

An Approach for Distributed State Estimation of LTI Systems

Aritra Mitra and Shreyas Sundaram

Abstract—We investigate the problem of distributed state estimation of a linear time-invariant (LTI) system by a network of sensors. We propose a new approach to designing distributed observers based on the following intuition: a given node (sensor) can reconstruct a certain portion of the state solely by using its own measurements together with an appropriate Luenberger observer. Hence it only needs to rely on information obtained from neighbors for estimating the portion of the state that is not locally detectable. We build on this intuition in this paper by extending the idea of the Kalman observable canonical decomposition to a setting with multiple sensors. We then construct local Luenberger observers at each node based on this decomposition, and use consensus dynamics to estimate the unobservable portions of the state at each node. This leads to an estimation scheme that achieves asymptotic state reconstruction at each node of the network for the most general class of LTI systems, sensor network topologies and sensor measurement structures.

I. INTRODUCTION

We consider a setting where the state of a linear dynamical system is monitored by a network of sensor nodes. The objective of each node is to asymptotically estimate the state of the dynamical system using its own (limited) measurements and via information exchange with neighbors. This is known as the *distributed state estimation* problem.

Different versions of this problem have been explored under varying assumptions on the system observability and network topology. For instance, [1]–[3] consider scalar stochastic dynamical systems over general graphs subject to local observability at each node, whereas [4], [5] investigate the distributed estimation problem for complete graphs. For more general stochastic systems, a Kalman filtering based approach was proposed in [6], [7], which relies on a two-step strategy - a Kalman filter based state estimate update rule, and a data fusion step based on average-consensus. The stability and performance issues of this method have been investigated in [8], [9]. A drawback of this method (and the ones in [2], [10], [11]), stems from the fact that they require a (theoretically) infinite number of data fusion iterations between two consecutive time steps of the plant dynamics in order to reach average consensus. More recently, finite-time data fusion relying on a two-time-scale strategy has been studied in [12] and [13], and an LMI-based approach has been employed in [14].

In [15] and [16], the authors propose a single-time-scale *scalar-gain estimator* for distributed observer design over undirected graphs. However, the tight coupling between the network topology and the plant dynamics typically limits the

set of unstable eigenvalues that can be accommodated by their method without violating the constraints imposed upon the range of the scalar gain parameter. In [17], the author approaches the observer design problem from a geometric perspective and provides separate necessary and sufficient conditions for consensus-based distributed observer design. In [18]–[20], the authors use single-time-scale algorithms, and work under the broadest assumptions, namely that the pair (\mathbf{A}, \mathbf{C}) is detectable, where \mathbf{A} represents the system matrix, and \mathbf{C} is the collection of all the node observation matrices. In all of these works, the authors rely on state augmentation for casting the distributed estimation problem as a problem of designing a decentralized stabilizing controller for an LTI plant, using the notion of fixed modes [21].

In this paper, we develop a new approach to designing distributed observers for LTI dynamical systems based on the following simple, yet key observation - for each node, there may be certain portions of the state that the node can reconstruct using only its local measurements. The node thus does so. For the remaining portion of the state space, the node relies on a consensus-based update rule. The key is that those nodes that can reconstruct certain states on their own act as “*source nodes*” (or “*leaders*”) in the consensus dynamics, leading the rest of the nodes to asymptotically estimate those states as well. These ideas, in a nutshell, constitute the essence of our distributed estimation strategy. We consider the most general category of systems and graphs (taken together) for which the distributed estimation problem can be solved, and develop an estimation scheme with the following appealing features: i) it provides theoretical guarantees regarding the design of asymptotically stable estimators; (ii) it results in a single-time-scale algorithm; (iii) it does not require any state augmentation; (iv) it requires only state estimates to be exchanged locally; and (v) it works under the broadest conditions on the system and network topology.

II. SYSTEM MODEL

A. Notation

A directed graph is denoted by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, \dots, N\}$ is the set of nodes and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ represents the edges. An edge from node j to node i , denoted by (j, i) , implies that node j can transmit information to node i . The neighborhood of the i -th node is defined as $\mathcal{N}_i \triangleq \{i\} \cup \{j \mid (j, i) \in \mathcal{E}\}$. The set of all eigenvalues of a matrix \mathbf{A} is denoted by $sp(\mathbf{A})$. For a set $\{\mathbf{A}_1, \dots, \mathbf{A}_n\}$ of matrices, we use $diag(\mathbf{A}_1, \dots, \mathbf{A}_n)$ to denote a block diagonal matrix with the \mathbf{A}_i 's along the diagonal. For a set $\mathcal{S} = \{s_1, \dots, s_p\} \subseteq \{1, \dots, N\}$, and a matrix $\mathbf{C} =$

$[\mathbf{C}_1^T \ \cdots \ \mathbf{C}_N^T]^T$, we define $\mathbf{C}_S \triangleq [\mathbf{C}_{s_1}^T \ \cdots \ \mathbf{C}_{s_p}^T]^T$. We use the star notation to avoid writing matrices that are either unimportant or that can be inferred from context. We use \mathbf{I}_r to indicate an identity matrix of dimension $r \times r$. Throughout the rest of this paper, we use the terms ‘nodes’ and ‘observers’ interchangeably.

B. Problem Formulation

Consider the discrete-time linear dynamical system¹

$$\mathbf{x}[k+1] = \mathbf{A}\mathbf{x}[k], \quad (1)$$

where $k \in \mathbb{N}$ is the discrete-time index, $\mathbf{x}[k] \in \mathbb{R}^n$ is the state vector and $\mathbf{A} \in \mathbb{R}^{n \times n}$ is the system matrix. The system is monitored by a network $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ consisting of N nodes. The i -th node has a measurement of the state, given by

$$\mathbf{y}_i[k] = \mathbf{C}_i \mathbf{x}[k], \quad (2)$$

where $\mathbf{y}_i[k] \in \mathbb{R}^{r_i}$ and $\mathbf{C}_i \in \mathbb{R}^{r_i \times n}$. We denote $\mathbf{y}[k] = [\mathbf{y}_1^T[k] \ \cdots \ \mathbf{y}_N^T[k]]^T$, and $\mathbf{C} = [\mathbf{C}_1^T \ \cdots \ \mathbf{C}_N^T]^T$.

The task of each node $i \in \{1, \dots, N\}$ is to estimate the entire system state $\mathbf{x}[k]$. However, if the pair $(\mathbf{A}, \mathbf{C}_i)$ is not detectable, then node i cannot estimate the true state of the plant solely based on its own local measurements, thereby necessitating information exchange with neighbors. Let $\hat{\mathbf{x}}_i[k]$ denote the state estimate of node i . We refer to the network of nodes maintaining and updating their estimates as a *distributed observer*. We shall borrow the following definitions from [20] for our analysis.

Definition 1 (Omniscience): A distributed observer is said to achieve *omniscience* if $\lim_{k \rightarrow \infty} \|\hat{\mathbf{x}}_i[k] - \mathbf{x}[k]\| = 0, \forall i \in \{1, \dots, N\}$, i.e., the state estimate maintained by each node asymptotically converges to the true state of the plant. \square

Definition 2 (Source Component): Given a directed graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, a source component $(\mathcal{V}_s, \mathcal{E}_s)$ is defined as a strongly connected component of \mathcal{G} such that there are no edges from $\mathcal{V} \setminus \mathcal{V}_s$ to \mathcal{V}_s . \square

Let there be p source components of \mathcal{G} , denoted by $\{(\mathcal{V}_i, \mathcal{E}_i)\}_{i \in \{1, \dots, p\}}$. The subsystem associated with the i -th source component is given by the pair $(\mathbf{A}, \mathbf{C}_{\mathcal{V}_i})$. For the subsequent development, it should be noted that by a system (\mathbf{A}, \mathbf{C}) , we refer to a pair (\mathbf{A}, \mathbf{C}) satisfying equations (1) and (2).

III. DESCRIPTION OF THE ESTIMATION SCHEME

In this section, we propose a distributed observer for achieving omniscience. For presenting the key ideas while reducing notational complexity, we shall make the following assumption.²

¹We omit noise terms in the dynamics for ease of exposition (e.g., as in [17]–[20]). However, it can be shown that the methods developed in this paper lead to bounded mean square estimation error in the presence of i.i.d. noise with bounded second moments. Further, it should be noted that our proposed techniques will be equally applicable to continuous-time systems, with straightforward modifications.

²We shall later argue that the development can be easily extended to arbitrary graph topologies.

Assumption 1: The graph \mathcal{G} is strongly connected, i.e., there exists a directed path from any node i to any other node j , where $i, j \in \mathcal{V}$.

We begin our development of the proposed estimation scheme by providing a generalization of the Kalman observable canonical form to a setting with multiple sensors.

A. Multi-Sensor Observable Canonical Decomposition

Given a system matrix \mathbf{A} and a set of N sensors, such that the i -th sensor has an observation matrix given by \mathbf{C}_i , we introduce the notion of a *multi-sensor observable canonical decomposition* in this section. The basic philosophy underlying such a decomposition is as follows: given a list of indexed sensors, perform an observable canonical decomposition with respect to the first sensor. Then, identify the observable portion of the state space with respect to sensor 2 *within the unobservable subspace of sensor 1*, and repeat the process until the last sensor is reached. Thus, one needs to perform N observable canonical decompositions, one for each sensor, with the last decomposition revealing the portions of the state space that can and cannot be observed using the cumulative measurements of all the sensors. In this context, consider the following result.

Proposition 1: Given a system matrix \mathbf{A} , and a set of N sensor observation matrices $\mathbf{C}_1, \mathbf{C}_2, \dots, \mathbf{C}_N$, define $\mathbf{C} \triangleq [\mathbf{C}_1^T \ \cdots \ \mathbf{C}_N^T]^T$. Then, there exists a similarity transformation matrix \mathcal{T} which transforms the pair (\mathbf{A}, \mathbf{C}) to $(\bar{\mathbf{A}}, \bar{\mathbf{C}})$, such that

$$\bar{\mathbf{A}} = \left[\begin{array}{c|ccc} \mathbf{A}_{1\mathcal{O}} & & & \mathbf{0} \\ \mathbf{A}_{2\mathcal{O}} & & & \mathbf{0} \\ \star & & \ddots & \vdots \\ & \star & \mathbf{A}_{(N-1)\mathcal{O}} & \mathbf{0} \\ & & \star & \mathbf{A}_{N\mathcal{O}} \\ & & & \star & \mathbf{A}_{U\mathcal{O}} \end{array} \right],$$

$$\bar{\mathbf{C}} = \left[\begin{array}{c} \bar{\mathbf{C}}_1 \\ \bar{\mathbf{C}}_2 \\ \vdots \\ \bar{\mathbf{C}}_N \end{array} \right] = \left[\begin{array}{c|ccc} \mathbf{C}_{1\mathcal{O}} & & & \mathbf{0} \\ \mathbf{C}_{21\mathcal{O}} & \mathbf{C}_{2\mathcal{O}} & & \mathbf{0} \\ \vdots & \vdots & & \vdots \\ \mathbf{C}_{N1\mathcal{O}} & \mathbf{C}_{N2\mathcal{O}} & \cdots & \mathbf{C}_{N(N-1)\mathcal{O}} & \mathbf{C}_{N\mathcal{O}} & | & \mathbf{0} \end{array} \right]. \quad (3)$$

Furthermore, the following properties hold: (i) the pair $(\mathbf{A}_{i\mathcal{O}}, \mathbf{C}_{i\mathcal{O}})$ is observable $\forall i \in \{1, 2, \dots, N\}$; and (ii) the matrix $\mathbf{A}_{U\mathcal{O}}$ describes the dynamics of the unobservable subspace of the pair (\mathbf{A}, \mathbf{C}) . \square

We provide a proof of the above result (which details the steps associated with the multi-sensor observable canonical decomposition) in [22]. In the following section, we discuss how the multi-sensor observable canonical decomposition is applicable to the problem of achieving omniscience.

B. Observer Design

Using the matrix \mathcal{T} identified in Proposition 1, we perform the coordinate transformation $\mathbf{x}[k] = \mathcal{T}\mathbf{z}[k]$ to obtain

$$\begin{aligned} \mathbf{z}[k+1] &= \bar{\mathbf{A}}\mathbf{z}[k] \\ \mathbf{y}_i[k] &= \bar{\mathbf{C}}_i \mathbf{z}[k], \quad \forall i \in \{1, \dots, N\}, \end{aligned} \quad (4)$$

where $\bar{\mathbf{A}} = \mathcal{T}^{-1}\mathbf{A}\mathcal{T}$ and $\bar{\mathbf{C}}_i = \mathbf{C}_i\mathcal{T} = [\mathbf{C}_{i1\mathcal{O}} \ \mathbf{C}_{i2\mathcal{O}} \ \cdots \ \mathbf{C}_{i(i-1)\mathcal{O}} \ \mathbf{C}_{i\mathcal{O}} \ | \ \mathbf{0}]$ are given by (3). The vector $\mathbf{z}[k]$ assumes the following structure (commensurate with the structure of $\bar{\mathbf{A}}$ in (3)):

$$\mathbf{z}[k] = \begin{bmatrix} \mathbf{z}_{\mathcal{O}}^{(1)}[k] \\ \vdots \\ \mathbf{z}_{\mathcal{O}}^{(N)}[k] \\ \mathbf{z}_{\mathcal{U}\mathcal{O}}[k] \end{bmatrix}. \quad (5)$$

Here, $\mathbf{z}_{\mathcal{U}\mathcal{O}}[k]$ is precisely the unobservable portion of the state $\mathbf{z}[k]$, with respect to the pair (\mathbf{A}, \mathbf{C}) . We call $\mathbf{z}_{\mathcal{O}}^{(j)}[k]$ the j -th sub-state, and $\mathbf{z}_{\mathcal{U}\mathcal{O}}[k]$ the unobservable sub-state. Notice that based on the multi-sensor observable canonical decomposition, there is a one-to-one correspondence between a node j and its associated sub-state $\mathbf{z}_{\mathcal{O}}^{(j)}[k]$. Accordingly, node j will be viewed as the source of information of its corresponding sub-state $\mathbf{z}_{\mathcal{O}}^{(j)}[k]$, and is tasked with the responsibility of estimating this sub-state. For each of the N sub-states, we thus have a unique source of information (based on the initial labeling of the nodes). However, there is no unique source of information for the unobservable sub-state $\mathbf{z}_{\mathcal{U}\mathcal{O}}[k]$, as this portion of the state does not correspond to the observable subspace of any of the nodes in the network. Each node will thus maintain an estimate of $\mathbf{z}_{\mathcal{U}\mathcal{O}}[k]$, which it updates as a linear function of its *own* estimates of each of the N sub-states $\mathbf{z}_{\mathcal{O}}^{(j)}[k], \forall j \in \{1, 2, \dots, N\}$.

Remark 1: It should be noted that a given sub-state $\mathbf{z}_{\mathcal{O}}^{(j)}[k]$ in equation (5) might be of zero dimension (i.e., the sub-state can be empty). For instance, this can happen if its corresponding source of information, namely node j , has no measurements, i.e., if $\mathbf{C}_j = 0$. \square

Define $\hat{\mathbf{z}}_{i\mathcal{O}}^{(j)}[k]$ as the estimate of the j -th sub-state maintained by the i -th node. First, based on equations (3), (4) and (5), we observe that the dynamics of the i -th sub-state are governed by the equations

$$\begin{aligned} \mathbf{z}_{\mathcal{O}}^{(i)}[k+1] &= \mathbf{A}_{i\mathcal{O}}\mathbf{z}_{\mathcal{O}}^{(i)}[k] + \sum_{j=1}^{i-1} \mathbf{A}_{ij}\mathbf{z}_{\mathcal{O}}^{(j)}[k], \\ \mathbf{y}_i[k] &= \mathbf{C}_{i\mathcal{O}}\mathbf{z}_{\mathcal{O}}^{(i)}[k] + \sum_{j=1}^{i-1} \mathbf{C}_{ij}\mathbf{z}_{\mathcal{O}}^{(j)}[k], \end{aligned} \quad (6)$$

where \mathbf{A}_{ij} are matrices represented by \star 's in (3). The reader is referred to the proof of Proposition 1 in [22] for a mathematical description of these matrices. Also, note that the unobservable sub-state $\mathbf{z}_{\mathcal{U}\mathcal{O}}[k]$ is governed by the dynamics

$$\mathbf{z}_{\mathcal{U}\mathcal{O}}[k+1] = \mathbf{A}_{\mathcal{U}\mathcal{O}}\mathbf{z}_{\mathcal{U}\mathcal{O}}[k] + \sum_{j=1}^N \mathbf{A}_j\mathbf{z}_{\mathcal{O}}^{(j)}[k], \quad (7)$$

where the matrices \mathbf{A}_j (represented by \star 's in equation (3)) describe the coupling that exists between the unobservable sub-state $\mathbf{z}_{\mathcal{U}\mathcal{O}}[k]$ and each of the N sub-states $\mathbf{z}_{\mathcal{O}}^{(j)}[k]$. The estimation policy adopted by the i -th node is as follows - it uses a Luenberger-style update rule for updating its associated sub-state $\hat{\mathbf{z}}_{i\mathcal{O}}^{(i)}[k]$, and a consensus based scheme for

updating all other sub-states $\hat{\mathbf{z}}_{i\mathcal{O}}^{(j)}[k]$, where $j \in \{1, \dots, N\} \setminus \{i\}$. Based on the dynamics (6), the Luenberger observer at node i is constructed as

$$\begin{aligned} \hat{\mathbf{z}}_{i\mathcal{O}}^{(i)}[k+1] &= \mathbf{A}_{i\mathcal{O}}\hat{\mathbf{z}}_{i\mathcal{O}}^{(i)}[k] + \sum_{j=1}^{i-1} \mathbf{A}_{ij}\hat{\mathbf{z}}_{i\mathcal{O}}^{(j)}[k] \\ &\quad + \mathbf{L}_i(\mathbf{y}_i[k] - (\mathbf{C}_{i\mathcal{O}}\hat{\mathbf{z}}_{i\mathcal{O}}^{(i)}[k] + \sum_{j=1}^{i-1} \mathbf{C}_{ij}\hat{\mathbf{z}}_{i\mathcal{O}}^{(j)}[k])), \end{aligned} \quad (8)$$

where $\mathbf{L}_i \in \mathbb{R}^{o_i \times r_i}$ is a gain matrix which needs to be designed (recall $\mathbf{z}_{\mathcal{O}}^{(i)} \in \mathbb{R}^{o_i}$). For estimation of the j -th sub-state, where $j \in \{1, \dots, N\} \setminus \{i\}$, the i -th node again mimics the first equation in (6), but this time relies on consensus dynamics of the form

$$\hat{\mathbf{z}}_{i\mathcal{O}}^{(j)}[k+1] = \underbrace{\mathbf{A}_{j\mathcal{O}} \sum_{l \in \mathcal{N}_i} w_{il}^j \hat{\mathbf{z}}_{i\mathcal{O}}^{(j)}[k]}_{\text{consensus term}} + \underbrace{\sum_{l=1}^{j-1} \mathbf{A}_{jl} \hat{\mathbf{z}}_{i\mathcal{O}}^{(l)}[k]}_{\text{coupling term}}, \quad (9)$$

where w_{il}^j is the weight the i -th node associates with the l -th node, for the estimation of the j -th sub-state. The weights are non-negative and satisfy

$$\sum_{l \in \mathcal{N}_i} w_{il}^j = 1, \quad \forall j \in \{1, \dots, N\} \setminus \{i\}. \quad (10)$$

In equation (9), the first term is a standard consensus term, while the second term has been introduced specifically to account for the coupling that exists between a given sub-state j , and sub-states 1 to $j-1$ (as given by (6)). Let $\hat{\mathbf{z}}_{i\mathcal{U}\mathcal{O}}[k]$ denote the estimate of the unobservable sub-state $\mathbf{z}_{\mathcal{U}\mathcal{O}}[k]$ maintained by the i -th node. Mimicking equation (7), each node i uses the following rule to update $\hat{\mathbf{z}}_{i\mathcal{U}\mathcal{O}}[k]$:

$$\hat{\mathbf{z}}_{i\mathcal{U}\mathcal{O}}[k+1] = \mathbf{A}_{\mathcal{U}\mathcal{O}}\hat{\mathbf{z}}_{i\mathcal{U}\mathcal{O}}[k] + \sum_{j=1}^N \mathbf{A}_j\hat{\mathbf{z}}_{i\mathcal{O}}^{(j)}[k]. \quad (11)$$

In summary, equations (8), (9) and (11) together form the observer for the state $\mathbf{z}[k] = \mathcal{T}^{-1}\mathbf{x}[k]$ maintained by each node i .

C. Error Dynamics at the i -th Node

Define $\mathbf{e}_{i\mathcal{O}}^{(j)}[k] \triangleq \hat{\mathbf{z}}_{i\mathcal{O}}^{(j)}[k] - \mathbf{z}_{\mathcal{O}}^{(j)}[k]$ as the error in estimation of the j -th sub-state by the i -th node. Using equations (6) and (8), we obtain the error in the Luenberger observer dynamics at the i -th node as

$$\mathbf{e}_{i\mathcal{O}}^{(i)}[k+1] = (\mathbf{A}_{i\mathcal{O}} - \mathbf{L}_i\mathbf{C}_{i\mathcal{O}})\mathbf{e}_{i\mathcal{O}}^{(i)}[k] + \sum_{j=1}^{i-1} (\mathbf{A}_{ij} - \mathbf{L}_i\mathbf{C}_{ij\mathcal{O}})\mathbf{e}_{i\mathcal{O}}^{(j)}[k]. \quad (12)$$

Similarly, noting that $\mathbf{A}_{j\mathcal{O}} = \mathbf{A}_{j\mathcal{O}} \sum_{l \in \mathcal{N}_i} w_{il}^j$ (based on equation (10)), and using equations (6) and (9), we obtain the following consensus error dynamics at node i , $\forall j \in \{1, \dots, N\} \setminus \{i\}$:

$$\mathbf{e}_{i\mathcal{O}}^{(j)}[k+1] = \mathbf{A}_{j\mathcal{O}} \sum_{l \in \mathcal{N}_i} w_{il}^j \mathbf{e}_{i\mathcal{O}}^{(j)}[k] + \sum_{l=1}^{j-1} \mathbf{A}_{jl} \mathbf{e}_{i\mathcal{O}}^{(l)}[k]. \quad (13)$$

Define $\mathbf{e}_{i\mathcal{O}}[k] \triangleq \hat{\mathbf{z}}_{i\mathcal{O}}[k] - \mathbf{z}_{i\mathcal{O}}[k]$ as the error in estimation of the unobservable sub-state $\mathbf{z}_{i\mathcal{O}}[k]$ by the i -th node. Using (7) and (11), we obtain the following error dynamics for the unobservable sub-state at node i :

$$\mathbf{e}_{i\mathcal{O}}[k+1] = \mathbf{A}_{i\mathcal{O}}\mathbf{e}_{i\mathcal{O}}[k] + \sum_{j=1}^N \mathbf{A}_j \mathbf{e}_{i\mathcal{O}}^{(j)}[k]. \quad (14)$$

IV. ANALYSIS OF THE ESTIMATION SCHEME

We now state and prove our main result which provides necessary and sufficient conditions for our proposed observer to achieve omniscience.

Theorem 1: Consider a system (\mathbf{A}, \mathbf{C}) , and a graph \mathcal{G} satisfying Assumption 1. Then, for each node $i \in \{1, 2, \dots, N\}$, there exists a choice of observer gain matrix \mathbf{L}_i , and consensus weights w_{il}^j , $j \in \{1, 2, \dots, N\} \setminus \{i\}$, $l \in \mathcal{N}_i$, such that the distributed observer given by equations (8), (9), and (11) achieves omniscience *if and only if* the pair (\mathbf{A}, \mathbf{C}) is detectable. \square

Proof: “ \Leftarrow ” Consider the composite error in estimation of sub-state j by all of the nodes in \mathcal{V} , defined as

$$\mathbf{E}_{\mathcal{O}}^{(j)}[k] \triangleq \begin{bmatrix} \mathbf{e}_{1\mathcal{O}}^{(j)}[k] \\ \mathbf{e}_{2\mathcal{O}}^{(j)}[k] \\ \vdots \\ \mathbf{e}_{N\mathcal{O}}^{(j)}[k] \end{bmatrix}. \quad (15)$$

We will prove that $\mathbf{E}_{\mathcal{O}}^{(j)}[k]$ converges to zero asymptotically $\forall j \in \{1, \dots, N\}$ (recall that there are precisely N nodes in the network, each responsible for estimating a certain sub-state). We prove by induction on j . Consider the base case $j = 1$, i.e., the estimation of the first sub-state. Let the index set $\{1, k_1, k_2, \dots, k_{N-1}\}$ represent a topological ordering³ consistent with a spanning tree rooted at node 1 (the source of information for sub-state 1). Note that based on Assumption 1, it is always possible to find such a spanning tree. Next, consider the composite error vector

$$\bar{\mathbf{E}}_{\mathcal{O}}^{(1)}[k] = \begin{bmatrix} \mathbf{e}_{1\mathcal{O}}^{(1)}[k] \\ \mathbf{e}_{k_1\mathcal{O}}^{(1)}[k] \\ \vdots \\ \mathbf{e}_{k_{N-1}\mathcal{O}}^{(1)}[k] \end{bmatrix} = \begin{bmatrix} \mathbf{e}_{1\mathcal{O}}^{(1)}[k] \\ \tilde{\mathbf{E}}_{\mathcal{O}}^{(1)}[k] \end{bmatrix}, \quad (16)$$

where $\tilde{\mathbf{E}}_{\mathcal{O}}^{(1)}[k] \triangleq [\mathbf{e}_{k_1\mathcal{O}}^{(1)}[k]^T \dots \mathbf{e}_{k_{N-1}\mathcal{O}}^{(1)}[k]^T]^T$. Note that $\bar{\mathbf{E}}_{\mathcal{O}}^{(1)}[k]$ is simply a permutation of the rows of $\mathbf{E}_{\mathcal{O}}^{(1)}[k]$. Based on the error dynamics equations given by (12) and (13), we obtain

³Such an ordering results when a standard Breadth-First Search (BFS) [23] algorithm is applied to the graph \mathcal{G} , with node 1 as the root node of the tree. Specifically, the order represents the order in which the nodes are added to the spanning tree when the BFS algorithm is implemented, i.e., node k_1 would be added first, followed by node k_2 and so on. This ordering naturally leads to a lower triangular adjacency matrix for the constructed spanning tree.

$$\begin{bmatrix} \mathbf{e}_{1\mathcal{O}}^{(1)}[k+1] \\ \bar{\mathbf{E}}_{\mathcal{O}}^{(1)}[k+1] \end{bmatrix} = \underbrace{\begin{bmatrix} (\mathbf{A}_{1\mathcal{O}} - \mathbf{L}_1 \mathbf{C}_{1\mathcal{O}}) & \mathbf{0} \\ \mathbf{W}_{21}^1 \otimes \mathbf{A}_{1\mathcal{O}} & \mathbf{W}_{22}^1 \otimes \mathbf{A}_{1\mathcal{O}} \end{bmatrix}}_{\mathbf{M}_1} \begin{bmatrix} \mathbf{e}_{1\mathcal{O}}^{(1)}[k] \\ \bar{\mathbf{E}}_{\mathcal{O}}^{(1)}[k] \end{bmatrix},$$

where the entries of the weight matrix $\mathbf{W}^1 = [\mathbf{W}_{21}^1 \quad \mathbf{W}_{22}^1]$ are populated by the appropriate weights defined by equation (13) (note that $\mathbf{W}^1 \in \mathbb{R}^{(N-1) \times N}$ and \mathbf{W}_{21}^1 is the first column of \mathbf{W}^1). Notice that $sp(\mathbf{M}_1) = sp(\mathbf{A}_{1\mathcal{O}} - \mathbf{L}_1 \mathbf{C}_{1\mathcal{O}}) \cup sp(\mathbf{W}_{22}^1 \otimes \mathbf{A}_{1\mathcal{O}})$. By construction, the pair $(\mathbf{A}_{1\mathcal{O}}, \mathbf{C}_{1\mathcal{O}})$ is observable. Thus, it is always possible to find a gain matrix \mathbf{L}_1 such that $(\mathbf{A}_{1\mathcal{O}} - \mathbf{L}_1 \mathbf{C}_{1\mathcal{O}})$ is Schur stable. Next, we impose the constraint that for the estimation of sub-state 1, non-zero consensus weights are assigned to only the branches of the spanning tree consistent with the ordering $\{1, k_1, k_2, \dots, k_{N-1}\}$, i.e., a node listens to only its parent in such a tree. In this way, \mathbf{W}_{22}^1 becomes lower triangular with eigenvalues equal to zero, without violating the stochasticity condition imposed on \mathbf{W}^1 by equation (10). We conclude that by an appropriate choice of consensus weights, we can achieve $\Lambda_U(\mathbf{W}_{22}^1 \otimes \mathbf{A}_{1\mathcal{O}}) = \emptyset$ (even if $\Lambda_U(\mathbf{A}_{1\mathcal{O}}) \neq \emptyset$).⁴

Thus, \mathbf{M}_1 can be made Schur stable and hence $\lim_{k \rightarrow \infty} \bar{\mathbf{E}}_{\mathcal{O}}^{(1)}[k] = \mathbf{0}$, implying $\lim_{k \rightarrow \infty} \mathbf{E}_{\mathcal{O}}^{(1)}[k] = \mathbf{0}$ (one is just a permutation of the other). Thus, the base case is proven. Next, suppose that $\mathbf{E}_{\mathcal{O}}^{(j)}[k]$ converges to zero asymptotically $\forall j \in \{1, \dots, p-1\}$, where $1 \leq p-1 \leq N-1$. Consider the following composite error vector for the p -th sub-state:

$$\bar{\mathbf{E}}_{\mathcal{O}}^{(p)}[k] = \begin{bmatrix} \mathbf{e}_{p\mathcal{O}}^{(p)}[k] \\ \mathbf{e}_{m_1\mathcal{O}}^{(p)}[k] \\ \vdots \\ \mathbf{e}_{m_{N-1}\mathcal{O}}^{(p)}[k] \end{bmatrix} = \begin{bmatrix} \mathbf{e}_{p\mathcal{O}}^{(p)}[k] \\ \tilde{\mathbf{E}}_{\mathcal{O}}^{(p)}[k] \end{bmatrix}, \quad (17)$$

where the index set $\{p, m_1, m_2, \dots, m_{N-1}\}$ represents a topological ordering of the nodes of \mathcal{V} to obtain a spanning tree rooted at node p (the source of information for sub-state p), and $\tilde{\mathbf{E}}_{\mathcal{O}}^{(p)}[k] \triangleq [\mathbf{e}_{m_1\mathcal{O}}^{(p)}[k]^T \dots \mathbf{e}_{m_{N-1}\mathcal{O}}^{(p)}[k]^T]^T$. From the error dynamics equations given by (12) and (13), we obtain

$$\bar{\mathbf{E}}_{\mathcal{O}}^{(p)}[k+1] = \mathbf{M}_p \bar{\mathbf{E}}_{\mathcal{O}}^{(p)}[k] + \sum_{l=1}^{p-1} \mathbf{H}_{pl} \bar{\mathbf{E}}_{\mathcal{O}}^{(pl)}[k], \quad (18)$$

where

$$\mathbf{M}_p = \begin{bmatrix} (\mathbf{A}_{p\mathcal{O}} - \mathbf{L}_p \mathbf{C}_{p\mathcal{O}}) & \mathbf{0} \\ \mathbf{W}_{21}^p \otimes \mathbf{A}_{p\mathcal{O}} & \mathbf{W}_{22}^p \otimes \mathbf{A}_{p\mathcal{O}} \end{bmatrix}, \quad (19)$$

$$\mathbf{H}_{pl} = \text{diag}(\mathbf{A}_{pl} - \mathbf{L}_p \mathbf{C}_{pl\mathcal{O}}, \mathbf{I}_{N-1} \otimes \mathbf{A}_{pl}), \quad (20)$$

$$\bar{\mathbf{E}}_{\mathcal{O}}^{(pl)}[k] = \begin{bmatrix} \mathbf{e}_{p\mathcal{O}}^{(l)}[k] \\ \mathbf{e}_{m_1\mathcal{O}}^{(l)}[k] \\ \vdots \\ \mathbf{e}_{m_{N-1}\mathcal{O}}^{(l)}[k] \end{bmatrix}. \quad (21)$$

⁴Here, we use the result that if $\mathbf{A} \in \mathbb{R}^{n \times n}$ and $\mathbf{B} \in \mathbb{R}^{m \times m}$, then the eigenvalues of the Kronecker product $\mathbf{A} \otimes \mathbf{B} \in \mathbb{R}^{mn \times mn}$ are the mn numbers $\lambda_i(\mathbf{A})\lambda_j(\mathbf{B})$, ($i = 1, \dots, n; j = 1, \dots, m$) [24].

By following the same train of logic as the base case, one concludes that \mathbf{M}_p can be made Schur stable by appropriate choices of the observer gain matrix \mathbf{L}_p , and consensus weight matrix $\mathbf{W}^p = [\mathbf{W}_{21}^p \quad \mathbf{W}_{22}^p]$ (note that $\mathbf{W}^p \in \mathbb{R}^{(N-1) \times N}$ and \mathbf{W}_{21}^p is the first column of \mathbf{W}^p). Specifically, non-zero weights are assigned in \mathbf{W}^p only on the branches of the tree rooted at node p , consistent with the topological ordering. Notice that $\bar{\mathbf{E}}_{\mathcal{O}}^{(p)}[k]$ is simply a permutation of the rows of $\mathbf{E}_{\mathcal{O}}^{(l)}[k]$ (permuted to match the order of indices in $\bar{\mathbf{E}}_{\mathcal{O}}^{(p)}[k]$). Further, based on our induction hypothesis, $\mathbf{E}_{\mathcal{O}}^{(l)}[k]$ converges to zero asymptotically (since $1 \leq l \leq p-1$). Thus, by Input to State Stability (ISS), we conclude that $\bar{\mathbf{E}}_{\mathcal{O}}^{(p)}[k]$, and hence $\mathbf{E}_{\mathcal{O}}^{(p)}[k]$, converges to zero asymptotically. We have thus proven that the composite estimation error for every sub-state asymptotically approaches zero, i.e., $\lim_{k \rightarrow \infty} \mathbf{e}_{i\mathcal{O}}^{(j)}[k] = \mathbf{0}, \forall i, j \in \{1, \dots, N\}$.

Finally, consider the error in estimation of the unobservable sub-state $\mathbf{z}_{\mathcal{UO}}[k]$ (given by equation (14)). As the pair (\mathbf{A}, \mathbf{C}) is detectable, it follows from Proposition 1 that the matrix $\mathbf{A}_{\mathcal{UO}}$ in (14) is stable. Invoking ISS, we have that $\lim_{k \rightarrow \infty} \mathbf{e}_{\mathcal{UO}}[k] = \mathbf{0}, \forall i \in \{1, \dots, N\}$. We conclude that every node in the network can asymptotically estimate $\mathbf{z}[k]$, and hence $\mathbf{x}[k]$, as $\mathbf{x}[k] = \mathcal{T}\mathbf{z}[k]$.

“ \implies ” The proof of necessity follows from standard detectability arguments. Specifically, if the pair (\mathbf{A}, \mathbf{C}) is not detectable, then there exists some initial system state that causes the system outputs of all the nodes in the graph to be zero for all time, without the state decaying to zero. Thus, no node in \mathcal{V} can ever distinguish this case from the one where the initial state is identically zero. Thus, omniscience cannot be achieved. It is easy to see that the detectability of the pair (\mathbf{A}, \mathbf{C}) for strongly connected graphs is in fact necessary for any estimation scheme to achieve omniscience.⁵ ■

A. Summary of the Estimation Scheme

Our proposed distributed observer scheme for strongly connected graphs can be broadly decomposed into two main phases, namely, the design phase, and the distributed estimation phase, summarized as follows.

Design Phase:

- The multi-sensor observable canonical decomposition is performed, yielding the state $\mathbf{z}[k] = \mathcal{T}^{-1}\mathbf{x}[k]$.
- Based on this transformation, each node is associated with a sub-state of $\mathbf{z}[k]$ that it is responsible for estimating. Recall that there are precisely N sub-states, one corresponding to each node in the network; some of these sub-states might be empty.
- For the estimation of a given sub-state, we construct a spanning tree rooted at the specific node which acts as the source of information for that sub-state. The resulting spanning tree guides the construction of the consensus weight matrix to be used for the estimation of that particular sub-state. We construct one spanning tree for the estimation of each non-empty sub-state.

⁵A more general version of this result for arbitrary graph topologies will be presented in a later section.

Estimation Phase (Run-time):

- Each node employs a Luenberger observer for constructing an estimate of its corresponding sub-state, and runs consensus dynamics for estimating the sub-states corresponding to the remaining nodes in the network. These dynamics are given by equations (8), (9) and (11).

Remark 2: While the observer design procedure we have outlined (involving the multi-sensor decomposition, design of local observer gains, construction of spanning trees and selection of consensus weights) can be readily implemented in a centralized manner, it may also be possible to perform these steps in a distributed fashion. This would require the nodes to assign themselves unique identifiers (or labels) and execute the multi-sensor decomposition in a round-robin fashion, followed by a distributed construction of spanning trees. □

B. A Compact Representation of the Proposed Observer

In this section, we combine the update equations (8), (9) and (11) to obtain a compact representation of our distributed observer. To do so, we need to first introduce some notation. Accordingly, let $\mathbf{B}_j = [\mathbf{0} \dots \mathbf{I}_{o_j} \dots \mathbf{0}]$ be the matrix which extracts the j -th sub-state from the transformed state vector $\mathbf{z}[k]$, i.e., $\mathbf{z}_{\mathcal{O}}^{(j)}[k] = \mathbf{B}_j\mathbf{z}[k]$. Similarly, let $\mathbf{B}_{\mathcal{UO}}$ be such that $\mathbf{z}_{\mathcal{UO}}[k] = \mathbf{B}_{\mathcal{UO}}\mathbf{z}[k]$. Define $\mathbb{B} \triangleq \text{diag}(\mathbf{B}_1, \dots, \mathbf{B}_N, \mathbf{B}_{\mathcal{UO}})$. Next, notice that the transformed system matrix $\bar{\mathbf{A}}$ in equation (3) can be written as $\bar{\mathbf{A}} = \bar{\mathbf{A}}_1 + \bar{\mathbf{A}}_2$, where $\bar{\mathbf{A}}_2 = \text{diag}(\mathbf{A}_{1\mathcal{O}}, \dots, \mathbf{A}_{N\mathcal{O}}, \mathbf{A}_{\mathcal{UO}})$, and $\bar{\mathbf{A}}_1$ is a block lower-triangular matrix given by $\bar{\mathbf{A}} - \bar{\mathbf{A}}_2$. Let \mathbf{w}_{il} (where $l \in \mathcal{N}_i \setminus \{i\}$) be the vector of weights node i associates with a neighbor l for the estimation of the transformed state $\mathbf{z}[k]$. Based on our estimation scheme, note that at any given time-step k , node i does not use the estimates received from its neighbors at time-step k for estimating $\mathbf{z}_{\mathcal{O}}^{(i)}[k]$ and $\mathbf{z}_{\mathcal{UO}}[k]$, and hence these weight vectors assume the following form: $\mathbf{w}_{il} = [w_{il}^1, \dots, w_{il}^{i-1}, 0, w_{il}^{i+1}, \dots, w_{il}^N, 0]^T, \forall l \in \mathcal{N}_i \setminus \{i\}$. Also, notice that the element w_{il}^j is not present in the vector if the j -th sub-state is empty (i.e., of dimension 0). Similarly, let \mathbf{w}_{ii} be a vector with a ‘1’ in the elements corresponding to the i -th sub-state and the unobservable sub-state $\mathbf{z}_{\mathcal{UO}}[k]$, and zeroes at all other positions. Finally, defining $\mathbb{H}_i \triangleq [\mathbf{0}^T \dots \mathbf{L}_i^T \dots \mathbf{0}^T]^T$, using equations (8), (9) and (11), and noting that $\mathbf{z}[k] = \mathcal{T}^{-1}\mathbf{x}[k]$, we obtain the following overall state estimate update rule at node i :

$$\hat{\mathbf{x}}_i[k+1] = \mathcal{T}\bar{\mathbf{A}}_1\mathcal{T}^{-1}\hat{\mathbf{x}}_i[k] + \underbrace{\mathcal{T}\mathbb{H}_i(\mathbf{y}_i[k] - \mathbf{C}_i\hat{\mathbf{x}}_i[k])}_{\text{innovation term}} + \underbrace{\sum_{l \in \mathcal{N}_i} \mathbb{G}_{il}\hat{\mathbf{x}}_l[k]}_{\text{“consensus term”}} \quad (22)$$

where $\hat{\mathbf{x}}_i[k]$ denotes the estimate of the state $\mathbf{x}[k]$ maintained by node i , and

$$\mathbf{C}_i = (\mathbf{C}_{i\mathcal{O}}\mathbf{B}_i + \sum_{j=1}^{i-1} \mathbf{C}_{ij\mathcal{O}}\mathbf{B}_j)\mathcal{T}^{-1}, \mathbb{G}_{il} = \mathcal{T}\bar{\mathbf{A}}_2\mathbb{B}(\mathbf{w}_{il} \otimes \mathcal{T}^{-1}). \quad (23)$$

Remark 3: From the structure of our overall estimator at node i (equation (22)), it is easy to see that the estimator maintained at each node has dimension equal to n (i.e., equal

to that of the state). Thus, our approach alleviates the need to construct augmented observers, as in [20]. \square

In the following section, we discuss how our estimation scheme for strongly connected graphs can be extended to arbitrary directed network topologies.

C. Extension to General Network Topologies

Our distributed observer design can be extended to general directed networks by first decomposing \mathcal{G} into its strong components, and identifying each of the source components. Next, within a given source component, one simply follows the observer design procedure outlined in Section IV-A for a strongly connected graph, to obtain an estimator of the form (22) for each node within the source component. Define $\mathcal{S} \triangleq \bigcup_{i=1}^p \mathcal{V}_i$ to be the set of all nodes that belong to the source components of \mathcal{G} . Let each node in $i \in \mathcal{V} \setminus \mathcal{S}$ employ a pure consensus strategy of the form

$$\hat{\mathbf{x}}_i[k+1] = \mathbf{A} \sum_{j \in \mathcal{N}_i} w_{ij} \hat{\mathbf{x}}_j[k], \quad (24)$$

where $\hat{\mathbf{x}}_i[k]$ represents an estimate of the state maintained by the i -th node. The weights w_{ij} are non-negative and satisfy

$$\sum_{j \in \mathcal{N}_i} w_{ij} = 1, \quad \forall i \in \mathcal{V} \setminus \mathcal{S}. \quad (25)$$

The design of consensus weights for nodes in $\mathcal{V} \setminus \mathcal{S}$ is based on the observation that the set $\mathcal{V} \setminus \mathcal{S}$ can be spanned by a disjoint union of trees rooted in \mathcal{S} . By assigning consensus weights to only the branches of these trees (without violating the stochasticity condition imposed by equation (25)), one obtains stable estimation error dynamics for each node in $\mathcal{V} \setminus \mathcal{S}$ (the details are similar to the proof of Theorem 1).

Given the above strategy, the following result immediately holds.

Theorem 2: Consider a system (\mathbf{A}, \mathbf{C}) and a graph \mathcal{G} . Let each node in \mathcal{S} run an observer of the form (22), and each node in $\mathcal{V} \setminus \mathcal{S}$ run the consensus dynamics given by (24). Then, there exists a choice of consensus weights and observer gain matrices such that the proposed distributed observer achieves omniscience *if and only if* the sub-system associated with every source component is detectable, i.e., the pair $(\mathbf{A}, \mathbf{C}_{\mathcal{V}_i})$ is detectable $\forall i \in \{1, \dots, p\}$. \square

Remark 4: As in the proof of necessity of Theorem 1, it can be argued that the detectability of subsystems associated with each source component of a given directed network is in fact necessary for **any** estimation scheme to achieve omniscience. \square

V. CONCLUSION

In this paper, we proposed a novel method for constructing distributed observers for linear dynamical systems. Our approach led to a class of observers that not only achieve asymptotic state reconstruction for the most general class of system dynamics and network topologies, but also enjoy a variety of appealing features. Extensions of our approach to cases with faulty or adversarial nodes would be of interest; an initial exploration along these lines is provided in [25].

REFERENCES

- [1] D. Acemoglu, A. Nedić, and A. Ozdaglar, "Convergence of rule-of-thumb learning rules in social networks," in *Proc. of the 47th IEEE Conference on Decision and Control*, 2008, pp. 1714–1720.
- [2] R. Carli, A. Chiuso, L. Schenato, and S. Zampieri, "Distributed Kalman filtering based on consensus strategies," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 4, pp. 622–633, 2008.
- [3] A. Speranzon, C. Fischione, and K. H. Johansson, "Distributed and collaborative estimation over wireless sensor networks," in *Proc. of the 45th IEEE Conference on Decision and Control*, 2006, pp. 1025–1030.
- [4] J. L. Speyer, "Computation and transmission requirements for a decentralized linear-quadratic-Gaussian control problem," *IEEE Transactions on Automatic Control*, vol. 24, no. 2, pp. 266–269, 1979.
- [5] B. Rao, H. Durrant-Whyte, and J. Sheen, "A fully decentralized multi-sensor system for tracking and surveillance," *The International Journal of Robotics Research*, vol. 12, no. 1, pp. 20–44, 1993.
- [6] R. Olfati-Saber, "Distributed Kalman filter with embedded consensus filters," in *Proc. of the 44th IEEE Conference on Decision and Control and European Control Conference*, 2005, pp. 8179–8184.
- [7] —, "Distributed Kalman filtering for sensor networks," in *Proc. of the 46th IEEE Conference on Decision and Control*, 2007, pp. 5492–5498.
- [8] —, "Kalman-consensus filter: Optimality, stability, and performance," in *Proc. of the 48th IEEE Conference on Decision and Control held jointly with the 28th Chinese Control Conference*, 2009, pp. 7036–7042.
- [9] M. Kamgarpour and C. Tomlin, "Convergence properties of a decentralized Kalman filter," in *Proc. of the 47th IEEE Conference on Decision and Control*, 2008, pp. 3205–3210.
- [10] U. Khan and J. M. Moura, "Distributing the Kalman filter for large-scale systems," *IEEE Transactions on Signal Processing*, vol. 56, no. 10, pp. 4919–4935, 2008.
- [11] E. J. Msechu, S. Roumeliotis, A. Ribeiro, and G. B. Giannakis, "Decentralized quantized Kalman filtering with scalable communication cost," *IEEE Transactions on Signal Processing*, vol. 56, no. 8, pp. 3727–3741, 2008.
- [12] U. A. Khan and A. Jadbabaie, "On the stability and optimality of distributed Kalman filters with finite-time data fusion," in *Proc. of the American Control Conference*, 2011, pp. 3405–3410.
- [13] D. W. Casbeer and R. Beard, "Distributed information filtering using consensus filters," in *Proc. of the American Control Conference*, 2009, pp. 1882–1887.
- [14] I. Matei and J. S. Baras, "Consensus-based linear distributed filtering," *Automatica*, vol. 48, no. 8, pp. 1776–1782, 2012.
- [15] U. Khan, S. Kar, A. Jadbabaie, and J. M. Moura, "On connectivity, observability, and stability in distributed estimation," in *Proc. of the 49th IEEE Conference on Decision and Control*, 2010, pp. 6639–6644.
- [16] U. A. Khan and A. Jadbabaie, "Collaborative scalar-gain estimators for potentially unstable social dynamics with limited communication," *Automatica*, vol. 50, no. 7, pp. 1909–1914, 2014.
- [17] V. Ugrinovskii, "Conditions for detectability in distributed consensus-based observer networks," *arXiv preprint arXiv:1303.6397*, 2013.
- [18] S. Park and N. C. Martins, "An augmented observer for the distributed estimation problem for LTI systems," in *Proc. of the American Control Conference*, 2012, pp. 6775–6780.
- [19] —, "Necessary and sufficient conditions for the stabilizability of a class of LTI distributed observers," in *Proc. of the 51st IEEE Conference on Decision and Control*, 2012, pp. 7431–7436.
- [20] —, "Design of distributed LTI observers for state omniscience," *IEEE Transactions on Automatic Control*, to appear.
- [21] B. D. Anderson and D. J. Clements, "Algebraic characterization of fixed modes in decentralized control," *Automatica*, vol. 17, no. 5, pp. 703–712, 1981.
- [22] A. Mitra and S. Sundaram, "Distributed observers for LTI systems," *arXiv preprint arXiv:1608.01429*, 2016.
- [23] J. A. Bondy and U. S. R. Murty, *Graph Theory with Applications*. Citeseer, 1976, vol. 290.
- [24] M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*. Courier Corporation, 1992, vol. 14.
- [25] A. Mitra and S. Sundaram, "Secure distributed observers for a class of linear time invariant systems in the presence of byzantine adversaries," in *Proc. of the 55th IEEE Conference on Decision and Control*, 2016, to appear.