_{i=m_1+1}^{m}\pi_iz_i'\tilde{A}_iz_i + \sum_{i=m_1+1}^{m}\pi_iz_i'\hat{A}_iQ_ie_i - \sum_{i=m_1+1}^{m}\pi_ig_i(0)z_i'V_i'W_ie \tag{48}$$

By the submultiplicity of the matrix two norm, $z_i'\tilde{A}_iz_i' \leq |\tilde{A}_i|_2|z_i|_2^2$. From this, (47), (48), (40) and (28),

$$\begin{aligned}\dot{R}_1 + \dot{R}_2 &\leq z'\mathcal{W}_1z + \sum_{i=1}^{m}\pi_iz_i'\hat{A}_iQ_ie_i - \sum_{i=1}^{m}\pi_ipz_i'V_i'W_ie - \sum_{i=m_1+1}^{m}\pi_i(g_i(0)-p)z_i'V_i'W_ie\\ &= z'\mathcal{W}_1z + z'H\hat{A}Qe - pz'HV'(\bar{S}-I_{mn})e - \sum_{i=m_1+1}^{m}\pi_i(g_i(0)-p)z_i'V_i'W_ie\end{aligned}$$

It can be observed that $z'HV'(\bar{S}-I_{mn})e = z'HV'(\bar{S}-I_{mn})(VV'e + Q'Qe) = z'HV'(\bar{S}-I_{mn})Vz + z'HV'(\bar{S}-I_{mn})Q'Qe$. From this and (41), $z'HV'((S-I_m)\otimes I_n)e = z'Fz + z'HV'((S-I_m)\otimes I_n)Q'Qe$. Thus

$$\begin{aligned}\dot{R}_1 + \dot{R}_2 &\leq z'(\mathcal{W}_1 - pF)z + z'H\hat{A}Qe - pz'HV'(\bar{S}-I_{mn})Q'Qe - \sum_{i=m_1+1}^{m}\pi_i(g_i(0)-p)z_i'V_i'W_ie\\ &= -z'\mathcal{W}_2z + z'H\hat{A}Qe - pz'HV'(\bar{S}-I_{mn})Q'Qe - \sum_{i=m_1+1}^{m}\pi_i(g_i(0)-p)z_i'V_i'W_ie\end{aligned}$$

From this, (42), (43), and (44),

$$\dot{R}_1 + \dot{R}_2 \leq -z'\mathcal{W}_2z + \frac{\beta_1}{2}|\hat{A}'Hz|_2^2 + \frac{\beta_2}{2}|Q(\bar{S}-I_{mn})'VHz|_2^2 + \sum_{i=m_1+1}^{m}\frac{\lambda_i}{2}|\pi_i(g_i(0)-p)z_i|_2^2 + \frac{1}{2\beta_1}|Qe|_2^2 + \frac{1}{2\beta_2}|Qe|_2^2 + \sum_{i=m_1+1}^{m}\frac{1}{2\lambda_i}|V_i'W_ie|_2^2 \tag{49}$$

It is direct to get that

$$\dot{R}_3 + \dot{R}_4 = -\sum_{i=m_1+1}^{m}\alpha_i|V_i'W_ie|_2^2 - \alpha_0|Qe|_2^2 \tag{50}$$

From (49), (50), and (46),

$$\dot{R} \leq -z'\mathcal{W}_2z + \frac{\beta_1}{2}|\hat{A}'Hz|_2^2 + \frac{\beta_2}{2}|Q((S-I)\otimes I)'VHz|_2^2 + \sum_{i=m_1+1}^{m}\frac{\lambda_i}{2}|\pi_i(g_i(0)-p)z_i|_2^2$$

According to (45),

$$\dot{R} \leq -z'\mathcal{W}_3z < 0 \tag{51}$$

Since Qe is exponentially convergent, the limit of $R_4(t)$ as t goes to infinity exists. Due to the assumption that for $i \in \mathcal{V}_b$, the g_i are bounded. Thus R_3 is bounded. Therefore, R is lower bounded. From this and (51) it follows that $z \in \mathcal{L}^2$. From this and the fact that $Qe \in \mathcal{L}^2$, we conclude $e \in \mathcal{L}^2$. This with the definition of $\dot{g}_i$ imply that all g_i for $i \in \mathbf{m}$ are bounded. Thus by contradiction all g_i are bounded.

Next, we want to show that e converges to zero. According to Theorem 1, let $G_1 = g_1I$ be a matrix chosen so that $\bar{A} - G_1P(I_{mn} - \bar{S}(t))$ is a stable matrix. From (38), the dynamic of e can be rewritten as

$$\dot{e} = (\bar{A} - G_1P(I_{mn} - \bar{S}(t)))e + (G_1 - G)P(I_{mn} - \bar{S}(t))e \tag{52}$$

Since $(I_{mn} - \bar{S}(t))e \in \mathcal{L}^2$, $G_1 - G$ is bounded, and $\bar{A} - G_1P(I_{mn} - \bar{S}(t))$ is stable, thus $\dot{e} \in \mathcal{L}^2$. Thus (52) is input-to-state stable which implies that e must converge to zero asymptotically. ■

Appendix B. Proofs for discrete-time distributed estimator

Proof of Lemma 2. Since M is an $m \times m$ row stochastic matrix which ahs a strongly connected graph, M is irreducible {Theorem 6.2.24, Horn and Johnson (1985)}. Thus by the Perron–Frobenius Theorem there must be a positive vector π such that $M'\pi = \pi$. Without loss of generality, assume π is normalized so that the sum of its entries equals 1; i.e., π is a probability vector. Let Π_M be that diagonal matrix whose diagonal entries are the entries of π. Then $\Pi_M\mathbf{1} = \pi$. Since $M\mathbf{1} = \mathbf{1}$, $\Pi_M\mathbf{1} = \pi$, and $M'\pi = \pi$, it must be true that $M'\Pi_MM\mathbf{1} = \pi$ and thus that $L_M\mathbf{1} = 0$. Thus L_M can also be written as $L_M = D - \hat{A}$ where D is a diagonal matrix whose diagonal entries are the diagonal entries of L_M and $\hat{A}$ is the nonnegative matrix $\hat{A} = D - L_M$. Arguing as in the proof of Proposition 2, it can be shown that L_M is positive-semidefinite.

Now suppose that the diagonal entries of M are all positive. Then the diagonal entries of $M'\Pi_M$ must also all be positive. It follows that every arc in the graph of M' must be an arc in the graph of $M'\Pi_MM$ so the graph of $M'\Pi_MM$ must be strongly connected. Since $I - \Pi_M$ is a nonnegative matrix, the graph of $M'\Pi_MM$ must be a spanning subgraph of the graph of $I - \Pi_M + M'\Pi_MM$. Since $I - L_M = I - \Pi_M + M'\Pi_MM$ and the graph of $M'\Pi M$ is strongly connected, the graph of $I - L_M$ must be strongly connected as well. But $I - L_M$ is a nonnegative matrix so it must be irreducible. In addition, since $(I - L_M)\mathbf{1} = \mathbf{1}$, the row sums of $(I - L_M)$ all equal one. Therefore the infinity norm of $I - L_M$ is one so its spectral radius is no greater than 1. Moreover 1 is an eigenvalue of $I - L_M$. Thus by the Perron–Frobenius Theorem, the geometric multiplicity of this eigenvalue is one. It follows that the geometric multiplicity of the eigenvalue of L_M at 0 is also one; ie, the dimension of the kernel of L_M is one as claimed. ■

Proof of Proposition 4. Fix τ and write S for $S(\tau)$ and $\bar{S}$ for $\bar{S}(\tau)$. Note that the graph of S', namely $\mathbb{N}$, is strongly connected. In view of Lemma 2, the matrix $L = \Pi_S - S'\Pi_S S$ is positive semi-definite and $L\mathbf{1} = 0$. Moreover, since the diagonal entries of S and thus S' are all positive, the kernel of L is one-dimensional.

Write R for $R(\tau)$. To prove the proposition it is enough to show that the matrix

$$Q = R - (V'\bar{S}'V)R(V'\bar{S}V) \tag{53}$$

is positive definite.

To proceed, set $\bar{L} = L \otimes I_n$ in which case $\bar{L}$ is positive semi-definite because L is. Moreover, $\bar{L} = \bar{\Pi} - \bar{S}'\bar{\Pi}\bar{S}$ where $\bar{\Pi} = \Pi_S \otimes I_n$. Note that $VRV' = P\bar{\Pi}P$ where P is the orthogonal projection matrix $P = VV'$. Clearly $VRV' = P\bar{\Pi}^{\frac{1}{2}}\bar{\Pi}^{\frac{1}{2}}P$. Note that both P and $\bar{\Pi}^{\frac{1}{2}}$ are block diagonal matrices with corresponding diagonal blocks of the same size. Because of this and the fact that each diagonal block in $\bar{\Pi}^{\frac{1}{2}}$ is a scalar times the identity matrix, it must be true that P and $\bar{\Pi}^{\frac{1}{2}}$ commute; thus $P\bar{\Pi}^{\frac{1}{2}} = \bar{\Pi}^{\frac{1}{2}}P$. From this and the fact that P is idempotent, it follows that $VRV' = \bar{\Pi}^{\frac{1}{2}}P\bar{\Pi}^{\frac{1}{2}}$. Clearly $\bar{\Pi}^{\frac{1}{2}}P\bar{\Pi}^{\frac{1}{2}} \le \bar{\Pi}^{\frac{1}{2}}\bar{\Pi}^{\frac{1}{2}}$ so $VRV' \le \bar{\Pi}$. It follows using (53) that $Q \ge R - V'\bar{S}'\bar{\Pi}\bar{S}V = R + V'\bar{L}V - V'\bar{\Pi}V$. Therefore

$$Q \ge V'\bar{L}V \tag{54}$$

In view of this, to complete the proof it is enough to show that $V'\bar{L}V$ is positive definite. This can be shown by the same proof of Proposition 2.

Therefore Q is positive definite because of (54). From this and (53) it follows that (27) is true. ■

Proof of Theorem 5. First it will be assumed that each $\bar{K}_i$ has been selected so that the matrix $\bar{A}_V$ defined by (10), is such that $\bar{A}_V^\tau$ converges to zero as $\tau \to \infty$ as fast as λ^τ does. This can be done using standard spectrum assignment techniques to make the spectral radius of $\bar{A}_V$ at least as small as λ. In view of (26), it is clear that to assign the convergence rate of the state transition matrix of $\bar{A}(I_{mn} - P(I_{mn} - \bar{S}(\tau)))^q$ it is necessary and sufficient to control the convergence rate of the state transition matrix of $A_V(\tau)$. This, as we will now show, can be accomplished by choosing q sufficiently large. We will actually detail two different ways to do this, each utilizing a different matrix norm. Both approaches will be explained next using the abbreviated notation $B(\tau) = V'\bar{S}(\tau)V$; note that with this simplification, $A_V(\tau) = \tilde{A}B^q(\tau)$ because of (25).

Weighted Two-Norm: For each fixed τ and each appropriately-sized matrix M, write $\|M\|_{R(\tau)}$ for the matrix norm induced by the vector norm $\|x\|_{R(\tau)} \stackrel{\Delta}{=} \sqrt{x'R(\tau)x}$. Note that $\|M\|_{R(\tau)}$ is the largest singular value of $R^{\frac{1}{2}}(\tau)MR^{-\frac{1}{2}}(\tau)$. Note in addition that

$$(R^{\frac{1}{2}}(\tau)B(\tau)R^{-\frac{1}{2}}(\tau))'(R^{\frac{1}{2}}(\tau)B(\tau)R^{-\frac{1}{2}}(\tau)) < I$$

because of (27). This shows that the largest singular value of $R^{\frac{1}{2}}(\tau)B(\tau)R^{-\frac{1}{2}}(\tau)$ is less than one. Therefore

$$\|B(\tau)\|_{R(\tau)} < 1 \tag{55}$$

(a) $\mathbb{N}$ is constant

In this case both $B(\tau)$ and $R(\tau)$ are constant, so it is sufficient so choose q so that $\|\tilde{A}B^q(\tau)\|_{R(\tau)} \le \lambda$. Since $\|\cdot\|_{R(\tau)}$ is submultiplicative, this can be done by choosing q so that

$$\|B(\tau)\|_{R(\tau)}^q \le \frac{\lambda}{\|\tilde{A}\|_{R(\tau)}} \tag{56}$$

This can always be accomplished because of (55).

(b) $\mathbb{N}$ changes with time

In this case it is not possible to use the weighted two-norm $\|\cdot\|_{R(\tau)}$ because it is time-dependent. A simple fix, but perhaps not the most efficient one, would be to use the standard two-norm $|\cdot|_2$ instead since it does not depend on time. Using this approach, the first step would be to first choose, for each fixed τ, an integer $p_1(\tau)$ large enough so that $|B^{p_1(\tau)}(\tau)|_2 < 1$. Such values of $p_1(\tau)$ must exist because each $B(\tau)$ is a discrete-time stability matrix or equivalently, a matrix with a spectral radius less than 1. Computing such a value amounts to looking at the largest singular value of $B^{p_1(\tau)}(\tau)$ for successively largest values of $p_1(\tau)$ until that singular value is less than 1. Having accomplished this, a number p can easily be computed so that $|B^p(\tau)|_2 < 1\ \forall\tau$ since there are only a finite number of distinct strongly connected graphs on m vertices and consequently only a finite number of distinct matrices $B(\tau)$ in the set $\mathcal{B} = \{B(\tau) : \tau \ge 0\}$. Choosing p to be the maximum of the $p_1(\tau)$ with respect to τ is thus a finite computation. The next step would be to compute an integer $\bar{p}$ large enough so that each $|\tilde{A}(B^p(\tau))^{\bar{p}}|_2 \le \lambda$. A value of q with the required property would then be $q = p\bar{p}$.

Mixed Matrix Norm: There is a different way to choose q which does not make use of either Lemma 2 or Proposition 4. The approach exploits the "mixed matrix norm" introduced in Mou, Liu, and Morse (2015). To define this norm requires several steps. To begin, let $|\cdot|_\infty$ denote the standard induced infinity norm and write $\mathbb{R}^{mn\times mn}$ for the vector space of all $m \times m$ block matrices $M = [M_{ij}]$ whose ijth entry is a matrix $M_{ij} \in \mathbb{R}^{n\times n}$. With $n_i = \dim \mathcal{V}_i$, $i \in \mathbf{m}$, and $\bar{n} = n_1 + n_2 + \cdots n_m$, write $\mathbb{R}^{mn\times\bar{n}}$ for the vector space of all $m \times m$ block matrices $M = [M_{ij}]$ whose ijth entry is a matrix $M_{ij} \in \mathbb{R}^{n\times n_j}$. Similarly write $\mathbb{R}^{\bar{n}\times mn}$ for the vector space of all $m \times m$ block matrices $M = [M_{ij}]$ whose ijth entry is a matrix $M_{ij} \in \mathbb{R}^{n_i\times n}$. Finally write $\mathbb{R}^{\bar{n}\times\bar{n}}$ for the vector space of all $m \times m$ block matrices $M = [M_{ij}]$ whose ijth entry is a matrix $M_{ij} \in \mathbb{R}^{n_i\times n_j}$.

Note that $B \in \mathbb{R}^{mn\times mn}$, $\tilde{A} \in \mathbb{R}^{\bar{n}\times\bar{n}}$, $V \in \mathbb{R}^{mn\times\bar{n}}$, and $V' \in \mathbb{R}^{\bar{n}\times mn}$. For M in any one of these four spaces, the *mixed matrix norm* (Mou et al., 2015) of M, written $\|M\|$, is

$$\|M\| = \|\langle M\rangle\|_\infty \tag{57}$$

where $\langle M\rangle$ is the matrix in $\mathbb{R}^{m\times m}$ whose ijth entry is $\|M_{ij}\|_2$. It is very easy to verify that $\|\cdot\|$ is in fact a norm. It is even sub-multiplicative whenever matrix multiplication is defined. Note in addition that $\|V\| = 1$ and $\|V'\| = 1$ because the columns of each V_i form an orthonormal set.

Recall that $P = VV'$ is an orthogonal projection matrix. Using this, the definition of $B(\tau)$ and the fact that $PV = V$, it is easy to see that for any integer $p > 0$, $B^p(\tau) = V'(P\bar{S}(\tau)P)^pV$. Thus $\|B^p(\tau)\| \le \|(P\bar{S}(\tau)P)^p\|$. Using this and the fact that the graph of S' is strongly connected, one can conclude that for $p \ge (m-1)^2$, $\|(P\bar{S}(\tau)P)^p\| < 1$. This is a direct consequence of Proposition 2 of Mou et al. (2015). Thus

$$\|B^p(\tau)\| < 1, \quad p \ge (m-1)^2 \tag{58}$$

(a) $\mathbb{N}$ is constant

In this case $B(\tau)$ is constant so it is sufficient to choose q so that $\|\tilde{A}B^q(\tau)\| \le \lambda$. This can be done by choosing $q = p\bar{p}$ where $p \ge (m-1)^2$ and $\bar{p}$ is such that

$$\|B^p(\tau)\|^{\bar{p}} \le \frac{\lambda}{\|\tilde{A}\|} \tag{59}$$

This can always be accomplished because of (58).

(b) $\mathbb{N}$ changes with time

Note that (58) holds for all τ. Assuming p is chosen so that $p \ge (m-1)^2$ it is thus possible to find, for each τ, a positive integer $\bar{p}(\tau)$, for which

$$\|B^p(\tau)\|^{\bar{p}(\tau)} \le \frac{\lambda}{\|\tilde{A}\|} \tag{60}$$

Having accomplished this, a number $\bar{p}$ can easily be computed so that

$$\|B^p(\tau)\|^{\bar{p}} \leq \frac{\lambda}{\|\tilde{A}\|} \tag{61}$$

holds for all τ, since there are only a finite number of distinct strongly connected graphs on m vertices and consequently only a finite number of distinct matrices $B(\tau)$ in the set $\mathcal{B}$ defined earlier. Choosing $\bar{p}$ to be the maximum of $\bar{p}(\tau)$ with respect to τ is thus a finite computation. A value of q with the required property would then be $q = p\bar{p}$. ■

References

Açikmeşe, B., Mandić, M., & Speyer, J. L. (2014). Decentralized observers with consensus filters for distributed discrete-time linear systems. *Automatica*, *50*(4), 1037–1052.

Anderson, B. D. O., & Gevers, M. (1998). Fundamental problems in adaptive control. In *Perspectives in control: theory and applications* (pp. 9–21). London: Springer London.

Doostmohammadian, M., & Khan, U. A. (2013). On the genericity properties in distributed estimation: Topology design and sensor placement. *IEEE Journal of Selected Topics in Signal Processing*, *7*(2), 195–204.

Dorfler, F., Pasqualetti, F., & Bullo, F. (2013). Continuous-time distributed observers with discrete communication. *IEEE Journal of Selected Topics in Signal Processing*, *7*(2), 296–304.

Godsil, C., & Royle, G. (2001). *Algebraic graph theory*. Springer.

Han, W., Trentelman, H. L., Wang, Z., & Shen, Y. (2018). Towards a minimal order distributed observer for linear systems. *Systems & Control Letters*, 59–65.

Han, W., Trentelman, H. L., Wang, Z., & Shen, Y. (2019). A simple approach to distributed observer design for linear systems. *IEEE Transactions on Automatic Control*, *64*(1), 329–336.

Hespahna, J. P., & Morse, A. S. (1999). Switched systems with average dwell time. In *Proceedings of the IEEE conference on decision and control* (pp. 2655–2660).

Horn, R. A., & Johnson, C. R. (1985). *Matrix analysis*. Cambridge, England: Cambridge University Press.

Khan, U. A., & Jadbabaie, A. (2011). On the stability and optimality of distributed Kalman filters with finite-time data fusion. In *Proceedings of the 2011 American control conference* (pp. 3405–3410).

Khan, U. A., & Jadbabaie, A. (2014). Collaborative scalar-gain estimators for potentially unstable social dynamics with limited communication. *Automatica*, *50*, 1909–1914.

Kim, T., Lee, C., & Shim, H. (2020). Completely decentralized design of distributed observer for linear systems. *IEEE Transactions on Automatic Control*, *65*(11), 4664–4678.

Kim, T., Shim, H., & Cho, D. D. (2016). Distributed Luenberger observer design. In *2016 IEEE 55th conference on decision and control* (pp. 6928–6933).

Lee, J. G., & Shim, H. (2020). A distributed algorithm that finds almost best possible estimate under non-vanishing and time-varying measurement noise. *IEEE Control Systems Letters*, *4*(1), 229–234.

Li, Y., Phillips, S., & Sanfelice, R. G. (2018). Robust distributed estimationfor linear systems under intermittent information. *IEEE Transactions on Automatic Control*, *63*(4), 973–988.

Mitra, A., Richards, J. A., Bagchi, S., & Sundaram, S. (2019). Finite-time distributed state estimation over time-varying graphs: Exploiting the age-of-information. In *2019 American control conference* (pp. 4006–4011).

Mitra, A., Richards, J. A., Bagchi, S., & Sundaram, S. (2022). Distributed state estimation over time-varying graphs: Exploiting the age-of-information. *IEEE Transactions on Automatic Control*, *67*(12), 6349–6365.

Mitra, A., & Sundaram, S. (2018). Distributed observers for LTI systems. *IEEE Transactions on Automatic Control*, *63*(11), 3689–3704.

Morse, A. (1993). Dwell-time switching. In *Proceedings of the 2nd European control conference* (pp. 176–181).

Mou, S., Liu, J., & Morse, A. S. (2015). A distributed algorithm for solving a linear algebraic equation. *IEEE Transactions on Automatic Control*, *60*(11), 2863–2878.

Olfati-Reza, R. (2009). Kalman-consensus filter: Optimality, stability, and performance. In *Proc IEEE CDC* (pp. 7036–7042).

Olfati-Reza, R., & Shamma, J. S. (2005). Consensus filters for sensor networks and distributed sensor fusion. In *Proc IEEE CDC* (pp. 6698–6703).

Olfati-Saber, R. (2005). Distributed Kalman filter with embedded consensus filters. In *Proceedings of the 44th IEEE Conference on decision and control, and the european control conference* (pp. 8179–8184).

Olfati-Saber, R. (2007). Distributed Kalman filtering for sensor networks. In *Proceedings of the 46th IEEE conference on decision and control* (pp. 5492–5498).

Park, S., & Martins, N. C. (2012a). An augmented observer for the distributed estimation problem for LTI systems. In *Proceedings the 2012 American control conference* (pp. 6775–6780).

Park, S., & Martins, N. C. (2012b). Necessary and sufficient conditions for the stabilizability of a class of LTI distributed observers. In *Proceedings of the 51st IEEE conference on decision and control* (pp. 7431–7436).

Park, S., & Martins, N. C. (2017). Design of distributed LTI observers for state omniscience. *IEEE Transactions on Automatic Control*, 561–576.

Rego, F. C., Pu, Y., Alessandretti, A., Aguiar, A. P., Pascoal, A. M., & Jones, C. N. (2021). A distributed Luenberger observer for linear state feedback systems with quantized and rate-limited communications. *IEEE Transactions on Automatic Control*, *66*(9), 3922–3937.

Ugrinovskii, V. A. (2013). Distributed robust estimation over randomly switching networks using H_∞ consensus. *Automatica*, *49*, 160–168.

Wang, L., Liu, J., & Morse, A. S. (2019). A distributed observer for a continuous-time linear system. In *Proceedings of 2019 American control conference* (pp. 86–89). Philadelphia, PA, USA.

Wang, L., Liu, J., Morse, A. S., & Anderson, B. D. O. (2019). A distributed observer for a discrete-time linear system. In *Control and decision conference*.

Wang, L., & Morse, A. S. (2017). A distributed observer for a time-invariant linear system. In *Proceedings of the 2017 American control conference* (pp. 2020–2025).

Wang, L., & Morse, A. S. (2018). A distributed observer for an time-invariant linear system. *IEEE Transactions on Automatic Control*, *63*(7), 2123–2130.

Wang, L., Morse, A. S., Fullmer, D., & Liu, J. (2017). A hybrid observer for a distributed linear system with a changing neighbor graph. In *Proceedings of the 2017 IEEE conference on decision and control* (pp. 1024–1029).

Wang, L., Morse, A. S., & Liu, J. (2022). A hybrid observer for estimating the state of a distributed linear system. *Automatica*, *146*, Article 110633.

Xiao, L., Boyd, S., & Lall, S. (2005). A scheme for robust distributed sensor fusion based on average consensus. In *Proceedings of the 4th international conference on information processing in sensor networks* (pp. 63–70).

Lili Wang received the B.E. and M.S. degrees in electrical engineering from Zhejiang University, Zhejiang, China, in 2011 and 2014, respectively, and the Ph.D. degree in electrical engineering from Yale University, New Haven, CT, USA, in 2020. She is currently a Post Doctoral Research Assistant with Lilian Gilbreth fellowship in the College of Engineering, Purdue University, West Lafayette, IN, USA. Her research interests include the topic of cooperative multiagent systems, distributed computation and estimation, distributed control, and social networks.

Ji Liu received the B.S. degree in information engineering from Shanghai Jiao Tong University, Shanghai, China, in 2006 and the Ph.D. degree in electrical engineering from Yale University, New Haven, CT, USA, in 2013. He is currently an Assistant Professor with the Department of Electrical and Computer Engineering, Stony Brook University, Stony Brook, NY, USA. His current research interests include distributed control and optimization, distributed reinforcement learning, and resiliency of distributed algorithms.

Brian D.O. Anderson was born in Sydney, Australia, and educated at Sydney University in mathematics and electrical engineering, with Ph.D. in electrical engineering from Stanford University. Following graduation, he joined the faculty at Stanford University and worked in Vidar Corporation of Mountain View, California, as a staff consultant. He then returned to Australia to become a department chair in electrical engineering at the University of Newcastle. From there, he moved to the Australian National University in 1982, as the first engineering professor at that university where he is now Emeritus Professor. During his period in academia, he spent significant time working for the Australian Government, with this service including membership of the Prime Minister's Science Council under the chairmanship of three prime ministers. He also served on advisory boards or boards of various companies, including the board of the world's major supplier of cochlear implants, Cochlear Corporation, where he was a director for ten years. His awards include the Quazza Medal of the International Federation of Automatic Control (IFAC) in 1999, IEEE Control Systems Award of 1997, the 2001 IEEE James H. Mulligan, Jr. Education Medal, and the Bode Prize of the IEEE Control System Society in 1992, as well as IEEE and other best paper prizes. He is a Fellow of the Australian Academy of Science, the Australian Academy of Technological Sciences and Engineering, the Royal Society (London), and a foreign member of the US National Academy of Engineering. He holds honorary doctorates from a number of universities, including Université Catholique de Louvain, Belgium, and ETH, Zürich. He served as IFAC President from 1990 to 1993, having had earlier periods in various IFAC roles, including editor of Automatica. He was also President of the Australian Academy of Science from 1998 to 2002. His current research interests are in distributed control, social networks and econometric modeling.

A. Stephen Morse was born in Mt. Vernon, New York. He received a BSEE degree from Cornell University, a MS degree from the University of Arizona, and a Ph.D. degree from Purdue University. For three years he was associated with the Office of Control Theory and Application {OCTA} at the NASA Electronics Research Center in Cambridge, Mass. Since 1970 he has been with Yale University where he is presently the Dudley Professor of Engineering. His main interest is in system theory and he has done research in network synthesis, optimal control, multivariable control, adaptive control, urban transportation, vision-based control, hybrid and nonlinear systems, sensor networks, and coordination and control of large grouping of mobile autonomous agents. He is the recipient or co-recipient of several awards including the IEEE Control Systems Society's 1993 and 2005 George S. Axelby Best Paper Awards, the Automatica Theory/Methodology Prize, the 1999 IEEE Technical Field Award for Control Systems, the American Automatic Control Council's 2013 Richard E. Bellman Control Heritage Award, and the IFAC's 2023 Giorgio Quazza Medal. He is a Life Fellow of the IEEE, an IFAC Fellow, and a past Distinguished Lecturer of the IEEE Control System Society. He is a member of the US National Academy of Engineering and the Connecticut Academy of Science and Engineering.