

Finite-Time Distributed State Estimation over Time-Varying Graphs: Exploiting the Age-of-Information

Aritra Mitra, John A. Richards, Saurabh Bagchi and Shreyas Sundaram

Abstract—We study the problem of collaboratively estimating the state of a discrete-time LTI process by a network of sensor nodes interacting over a time-varying directed communication graph. Existing approaches to this problem either (i) make restrictive assumptions on the dynamical model, or (ii) make restrictive assumptions on the sequence of communication graphs, or (iii) require multiple consensus iterations between consecutive time-steps of the dynamics, or (iv) require higher-dimensional observers. In this paper, we develop a distributed observer that operates on a single time-scale, is of the same dimension as that of the state, and works under mild assumptions of joint observability of the sensing model, and joint strong-connectivity of the sequence of communication graphs. Our approach is based on the notion of a novel “freshness-index” that keeps track of the age-of-information being diffused across the network. In particular, such indices enable nodes to reject stale information regarding the state of the system, and in turn, help achieve stability of the estimation error dynamics. Based on the proposed approach, the estimate of each node can be made to converge to the true state exponentially fast, at any desired convergence rate. In fact, we argue that finite-time convergence can also be achieved through a suitable selection of the observer gains. Our proof of convergence is self-contained, and employs simple arguments from linear system theory and graph theory.

I. INTRODUCTION

Given a discrete-time LTI system $\mathbf{x}[k+1] = \mathbf{A}\mathbf{x}[k]$, and a linear measurement model $\mathbf{y}[k] = \mathbf{C}\mathbf{x}[k]$, a classical result in control theory states that if the pair (\mathbf{A}, \mathbf{C}) is observable, then one can generate an estimate $\hat{\mathbf{x}}[k]$ that converges exponentially fast to $\mathbf{x}[k]$ at any desired convergence rate. Over the last couple of decades, significant effort has been directed towards studying the distributed counterpart of the above problem, wherein observations of the process are distributed among a set of sensors modeled as nodes of a communication graph [1]–[10]. The fundamental problem of identifying the minimal requirements on the measurement structure of the nodes and the underlying communication graph that guarantee the existence of a distributed observer, was solved only recently in [5]–[9] for static graphs.

A. Mitra, S. Bagchi, and S. Sundaram are with the School of Electrical and Computer Engineering at Purdue University. J. A. Richards is with Sandia National Laboratories. Email: {mitra14, sbagchi, sundara2}@purdue.edu, jaricha@sandia.gov. This work was supported in part by NSF CAREER award 1653648, and by a grant from Sandia National Laboratories. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA0003525. The views expressed in the article do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

The approaches in [5]–[10] can be generally classified based on the following attributes. (i) Does the approach require multiple consensus iterations between consecutive time-steps of the dynamics?¹ (ii) What is the dimension of the estimator maintained by each node? (iii) Can the convergence rate be controlled? (iv) Is the approach robust to temporal variations in the underlying communication graph? The techniques proposed in [5]–[9] operate on a single-time-scale, those in [7]–[9] require observers of dimension no more than that of the state of the system, the ones in [6], [7], [10] can achieve any desired convergence rate, while the one in [10] can account for a fairly general class of time-varying graphs. The main **contribution** of this paper is the development of a distributed observer that shares each of the above positive attributes. Specifically, we develop a single-time-scale distributed state estimation algorithm in Section IV that requires each node to maintain an estimator of dimension equal to that of the state (along with some simple counters), and works under the basic assumptions of joint observability of the observation model, and joint strong-connectivity of the sequence of communication graphs.

The authors in [11] point out that even for the basic consensus problem, extending the stability analysis for static graphs to time-varying graphs is highly non-trivial. Arguably, the stability analysis for the distributed state estimation problem with time-varying graphs is even more challenging, since one has to account for potentially unstable external dynamics, a feature that is missing in the standard consensus problem. Nevertheless, in Section V, we establish using simple arguments from linear system theory and graph theory, that based on our approach, each node can track the true dynamics exponentially fast at any desired convergence rate. Additionally, we show how to design the observer gains so as to achieve convergence in finite time. The closest related work is [10], where the authors develop a solution that makes an elegant connection to the problem of distributed linear-equation solving [12]. In contrast to our technique, the one in [10] is inherently a two-time-scale approach, requires each node to maintain and update auxiliary state estimates, and works under the assumption that the communication graph is strongly-connected at each time-step.

The key idea behind our algorithm is the use of a suitably defined “freshness-index” that keeps track of the age-of-information being diffused across the network. Loosely speaking, such indices are a measure of the accuracy with

¹Such approaches, referred to as two-time-scale approaches, may prove to be computationally prohibitive for real-time applications.

which the information received by a node describes the physical process being observed. While the freshness-indices enable a node to reject stale information, the assumption of joint strong-connectivity ensures that fresh information is diffused across the network sufficiently often. These facts taken together help achieve stability of the estimation error process. Finally, we point out that while this is perhaps the first use of the notion of age-of-information (AoI) in a networked control/estimation setting, such a concept has been widely employed in the study of various queueing-theoretic problems arising in wireless networks [13], [14].²

II. PROBLEM FORMULATION AND BACKGROUND

We are interested in collaborative state estimation of a discrete-time LTI system of the form:

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

where $k \in \mathbb{N}$ is the discrete-time index, $\mathbf{A} \in \mathbb{R}^{n \times n}$ is the system matrix, and $\mathbf{x}[k] \in \mathbb{R}^n$ is the state of the system.³ A network of sensors, modeled as nodes of a communication graph, obtain partial measurements of the state of the above process as follows:

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

where $\mathbf{y}_i[k] \in \mathbb{R}^{r_i}$ represents the measurement vector of the i -th node at time-step k , and $\mathbf{C}_i \in \mathbb{R}^{r_i \times n}$ represents the corresponding observation matrix. Let $\mathbf{y}[k] = [\mathbf{y}_1^T[k] \ \dots \ \mathbf{y}_N^T[k]]^T$ and $\mathbf{C} = [\mathbf{C}_1^T \ \dots \ \mathbf{C}_N^T]^T$ represent the collective measurement vector at time-step k , and the collective observation matrix, respectively. The goal of each node i in the network is to generate an asymptotically correct estimate $\hat{\mathbf{x}}_i[k]$ of the true dynamics $\mathbf{x}[k]$. It may not be possible for any node i in the network to accomplish such a task in isolation, since the pair $(\mathbf{A}, \mathbf{C}_i)$ may not be detectable in general. Throughout the paper, we will only assume that the pair (\mathbf{A}, \mathbf{C}) is observable.

As is evident from the above discussion, information exchange among nodes is necessary to solve the problem at hand. At each time-step $k \in \mathbb{N}$, such interactions are modeled by a directed communication graph $\mathcal{G}[k] = (\mathcal{V}, \mathcal{E}[k])$, where $\mathcal{V} = \{1, \dots, N\}$ represents the set of nodes, and $\mathcal{E}[k]$ represents the edge set of $\mathcal{G}[k]$ at time-step k . Specifically, if $(i, j) \in \mathcal{E}[k]$, then node i can send information directly to node j at time-step k ; in such a case, node i will be called a neighbor of node j at time-step k . We will use $\mathcal{N}_i[k]$ to represent the set of all neighbors (excluding node i) of node i at time-step k . When $\mathcal{G}[k] = \mathcal{G} \ \forall k \in \mathbb{N}$, where \mathcal{G} is a static, directed communication graph, the necessary and sufficient condition (on the system and network) to solve the distributed state estimation problem is that each source

²The notion of age-of-information (AoI) was first introduced in [13] as a performance metric to keep track of real-time status updates in a communication system. In a wireless network, it measures the time elapsed since the generation of the packet most recently delivered to the destination. In Section IV, we will relate this concept to our present setting.

³We use \mathbb{N} and \mathbb{N}_+ to denote the set of non-negative integers and the set of positive integers, respectively.

component of \mathcal{G} be collectively detectable [5]. Our goal in this paper is to extend the above result to the scenario where the underlying communication graph is allowed to change over time. To this end, let the union graph over an interval $[k_1, k_2]$, $0 \leq k_1 < k_2$, indicate a graph with vertex set equal to \mathcal{V} , and edge set equal to the union of the edge sets of the individual graphs appearing over the interval $[k_1, k_2]$. Based on this convention, we will assume that the sequence of communication graphs $\{\mathcal{G}[k]\}_{k=0}^\infty$ is “jointly strongly-connected”, in the sense described below.

Assumption 1. (Joint Strong-Connectivity) *There exists $T \in \mathbb{N}_+$ such that the union graph over every interval of the form $[kT, (k+1)T]$ is strongly-connected, where $k \in \mathbb{N}$.*

For communication graphs satisfying the above assumption, our **objective** will be to design a distributed algorithm that ensures $\lim_{k \rightarrow \infty} \|\hat{\mathbf{x}}_i[k] - \mathbf{x}[k]\| = 0, \forall i \in \mathcal{V}$. To this end, we recall the following result from [8].

Lemma 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 $\mathbf{C} \triangleq [\mathbf{C}_1^T \ \dots \ \mathbf{C}_N^T]^T$. Suppose (\mathbf{A}, \mathbf{C}) is observable. Then, there exists a similarity transformation matrix \mathbf{T} that transforms the pair (\mathbf{A}, \mathbf{C}) to $(\bar{\mathbf{A}}, \bar{\mathbf{C}})$, such that*

$$\begin{aligned} \bar{\mathbf{A}} &= \left[\begin{array}{c|c|c} \mathbf{A}_{11} & \mathbf{0} & \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{0} \\ \vdots & \vdots & \ddots \\ \hline \mathbf{A}_{N1} & \mathbf{A}_{N2} & \dots & \mathbf{A}_{N(N-1)} & \mathbf{A}_{NN} \end{array} \right], \\ \bar{\mathbf{C}} &= \left[\begin{array}{c} \bar{\mathbf{C}}_1 \\ \bar{\mathbf{C}}_2 \\ \vdots \\ \bar{\mathbf{C}}_N \end{array} \right] = \left[\begin{array}{c|c|c} \mathbf{C}_{11} & \mathbf{0} & \\ \hline \mathbf{C}_{21} & \mathbf{C}_{22} & \mathbf{0} \\ \vdots & \vdots & \vdots \\ \hline \mathbf{C}_{N1} & \mathbf{C}_{N2} & \dots & \mathbf{C}_{N(N-1)} & \mathbf{C}_{NN} \end{array} \right], \end{aligned} \quad (3)$$

and the pair $(\bar{\mathbf{A}}_{ii}, \bar{\mathbf{C}}_{ii})$ is observable $\forall i \in \{1, 2, \dots, N\}$. \square

We use the matrix \mathbf{T} given by Lemma 1 to perform the coordinate transformation $\mathbf{x}[k] = \mathbf{T}\mathbf{z}[k]$, yielding:

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

where $\bar{\mathbf{A}} = \mathbf{T}^{-1}\mathbf{A}\mathbf{T}$ and $\bar{\mathbf{C}}_i = \mathbf{C}_i\mathbf{T}$ are given by (3). Commensurate with the structure of $\bar{\mathbf{A}}$, the vector $\mathbf{z}[k]$ is of the following form:

$$\mathbf{z}[k] = \left[\mathbf{z}^{(1)}[k]^T \ \dots \ \mathbf{z}^{(N)}[k]^T \right]^T, \quad (5)$$

where $\mathbf{z}^{(j)}[k]$ will be referred to as the j -th substate. By construction, since the pair $(\mathbf{A}_{jj}, \mathbf{C}_{jj})$ is locally observable w.r.t. the measurements of node j , node j will be viewed as the unique source node for substate j . In this sense, the role of node j will be to ensure that each non-source node $i \in \mathcal{V} \setminus \{j\}$ maintains an asymptotically correct estimate of substate j . For a time-invariant strongly-connected graph, this is achieved in [8] by first constructing a spanning tree rooted at node j , and then requiring nodes to only listen to their parents in such a tree for estimating substate j . The unidirectional flow of information (from the source node j to the rest of the network) so achieved guarantees stability of the error process. The above strategy is no longer applicable

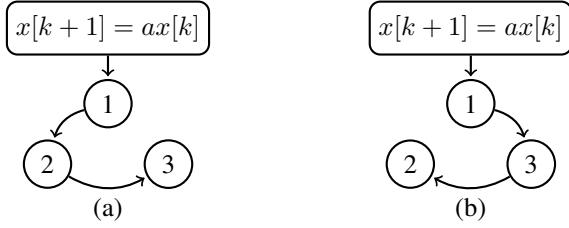


Fig. 1. An LTI system is monitored by a network of 3 nodes, where the communication graph $\mathcal{G}[k]$ switches between the two graphs shown above.

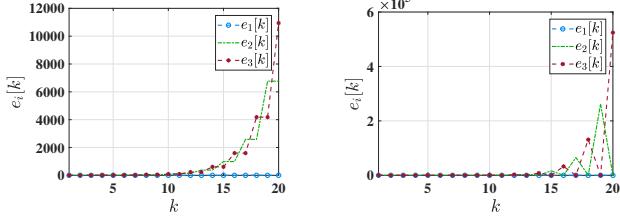


Fig. 2. Estimation error plots of the nodes for the model in Figure 1. Simulations are performed for a model where $a = 2$. The figure on the left corresponds to the case where consensus weights are distributed uniformly among neighbors, while the one on the right is the case where weights are placed along a tree rooted at node 1.

when the underlying communication graph is time-varying, for the following reasons. (i) For a given substate j , there may not exist a common spanning tree rooted at node j in each graph $\mathcal{G}[k]$, $k \in \mathbb{N}$. (ii) Requiring a specific spanning tree rooted at node j to repeat after a finite duration of time is restrictive, and qualifies as only a special case of Assumption 1. (iii) Suppose for simplicity that $\mathcal{G}[k]$ is strongly-connected at each time-step (as in [10]), and hence, there exists a spanning tree $\mathcal{T}_j[k]$ rooted at node j in each such graph. For estimating substate j , suppose consensus at time-step k is performed along the spanning tree $\mathcal{T}_j[k]$; the next section shows that switching between such spanning trees can lead to unstable error processes. In Section IV, we develop a fairly simple algorithm that bypasses each of the above problems. The next section discusses the intuition behind this algorithm.

III. ILLUSTRATIVE EXAMPLE

Consider a network of 3 nodes monitoring a scalar unstable process $x[k+1] = ax[k]$, as shown in Figure 1. The communication graph $\mathcal{G}[k]$ switches between the two topologies shown in Figure 1. Specifically, $\mathcal{G}[k]$ is the graph in Figure 1(a) at all even time-steps, and the one in 1(b) at all odd time-steps. Node 1 is the only node with non-zero measurements, and thus acts as the source node for this network. Suppose for simplicity that it has perfect information of the state at all time-steps, i.e., $\hat{x}_1[k] = x[k], \forall k \in \mathbb{N}$. Given this setup, a standard consensus based state estimate update rule would take the form (see for example [5], [8], [10]):

$$\hat{x}_i[k+1] = a \left(\sum_{j \in \mathcal{N}_i[k] \cup \{i\}} w_{ij}[k] \hat{x}_j[k] \right), i \in \{2, 3\}, \quad (6)$$

where the weights $w_{ij}[k]$ are non-negative, and satisfy $\sum_{j \in \mathcal{N}_i[k] \cup \{i\}} w_{ij}[k] = 1, \forall k \in \mathbb{N}$. The key question is: how should the consensus weights be chosen to guarantee stability of the estimation errors of nodes 2 and 3? Even for this simple example, if such weights are chosen naively, then the errors may grow unbounded over time. To see this, consider the following two choices: (1) consensus weights are distributed evenly over the set $\mathcal{N}_i[k] \cup \{i\}$, and (2) consensus weights are placed along the tree rooted at node 1. In each case, the error dynamics are unstable, as depicted in Figure 2. To overcome this problem, suppose nodes 2 and 3 are aware of the fact that node 1 has perfect information of the state. Since nodes 2 and 3 have no measurements of their own, intuitively, it makes sense that they should place their consensus weights entirely on node 1 whenever possible. The trickier question for node 2 (resp., node 3) is to decide when it should listen to node 3 (resp., node 2). Let us consider the situation from the perspective of node 2. At time-step 0, it adopts the information of node 1, and hence, the error of node 2 is zero at time-step 1. However, the error of node 3 is not necessarily zero at time-step 1. Consequently, if node 2 places a non-zero consensus weight on the estimate of node 3 at time-step 1, its error at time-step 2 might assume a non-zero value. Clearly, at time-step 1, node 2 is better off rejecting the information from node 3, and simply running open-loop. The main take-away point here is that adoption or rejection of information from a neighbor should be based on the quality of information that such a neighbor has to offer. In particular, a node that has come in contact with node 1 more recently is expected to have better information about the state than the other. Thus, to dynamically evaluate the quality of an estimate, the above reasoning suggests the need to introduce a metric that keeps track of how delayed that estimate is w.r.t. the estimate of the source node 1. In the following section, we introduce such a metric.

IV. ALGORITHM

Building on the intuition developed in the previous section, we introduce a new approach to designing distributed observers for a general class of time-varying networks. The main idea is the use of a “freshness-index” that keeps track of how delayed the estimates of a node are w.r.t. the estimates of a source node. Specifically, for updating its estimate of $\mathbf{z}^{(j)}[k]$, each node $i \in \mathcal{V}$ maintains and updates at every time-step a freshness-index $\tau_i^{(j)}[k]$. At each time-step $k \in \mathbb{N}$, the index $\tau_i^{(j)}[k]$ plays the following role: it determines whether node i should adopt the information received from one of its neighbors in $\mathcal{N}_i[k]$, or run open-loop, for updating $\hat{\mathbf{z}}_i^{(j)}[k]$, where $\hat{\mathbf{z}}_i^{(j)}[k]$ represents the estimate of $\mathbf{z}^{(j)}[k]$ maintained by node i . In case it is the former, it also indicates which specific neighbor in $\mathcal{N}_i[k]$ that node i should listen to at time-step k ; this piece of information is particularly important for the problem under consideration, and ensures stability of the error process. We now formally describe the rules that govern the update of the estimates of the j -th substate $\mathbf{z}^{(j)}[k]$.

Initialization of Freshness-Indices: Each node $i \in \mathcal{V}$ maintains an index $\tau_i^{(j)}[k] \in \{\omega\} \cup \mathbb{N}$, where ω is a dummy

value. Specifically, $\tau_i^{(j)}[k] = \omega$ represents an “infinite-delay” w.r.t. the estimate of the source node for sub-state j , namely node j (i.e., it represents that node i has not received any information from node j regarding substate j up to time-step k). The indices $\tau_i^{(j)}[k]$ are initialized as: $\tau_j^{(j)}[0] = 0, \tau_i^{(j)}[0] = \omega, \forall i \in \mathcal{V} \setminus \{j\}$.

Update Rules for the Source Node: Node j maintains $\tau_j^{(j)}[k] = 0, \forall k \in \mathbb{N}$, and updates $\hat{\mathbf{z}}_j^{(j)}[k]$ as:

$$\hat{\mathbf{z}}_j^{(j)}[k+1] = (\mathbf{A}_{jj} - \mathbf{L}_j \mathbf{C}_{jj}) \hat{\mathbf{z}}_j^{(j)}[k] + \sum_{q=1}^{(j-1)} (\mathbf{A}_{jq} - \mathbf{L}_j \mathbf{C}_{jq}) \hat{\mathbf{z}}_j^{(q)}[k] + \mathbf{L}_j \mathbf{y}_j[k], \quad (7)$$

where \mathbf{L}_j is an output-injection gain to be decided later.

Update Rules for the Non-Source Nodes: For each non-source node $i \in \mathcal{V} \setminus \{j\}$, we consider two distinct cases based on the value of $\tau_i^{(j)}[k]$.

Case 1: $\tau_i^{(j)}[k] = \omega$. Define

$$\mathcal{M}_i^{(j)}[k] \triangleq \{l \in \mathcal{N}_i[k] : \tau_l^{(j)}[k] \neq \omega\}. \quad (8)$$

If $\mathcal{M}_i^{(j)}[k] \neq \emptyset$, let $u \in \operatorname{argmin}_{l \in \mathcal{M}_i^{(j)}[k]} \tau_l^{(j)}[k]$. Node i then updates $\tau_i^{(j)}[k]$ and $\hat{\mathbf{z}}_i^{(j)}[k]$ as follows:

$$\tau_i^{(j)}[k+1] = \tau_u^{(j)}[k] + 1, \quad (9)$$

$$\hat{\mathbf{z}}_i^{(j)}[k+1] = \mathbf{A}_{jj} \hat{\mathbf{z}}_u^{(j)}[k] + \sum_{q=1}^{(j-1)} \mathbf{A}_{jq} \hat{\mathbf{z}}_i^{(q)}[k]. \quad (10)$$

If $\mathcal{M}_i^{(j)}[k] = \emptyset$, then

$$\tau_i^{(j)}[k+1] = \omega, \quad (11)$$

$$\hat{\mathbf{z}}_i^{(j)}[k+1] = \mathbf{A}_{jj} \hat{\mathbf{z}}_i^{(j)}[k] + \sum_{q=1}^{(j-1)} \mathbf{A}_{jq} \hat{\mathbf{z}}_i^{(q)}[k]. \quad (12)$$

Case 2: $\tau_i^{(j)}[k] \neq \omega$. Define

$$\mathcal{F}_i^{(j)}[k] \triangleq \{l \in \mathcal{M}_i^{(j)}[k] : \tau_l^{(j)}[k] < \tau_i^{(j)}[k]\}, \quad (13)$$

where $\mathcal{M}_i^{(j)}[k]$ is as defined in (8). If $\mathcal{F}_i^{(j)}[k] \neq \emptyset$, let $u \in \operatorname{argmin}_{l \in \mathcal{F}_i^{(j)}[k]} \tau_l^{(j)}[k]$. Node i then updates $\tau_i^{(j)}[k]$ as per (9), and $\hat{\mathbf{z}}_i^{(j)}[k]$ as per (10). If $\mathcal{F}_i^{(j)}[k] = \emptyset$, then $\tau_i^{(j)}[k]$ is updated as

$$\tau_i^{(j)}[k+1] = \tau_i^{(j)}[k] + 1, \quad (14)$$

and $\hat{\mathbf{z}}_i^{(j)}[k]$ is updated as per (12).

The above steps describe an approach for estimating $\mathbf{z}[k]$, and hence $\mathbf{x}[k]$, since $\mathbf{x}[k] = \mathbf{T}\mathbf{z}[k]$. We now briefly describe each rule of the proposed algorithm. Consider any substate $j \in \{1, \dots, N\}$. For estimation of substate j , since delays are measured w.r.t. the source node j , node j maintains its freshness-index $\tau_j^{(j)}[k]$ at zero for all time, to indicate a zero delay w.r.t. itself. For updating its estimate of $\mathbf{z}^{(j)}[k]$, it uses only its own information, as is evident from (7). Every other node starts out with an “infinite-delay” w.r.t. the source, which is represented by the freshness-index taking on the value ω . The freshness-index of a node $i \in \mathcal{V} \setminus \{j\}$ changes from ω to a finite value when it comes in contact with a

neighbor with a finite delay, i.e., with a freshness-index that is not ω (see equation (8)). At this point, we say that $\tau_i^{(j)}[k]$ has been “triggered”. Once triggered, at each time-step k , a non-source node i will adopt the information of a neighbor $l \in \mathcal{N}_i[k]$ only if node l ’s estimate of $\mathbf{z}^{(j)}[k]$ is “more fresh” relative to its own, i.e., only if $\tau_l^{(j)}[k] < \tau_i^{(j)}[k]$ (see equation (13)). Among the set of neighbors in $\mathcal{M}_i^{(j)}[k]$ (if $\tau_i^{(j)}[k]$ has not yet been triggered), or in $\mathcal{F}_i^{(j)}[k]$ (if $\tau_i^{(j)}[k]$ has been triggered), node i only adopts the information (based on (10)) of the neighbor u with the least delay. At this point, the delay of node i matches that of node u , and this fact is captured by the update rule (9). In case node i has no informative neighbor, it increments its own freshness-index linearly by 1 (to capture the effect of its own information getting older) via the update rule (14), and runs open-loop based on (12).

V. MAIN RESULT AND ANALYSIS

The main result of the paper is as follows.

Theorem 1. *Given an LTI system (1), and a measurement model (2), suppose (\mathbf{A}, \mathbf{C}) is observable. Let the sequence of communication graphs $\{\mathcal{G}[k]\}_{k=0}^{\infty}$ satisfy Assumption 1. Then, the observer gains $\mathbf{L}_1, \dots, \mathbf{L}_N$ can be designed in a manner such that the estimation error of each node $i \in \mathcal{V}$ converges to zero exponentially fast at any desired convergence rate ρ , based on the proposed algorithm. \square*

To prove Theorem 1, we require the following intermediate results; proofs of these results are omitted here due to space constraints, but can be found in [15].

Lemma 2. *Consider any substate j , and suppose that at some time-step k , we have that $j \in \mathcal{M}_i^{(j)}[k]$, for some $i \in \mathcal{V} \setminus \{j\}$. Then, the proposed algorithm guarantees the following.*

- (i) *If $\tau_i^{(j)}[k] = \omega$, then $j = \operatorname{argmin}_{l \in \mathcal{M}_i^{(j)}[k]} \tau_l^{(j)}[k]$.*
- (ii) *If $\tau_i^{(j)}[k] \neq \omega$, then $j \in \mathcal{F}_i^{(j)}[k]$, and $j = \operatorname{argmin}_{l \in \mathcal{F}_i^{(j)}[k]} \tau_l^{(j)}[k]$.*

\square

Lemma 3. *Suppose Assumption 1 is met. Then, for each substate j , the proposed algorithm guarantees the following.*

$$\tau_i^{(j)}[k] \neq \omega, \forall k \geq (N-1)T, \forall i \in \mathcal{V}. \quad (15)$$

\square

Lemma 4. *Consider any substate j , and suppose that at some time-step k , we have $\tau_i^{(j)}[k] = m$, where $i \in \mathcal{V} \setminus \{j\}$, and $m \in \mathbb{N}_+$. Then, the proposed algorithm guarantees:*

$$\hat{\mathbf{z}}_i^{(j)}[k] = \mathbf{A}_{jj}^m \hat{\mathbf{z}}_j^{(j)}[k-m] + \sum_{q=1}^{(j-1)} \sum_{\tau=(k-m)}^{(k-1)} \mathbf{A}_{jj}^{(k-\tau-1)} \mathbf{A}_{jq} \hat{\mathbf{z}}_{v(\tau)}^{(q)}[\tau], \quad (16)$$

where $v(\tau) \in \mathcal{V} \setminus \{j\}$. \square

Lemma 5. *Suppose Assumption 1 is met. Then, for any substate j , the proposed algorithm guarantees the following:*

$$\tau_i^{(j)}[k] \leq 2(N-1)T, \forall k \geq (N-1)T, \forall i \in \mathcal{V}. \quad (17)$$

\square

Proof. (Theorem 1) Given a desired rate of convergence $\rho \in (0, 1)$, pick a set of positive scalars $\{\rho_1, \dots, \rho_N\}$, such that $\rho_1 < \rho_2 < \dots < \rho_N < \rho$. For each substate $j \in \{1, \dots, N\}$, let the corresponding source node j design the observer gain \mathbf{L}_j (featuring in equation (7)) in a manner such that the matrix $(\mathbf{A}_{jj} - \mathbf{L}_j \mathbf{C}_{jj})$ has distinct real eigenvalues with spectral radius equal to ρ_j . Such a choice of \mathbf{L}_j exists as the pair $(\mathbf{A}_{jj}, \mathbf{C}_{jj})$ is observable by construction. Then, there exists a set of positive scalars $\{\alpha_1, \dots, \alpha_N\}$, such that [16]:⁴

$$\|(\mathbf{A}_{jj} - \mathbf{L}_j \mathbf{C}_{jj})^k\| \leq \alpha_j \rho_j^k, \forall k \in \mathbb{N}. \quad (18)$$

For a particular substate j , let $\mathbf{e}_j^{(j)}[k] = \hat{\mathbf{z}}_i^{(j)}[k] - \mathbf{z}^{(j)}[k]$. Consider the first substate $j = 1$, and observe that based on (3), (4), and (7), the following is true: $\mathbf{e}_1^{(1)}[k+1] = (\mathbf{A}_{11} - \mathbf{L}_1 \mathbf{C}_{11}) \mathbf{e}_1^{(1)}[k]$. Thus, we obtain

$$\mathbf{e}_1^{(1)}[k] = (\mathbf{A}_{11} - \mathbf{L}_1 \mathbf{C}_{11})^k \mathbf{e}_1^{(1)}[0]. \quad (19)$$

Taking norms on both sides of (19), and using (18), yields:

$$\|\mathbf{e}_1^{(1)}[k]\| \leq c_1 \rho_1^k, \forall k \in \mathbb{N}, \quad (20)$$

where $c_1 \triangleq \alpha_1 \|\mathbf{e}_1^{(1)}[0]\|$. Based on Lemmas 3 and 4, and the fact that $\mathbf{z}^{(1)}[k] = (\mathbf{A}_{11})^m \mathbf{z}^{(1)}[k-m], \forall m \in \mathbb{N}$, the following is true for any non-source node $i \in \mathcal{V} \setminus \{1\}$:

$$\mathbf{e}_i^{(1)}[k] = (\mathbf{A}_{11})^{\tau_i^{(1)}[k]} \mathbf{e}_1^{(1)}[k - \tau_i^{(1)}[k]], \forall k \geq (N-1)T. \quad (21)$$

For each substate j , one can always find scalars $\beta_j, \gamma_j \geq 1$, such that $\|(\mathbf{A}_{jj})^k\| \leq \beta_j \gamma_j^k, \forall k \in \mathbb{N}$ [16]. Using this bound and the one in (20), the fact that $\gamma_1 \geq 1$ and $\rho_1 < 1$, the fact that $\tau_i^{(1)}[k] \leq 2(N-1)T, \forall k \geq (N-1)T$ based on Lemma 5, and the sub-multiplicative property of the 2-norm, we obtain the following by taking norms on both sides of (21):

$$\|\mathbf{e}_i^{(1)}[k]\| \leq \bar{c}_1 \rho_1^k, \forall k \geq (N-1)T, \forall i \in \mathcal{V}, \quad (22)$$

where

$$\bar{c}_1 \triangleq c_1 \beta_1 \left(\frac{\gamma_1}{\rho_1} \right)^{2\bar{T}}, \quad \bar{T} = (N-1)T. \quad (23)$$

Note that $\bar{c}_1 \geq c_1$, and hence the bound in (22) applies to node 1 as well (see equation (20)). Our goal is to now obtain a bound similar to that in (22) for each substate $j \in \{2, \dots, N\}$. To this end, let $g_{jq} = \|(\mathbf{A}_{jq} - \mathbf{L}_j \mathbf{C}_{jq})\|$, and $h_{jq} = \|\mathbf{A}_{jq}\|$. Define the following quantities recursively for $j \in \{2, \dots, N\}$.

$$\begin{aligned} c_j &\triangleq \frac{\alpha_j}{\rho_j^{(2j-3)\bar{T}}} \left(\|\mathbf{e}_j^{(j)}[(2j-3)\bar{T}]\| + \sum_{q=1}^{(j-1)} \frac{g_{jq} \bar{c}_q}{(\rho_j - \rho_q)} \rho_q^{(2j-3)\bar{T}} \right), \\ \bar{c}_j &\triangleq \beta_j \left(c_j \left(\frac{\gamma_j}{\rho_j} \right)^{2\bar{T}} + \sum_{q=1}^{(j-1)} \frac{h_{jq} \bar{c}_q}{(\gamma_j - \rho_q)} \left(\frac{\gamma_j}{\rho_q} \right)^{2\bar{T}} \right), \end{aligned} \quad (24)$$

where $c_1 \triangleq \alpha_1 \|\mathbf{e}_1^{(1)}[0]\|$, and \bar{c}_1 is as defined in (23). We now claim the following for each substate $j \in \{1, \dots, N\}$:

$$\|\mathbf{e}_i^{(j)}[k]\| \leq \bar{c}_j \rho_j^k, \forall k \geq (2j-1)\bar{T}, \forall i \in \mathcal{V}. \quad (25)$$

⁴We use $\|\mathbf{A}\|$ to refer to the induced 2-norm of a matrix \mathbf{A} .

To prove the above claim, we proceed via induction on the substate number j . Suppose the claim holds for all $q \in \{1, \dots, j-1\}$, where $2 \leq j \leq N$. To prove the desired result for substate j , observe that equations (3) and (4) yield:

$$\begin{aligned} \mathbf{z}^{(j)}[k+1] &= \mathbf{A}_{jj} \mathbf{z}^{(j)}[k] + \sum_{q=1}^{(j-1)} \mathbf{A}_{jq} \mathbf{z}^{(q)}[k] \\ &= (\mathbf{A}_{jj} - \mathbf{L}_j \mathbf{C}_{jj}) \mathbf{z}^{(j)}[k] + \sum_{q=1}^{(j-1)} (\mathbf{A}_{jq} - \mathbf{L}_j \mathbf{C}_{jq}) \mathbf{z}^{(q)}[k] + \mathbf{L}_j \mathbf{y}_j[k]. \end{aligned} \quad (26)$$

Based on the above equation and (7), we obtain:

$$\mathbf{e}_j^{(j)}[k+1] = (\mathbf{A}_{jj} - \mathbf{L}_j \mathbf{C}_{jj}) \mathbf{e}_j^{(j)}[k] + \sum_{q=1}^{(j-1)} (\mathbf{A}_{jq} - \mathbf{L}_j \mathbf{C}_{jq}) \mathbf{e}_j^{(q)}[k]. \quad (27)$$

Rolling out the above equation over time, we obtain:

$$\begin{aligned} \mathbf{e}_j^{(j)}[k] &= (\mathbf{A}_{jj} - \mathbf{L}_j \mathbf{C}_{jj})^{(k-(2(j-1)-1)\bar{T})} \mathbf{e}_j^{(j)}[(2(j-1)-1)\bar{T}] \\ &\quad + \sum_{q=1}^{(j-1)} \sum_{\tau=(2(j-1)-1)\bar{T}}^{k-1} (\mathbf{A}_{jj} - \mathbf{L}_j \mathbf{C}_{jj})^{(k-\tau-1)} (\mathbf{A}_{jq} - \mathbf{L}_j \mathbf{C}_{jq}) \mathbf{e}_j^{(q)}[\tau], \end{aligned} \quad (28)$$

where $k \geq (2(j-1)-1)\bar{T}$. Taking norms on both sides of the above equation, using the triangle inequality, and the sub-multiplicative property of the two-norm, we obtain:

$$\begin{aligned} \|\mathbf{e}_j^{(j)}[k]\| &\stackrel{(a)}{\leq} \alpha_j \rho_j^k \left(\frac{\|\mathbf{e}_j^{(j)}[(2j-3)\bar{T}]\|}{\rho_j^{(2j-3)\bar{T}}} + \frac{1}{\rho_j} \sum_{q=1}^{(j-1)} g_{jq} \sum_{\tau=(2j-3)\bar{T}}^{(k-1)} \rho_j^{-\tau} \|\mathbf{e}_j^{(q)}[\tau]\| \right) \\ &\stackrel{(b)}{\leq} \alpha_j \rho_j^k \left(\frac{\|\mathbf{e}_j^{(j)}[(2j-3)\bar{T}]\|}{\rho_j^{(2j-3)\bar{T}}} + \frac{1}{\rho_j} \sum_{q=1}^{(j-1)} g_{jq} \bar{c}_q \sum_{\tau=(2j-3)\bar{T}}^{(k-1)} \left(\frac{\rho_q}{\rho_j} \right)^\tau \right) \\ &\stackrel{(c)}{\leq} c_j \rho_j^k, \forall k \geq (2j-3)\bar{T}. \end{aligned} \quad (29)$$

In the above inequalities, (a) follows from (18) and the definition of g_{jq} , (b) follows by noting that $q \leq (j-1)$, $\tau \geq (2(j-1)-1)\bar{T}$, and then applying the induction hypothesis, (c) follows by simplifying the preceding inequality using $\rho_q < \rho_j$. We have thus obtained a bound on the estimation error of substate j for node j . To bound the estimation errors of substate j for each non-source node $i \in \mathcal{V} \setminus \{j\}$, note that equation (26) can be rolled out over time to yield:

$$\mathbf{z}^{(j)}[k] = \mathbf{A}_{jj}^m \mathbf{z}^{(j)}[k-m] + \sum_{q=1}^{(j-1)} \sum_{\tau=(k-m)}^{(k-1)} \mathbf{A}_{jj}^{(k-\tau-1)} \mathbf{A}_{jq} \mathbf{z}^{(q)}[\tau]. \quad (30)$$

Lemma 4 then yields for each $i \in \mathcal{V} \setminus \{j\}$, $\forall k \geq \bar{T}$:

$$\begin{aligned} \mathbf{e}_i^{(j)}[k] &= (\mathbf{A}_{jj})^{\tau_i^{(j)}[k]} \mathbf{e}_j^{(j)}[k - \tau_i^{(j)}[k]] \\ &\quad + \sum_{q=1}^{(j-1)} \sum_{\tau=(k-\tau_i^{(j)}[k])}^{(k-1)} \mathbf{A}_{jj}^{(k-\tau-1)} \mathbf{A}_{jq} \mathbf{e}_v^{(q)}[\tau]. \end{aligned} \quad (31)$$

Based on the induction hypothesis and equation (29), notice that each error term featuring in the RHS of the above equation will decay exponentially provided k is large enough. Specifically, suppose $k \geq (2j-1)\bar{T}$, in which case $k - \tau_i^{(j)}[k] \geq (2j-3)\bar{T}$, since $\tau_i^{(j)}[k] \leq 2\bar{T}, \forall k \geq \bar{T}, \forall i \in \mathcal{V}$

based on Lemma 5. For $k \geq (2j-1)\bar{T}$, taking norms on both sides of (31) yields:

$$\begin{aligned}
\|\mathbf{e}_i^{(j)}[k]\| &\stackrel{(a)}{\leq} \beta_j \left(c_j \left(\frac{\gamma_j}{\rho_j} \right)^{2\bar{T}} \rho_j^k + \gamma_j^{(k-1)} \sum_{q=1}^{(j-1)} h_{jq} \bar{c}_q \sum_{\tau=(k-\tau_i^{(j)}[k])}^{(k-1)} \left(\frac{\rho_q}{\gamma_j} \right)^\tau \right) \\
&\stackrel{(b)}{\leq} \beta_j \left(c_j \left(\frac{\gamma_j}{\rho_j} \right)^{2\bar{T}} \rho_j^k + \gamma_j^{(k-1)} \sum_{q=1}^{(j-1)} h_{jq} \bar{c}_q \sum_{\tau=(k-2\bar{T})}^{(k-1)} \left(\frac{\rho_q}{\gamma_j} \right)^\tau \right) \\
&\stackrel{(c)}{\leq} \beta_j \left(c_j \left(\frac{\gamma_j}{\rho_j} \right)^{2\bar{T}} \rho_j^k + \sum_{q=1}^{(j-1)} \frac{h_{jq} \bar{c}_q}{(\gamma_j - \rho_q)} \left(\frac{\gamma_j}{\rho_q} \right)^{2\bar{T}} \rho_q^k \right) \\
&\stackrel{(d)}{\leq} \bar{c}_j \rho_j^k, \forall k \geq (2j-1)\bar{T}.
\end{aligned} \tag{32}$$

In the above inequalities, (a) follows from the induction hypothesis, equation (29), the bounds on the growth and the norm of \mathbf{A}_{jj} , and by noting that $\tau_i^{(j)}[k] \leq 2\bar{T}$ (based on Lemma 5), $\rho_j < 1$, $\gamma_j \geq 1$, (b) follows by suitably changing the lower limit of the inner summation (over time), a change that is warranted since each term in the summation is non-negative, (c) follows by simplifying the preceding inequality, and (d) follows by noting that $\rho_q < \rho_j, \forall q \in \{1, \dots, j-1\}$ by design. Note that the bound obtained in (32) for each non-source node $i \in \mathcal{V} \setminus \{j\}$ applies also to the source node j , since $\bar{c}_j \geq c_j$. Let $\mathbf{e}_i[k] = \hat{\mathbf{z}}_i[k] - \mathbf{z}[k]$. For any node $i \in \mathcal{V}$, we then obtain the desired result as follows:

$$\|\mathbf{e}_i[k]\| = \sqrt{\sum_{j=1}^N \|\mathbf{e}_i^{(j)}[k]\|^2} \leq \left(\sqrt{\sum_{j=1}^N \bar{c}_j^2} \right) \rho^k, \forall k \geq (2N-1)(N-1)\bar{T}. \tag{33}$$

In particular, we obtain $\lim_{k \rightarrow \infty} \|\mathbf{e}_i[k]\| = 0, \forall i \in \mathcal{V}$. \square

Corollary 1. (Finite-Time Convergence) Suppose the conditions stated in Theorem 1 are met. Then, the observer gains $\mathbf{L}_1, \dots, \mathbf{L}_N$ can be designed in a manner such that the estimation error of each node $i \in \mathcal{V}$ converges to zero in at most $n + 2N(N-1)\bar{T}$ time-steps. \square

Proof. (Sketch) For each substate $j \in \{1, \dots, N\}$, let the gain \mathbf{L}_j be chosen such that $(\mathbf{A}_{jj} - \mathbf{L}_j \mathbf{C}_{jj})$ has all its eigenvalues at 0. Such a choice of \mathbf{L}_j exists since $(\mathbf{A}_{jj}, \mathbf{C}_{jj})$ is observable by construction. Clearly, $(\mathbf{A}_{jj} - \mathbf{L}_j \mathbf{C}_{jj})$ is then a nilpotent matrix of index at most n_j , where $n_j = \dim(\mathbf{A}_{jj})$. Thus, it is easy to see that $\mathbf{e}_1^{(1)}[k] = \mathbf{0}, \forall k \geq n_1$, based on (19). Referring to equation (21), and noting that $\tau_i^{(1)}[k] \leq 2\bar{T}, \forall k \geq \bar{T}, \forall i \in \mathcal{V}$, we obtain: $\mathbf{e}_i^{(1)}[k] = \mathbf{0}, \forall k \geq n_1 + 2\bar{T}, \forall i \in \mathcal{V}$. One can easily generalize this argument to the remaining substates by using an inductive reasoning akin to that employed in the proof of Theorem 1. \square

Remark 1. Notice that given a desired convergence rate ρ , the general design approach described in the proof of Theorem 1 offers a considerable degree of freedom in choosing the parameters ρ_1, \dots, ρ_N , since they only need to satisfy $0 < \rho_1 < \rho_2 < \dots < \rho_N < \rho$. As such, this can be achieved in infinitely many ways, and the design flexibility so obtained in choosing the observer gains can be exploited to optimize transient performance, or performance against noise. In contrast, the proof of Corollary 1 highlights a

specific approach to obtain finite-time convergence. However, such an approach may lead to undesirable transient spikes in the estimation errors, owing to large observer gains. \square

VI. CONCLUSION

In this paper, we developed a new approach towards designing distributed observers that work under the basic assumption of joint observability, and can handle a very general class of time-varying graphs. Unlike existing literature, this is achieved without requiring multiple consensus iterations between consecutive time-steps of the dynamics. Instead, our main idea is based on introducing a metric that keeps track of the age-of-information being diffused across the network, and in turn, acts as measure of quality of such information. We established that any desired exponential convergence rate can be achieved based on our approach. Furthermore, we showed that one can even obtain finite-time convergence via an appropriate choice of the observer gains. Ongoing work involves extending the ideas in this paper to scenarios where certain sensors are potentially under attack [17].

REFERENCES

- [1] 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.
- [2] V. Ugrinovskii, "Conditions for detectability in distributed consensus-based observer networks," *IEEE Trans. on Autom. Control*, vol. 58, no. 10, pp. 2659–2664, 2013.
- [3] T. Kim, H. Shim, and D. D. Cho, "Distributed luenberger observer design," in *Proc. of the 55th IEEE Decision and Control Conference*, 2016, pp. 6928–6933.
- [4] S. Wang and W. Ren, "On the convergence conditions of distributed dynamic state estimation using sensor networks: A unified framework," *IEEE Trans. on Cont. Sys. Tech.*, vol. 26, no. 4, pp. 1300–1316, 2018.
- [5] S. Park and N. C. Martins, "Design of distributed LTI observers for state omniscience," *IEEE Trans. on Autom. Control*, vol. 62, no. 2, pp. 561–576, 2017.
- [6] L. Wang and A. S. Morse, "A distributed observer for a time-invariant linear system," *IEEE Trans. on Autom. Control*, vol. 63, no. 7, 2018.
- [7] W. Han, H. L. Trentelman, Z. Wang, and Y. Shen, "A simple approach to distributed observer design for linear systems," *IEEE Trans. on Autom. Control*, vol. 64, no. 1, pp. 329–336, 2019.
- [8] A. Mitra and S. Sundaram, "Distributed observers for LTI systems," *IEEE Trans. on Autom. Control*, vol. 63, no. 11, pp. 3689–3704, 2018.
- [9] F. F. Rego, A. P. Aguiar, A. M. Pascoal, and C. N. Jones, "A design method for distributed Luenberger observers," in *Proc. of the 56th IEEE Conference on Decision and Control*, 2017, pp. 3374 – 3379.
- [10] L. Wang, A. Morse, D. Fullmer, and J. Liu, "A hybrid observer for a distributed linear system with a changing neighbor graph," in *Proc. of the 56th IEEE Conf. on Decision and Control*, 2017, pp. 1024–1029.
- [11] M. Cao, A. S. Morse, and B. D. Anderson, "Reaching a consensus in a dynamically changing environment: A graphical approach," *SIAM J. Control Optim.*, vol. 47, no. 2, pp. 575–600, 2008.
- [12] S. Mou, J. Liu, and A. S. Morse, "A distributed algorithm for solving a linear algebraic equation," *IEEE Trans. on Autom. Control*, vol. 60, no. 11, pp. 2863–2878, 2015.
- [13] S. Kaul, R. Yates, and M. Gruteser, "Real-time status: How often should one update?" in *IEEE INFOCOM*, 2012, pp. 2731–2735.
- [14] R. Talak, S. Karaman, and E. Modiano, "Minimizing age-of-information in multi-hop wireless networks," in *Proc. Annual Allerton Conf. on Comm., Control, and Computing*, 2017, pp. 486–493.
- [15] A. Mitra, J. A. Richards, S. Bagchi, and S. Sundaram, "Finite-time distributed state estimation over time-varying graphs: Exploiting the age-of-information," *arXiv preprint arXiv:1810.06151*, 2018.
- [16] R. A. Horn, R. A. Horn, and C. R. Johnson, *Matrix analysis*. Cambridge university press, 1990.
- [17] A. Mitra and S. Sundaram, "Byzantine-resilient distributed observers for LTI systems," *arXiv preprint arXiv:1802.09651*, 2018.