A Fully Distributed LTI Estimation Scheme over Directed Graph Topologies

Evagoras Makridis[1], Camilla Fioravanti[2], Gabriele Oliva[2*], Maria Vrakopoulou[1], and Themistoklis Charalambous[1]

Abstract—**This paper introduces a distributed state estimation scheme for linear time-invariant (LTI) discrete-time systems, where observers, with partial observation of the system, communicate with each other over a directed and strongly connected (but not necessarily balanced) graph topology and execute a predefined number of average consensus steps in-between estimation steps. Our methodology departs from previous works in the literature in that it does not require any degree of centralized design nor relies on procedures that might be prone to numerical instability. By leveraging ratio consensus and matrix perturbation theory, we establish a convergence-guaranteeing condition for the number of consensus iterations needed between the steps of the distributed estimation process. This condition becomes the blueprint for a distributed initialization procedure, which allows the agents to collectively select an adequate number of ratio consensus steps.**

Index Terms—**Distributed Estimation, Ratio Consensus, Matrix Perturbation Analysis, LTI Systems.**

In recent years, the distributed estimation problem has gained significant attention in the context of multi-agent systems [1], [2], due to its importance in several applications of networked systems, such as communication, networked control, monitoring, and surveillance.

Several innovative solutions have been proposed, especially in the case of linear observers [3]–[10]. Studies such as [4], where Kalman's decomposition has been applied to ease the estimation task for the agents, or [7], which focuses on switching networks, represent important examples in this sense. Furthermore, works like those by Wang et al. [5], [6] and Savas et al. [8] have made significant contributions to both discrete and continuous-time distributed estimation theory. In particular, in [5], following the same path as the pioneering algorithm in [3], agents are allowed to perform a specified number of distributed agreement rounds between state updates, and the distributed estimators are designed so that the estimation error converges exponentially to zero at a fixed rate that is arbitrarily chosen, under the condition that the graph is always strongly connected. Interestingly, in [3] time-varying measurements and graph topologies are considered, although it is assumed that there is a subset of core measurements and that the agents are provided with global meta-information regarding such core measurements. Moreover, in [8] it is shown that, if the agents are allowed to execute a given number of average consensus steps in between estimation steps, then the overall performance of the distributed estimation scheme is greatly improved. Also, in [10] a bound on the number of steps required for stability is given, although the bound is based on knowledge of the second largest eigenvalues of the consensus dynamical matrix. However, the approaches in [3], [8], [10] assume that the agents interact over an undirected graph topology; furthermore, the schemes in [5], [8] assume that the linear estimators (gains) and the number of steps require knowledge of global information. In [11] we overcome these limitations by relying on minimum-time ratio consensus, which guarantees that agents reach an exact agreement in-between estimation steps. Moreover, we develop a distributed initialization procedure that allows the agents to distributively compute the gains. The approach in [11], although effective, relies on the construction of large Hankel matrices and on the identification of characteristic polynomials, a process that might be affected by numerical instability.

In this paper, we address the following question: *is it possible to agree on a number of consensus steps in-between estimation steps such that the convergence of the distributed estimation process is guaranteed in a fully distributed fashion?* Towards this end, we develop a distributed state estimation scheme for LTI discrete-time system where, between estimations, the agents execute a given number of asymptotic ratio consensus steps over a directed and strongly connected (but not necessarily balanced) graph. The main idea of the paper is that the dynamics of the estimation error can be represented as the sum of two terms: (i) a nominal term where the agents achieve perfect agreement in between estimation steps and (ii) a residual term that is due to the actual imperfect agreement. By handling the second term as a perturbation and resorting to matrix perturbation theory, we identify a condition on the number of ratio-consensus steps that guarantees convergence of the overall distributed estimation process. Interestingly, the identified condition translates into a distributed initialization procedure that allows the agents to collectively select an adequate number of ratio consensus steps.

The outline of the paper is as follows: Section I introduces the necessary notation and mathematical machinery; Section II

[1]Department of Electrical and Computer Engineering, School of Engineering, University of Cyprus, 2109 Aglantzia, Nicosia, Cyprus. E-mails: {surname.name}@ucy.ac.cy.

[2]Department of Engineering, University Campus Bio-Medico of Rome, Via Alvaro del Portillo, 21 - 00128 Roma, Italy. E-mails: {c.fioravanti, g.oliva}@unicampus.it.

The work of E. Makridis and T. Charalambous was partly funded by the project MINERVA, which received funding from the European Research Council (ERC) under the European Union's Horizon 2022 research and innovation programme (Grant Agreement No. 101044629).

* corresponding author.

states the problem at hand in this paper; Section III develops a sufficient condition for stability; Section IV provides a distributed initialization procedure that allows the agents to fulfill the necessary stability condition; Section V provides some simulations to corroborate the theoretical findings; finally, Section VI collects some conclusive remarks and future work directions.

I. PRELIMINARIES

A. Notation and Graph Theory

We denote vectors with boldface lowercase letters and matrices with uppercase letters. The transpose of matrix A and vector $\boldsymbol{x}$ are denoted as $A^\top$, $\boldsymbol{x}^\top$, respectively. We refer to the (i,j)-th entry of a matrix A by A_{ij}. We represent by $\mathbf{0}_n$ and $\mathbf{1}_n$ vectors with n entries, all equal to zero and to one, respectively. Given two matrices $A \in \mathbb{R}^{n\times m}$ and $B \in \mathbb{R}^{p\times q}$, we use $A \otimes B$ to denote their Kronecker product.

Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ be a directed graph (digraph) with N nodes $\mathcal{V} = \{v_1, v_2, \ldots, v_N\}$ and e edges $\mathcal{E} \subseteq \mathcal{V}\times\mathcal{V}$, where $(v_i, v_j) \in \mathcal{E}$ captures the existence of a link from node v_i to node v_j. A directed graph is *strongly connected* if each node can be reached by every other node via the edges, respecting their orientation. Let the in-neighborhood $\mathcal{N}_i^{\mathtt{in}}$ of a node $v_i \in \mathcal{V}$ be the set of nodes $v_j \in \mathcal{V}$ such that $(v_j, v_i) \in \mathcal{E}$; similarly, the out-neighborhood $\mathcal{N}_i^{\mathtt{out}}$ of a node $v_i \in \mathcal{V}$ is the set of nodes $v_j \in \mathcal{V}$ such that $(v_i, v_j) \in \mathcal{E}$. The *in-degree* $d_i^{\mathtt{in}}$ of a node v_i is the number of its incoming edges, *i.e.,* $d_i^{\mathtt{in}} = |\mathcal{N}_i^{\mathtt{in}}|$; similarly, the *out-degree* $d_i^{\mathtt{out}}$ of a node v_i is the number of its outgoing edges, *i.e.,* $d_i^{\mathtt{out}} = |\mathcal{N}_i^{\mathtt{out}}|$. The minimum distance from node v_i to v_j for $v_i, v_j \in \mathcal{V}, i \neq j$, is the shortest path from node v_i to v_j and is denoted $\mathtt{d}_{\min}(i,j)$, while in the absence of a directed path then $\mathtt{d}_{\min}(i,j) = \infty$. The diameter of $\mathcal{G}$ is defined as the longest shortest path between any two nodes *i.e.,* $D = \max_{i,j\in\mathcal{V}, i\neq j} \mathtt{d}_{\min}(i,j)$. We use $\|\cdot\|_2$, $\|\cdot\|_1$, and $\|\cdot\|_\infty$ to denote the Euclidean, one-, and infinity- norms, respectively.

B. Perturbation Theory

In the context of matrix analysis, perturbation theory aims to locate the eigenvalues of a perturbed matrix based on knowledge of both the unperturbed and perturbation matrix. Among several other results, the *Bauer-Fike Theorem* is particularly useful, since it applies to non-symmetric matrices and does not require the perturbation to be small.

Theorem 1 (Bauer and Fike [12]): Let M be an $n \times n$ diagonalizable matrix satisfying $M = V\Lambda V^{-1}$ and let E be an arbitrary $n \times n$ matrix. Every eigenvalue μ of $M + E$ satisfies the inequality $|\mu - \lambda| \leq \|V\|\|V^{-1}\|\|E\|$, where λ is some eigenvalue of M, and $\|\cdot\|$ is any operator norm.

C. Ratio Consensus

In this section we introduce the distributed ratio consensus algorithm [13], [14] to achieve average consensus over a set of N agents (associated with sensor measurements) exchanging information through a possibly unbalanced directed network. In particular, each agent $v_i \in \mathcal{V}$ maintains a local state variable $\chi_i[\cdot] \in \mathbb{R}^d$ and an auxiliary variable $\psi_i[\cdot] \in \mathbb{R}$ that compensates for the unbalanced directed network topology. At the beginning of the procedure, each agent v_i initializes $\chi_i[0] = \chi_{i0}$ and its auxiliary variable at $\psi_i[0] = 1$. Subsequently, at each iteration $m \geq 0$, each agent $v_i \in \mathcal{V}$ assigns a weight to its variables and transmits the weighted variables $p_{li}\chi_i[m]$ and $p_{li}\psi_i[m]$ to its out-neighbors $v_l \in \mathcal{N}_i^{\mathtt{out}}$ over the links $(v_i, v_l) \in \mathcal{E}$. The weight p_{li} can be assigned using the *equal-neighbor* model[1], *i.e.,* $p_{li} = 1/d_i^{\mathtt{out}}$ if $v_l \in \mathcal{N}_i^{\mathtt{out}}$ and $p_{li} = 0$ otherwise. Placing each weight p_{li} on the l-th row and i-th column of a matrix $P \in \mathbb{R}_{\geq 0}^{N\times N}$, we obtain a column-stochastic matrix. At each consensus iteration m, each node $v_i \in \mathcal{V}$ receives the state variables transmitted by its in-neighbors $p_{ij}\chi_j[m]$ and $p_{ij}\psi_j[m]$, and updates its own variables as follows:

$$\chi_i[m+1] = \sum_{j\in\mathcal{N}_i^{\mathtt{in}}} p_{ij}\chi_j[m], \quad \chi_i[0] = \chi_{i0}, \tag{1a}$$

$$\psi_i[m+1] = \sum_{j\in\mathcal{N}_i^{\mathtt{in}}} p_{ij}\psi_j[m], \quad \psi_i[0] = 1. \tag{1b}$$

To further analyze the ratio consensus algorithm, we concatenate the agents' individual variables into global network variables $\chi[m] = \big[\chi_1[m] \cdots \chi_N[m]\big]^\top \in \mathbb{R}^{N\times d}$ and $\psi[m] = \big[\psi_1[m] \cdots \psi_N[m]\big]^\top \in \mathbb{R}^N$, and thus we can rewrite the above distributed algorithm in its equivalent network matrix form:

$$\chi[m+1] = P\chi[m] = \Phi_P(m)\chi[0], \tag{2a}$$

$$\psi[m+1] = P\psi[m] = \Phi_P(m)\psi[0], \tag{2b}$$

where $\Phi_P(m) := P^{m+1}$. Now, rewriting the above iterations at node v_i in a equivalent way that was first presented in [15] (in which the weights involved in the update of the ratio are independent of $\chi[m]$, and they are instead based on $\psi[m]$), we obtain the ratio $z_i[m+1] = \chi_i[m+1]/\psi_i[m+1]$. For simplicity of exposition, assuming $d = 1$, we have that

$$\begin{aligned} z_i[m+1] &= \frac{\sum_{j\in\mathcal{N}_i^{\mathtt{in}}} p_{ij}\chi_j[m]}{\sum_{j\in\mathcal{N}_i^{\mathtt{in}}} p_{ij}\psi_j[m]} = \sum_{j\in\mathcal{N}_i^{\mathtt{in}}} \left(\frac{p_{ij}\chi_j[m]}{\sum_{l\in\mathcal{N}_i^{\mathtt{in}}} p_{il}\psi_l[m]}\right) \\ &= \sum_{j\in\mathcal{N}_i^{\mathtt{in}}} \left(\frac{p_{ij}\psi_j[m]}{\sum_{l\in\mathcal{N}_i^{\mathtt{in}}} p_{il}\psi_l[m]}\right) z_j[m] = \sum_{j\in\mathcal{N}_i^{\mathtt{in}}} s_{ij}[m] z_j[m]. \end{aligned} \tag{3}$$

Interestingly the time-varying matrix $S[m] \in \mathbb{R}^{N\times N}$ collecting the terms $s_{ij}[m]$ is nonnegative and row-stochastic [15]. Hence, one can also express the ratio consensus algorithm in its matrix form as

$$z[m+1] = S[m]z[m] = \Phi_S(m)\chi[0], \tag{4}$$

where $\Phi_S(m)$ is the backward product of the row-stochastic matrices $\Phi_S(m) = S[m]\cdots S[1]S[0]$, for $m \geq 0$. Notice that, in the general case where $d > 1$, it can be easily shown that

[1] This strategy ensures that the total mass of the variables is equally allocated to the out-neighbors of v_i, and that $\sum_{l\in\mathcal{N}_i^{\mathtt{out}}} p_{li} = 1$.

$\Phi_S(m)$ in Eq. (4), can be replaced by $\bar{\Phi}_S(m) = \Phi_S(m) \otimes I_d$.

II. Problem Formulation

Consider the following system dynamics with state $\boldsymbol{x}[k] \in \mathbb{R}^d$, and $\boldsymbol{y}_i[k] \in \mathbb{R}^q$, which evolve with time as follows:

$$\boldsymbol{x}[k+1] = A\boldsymbol{x}[k], \quad \boldsymbol{y}_i[k] = C_i\boldsymbol{x}[k], \tag{5}$$

where $\boldsymbol{y}_i[k]$ and C_i denote the observations and observation matrix of node $v_i \in \mathcal{V}$, respectively. In the following, we assume that the system is *jointly observable*, *i.e.,* the pair (A, C) with $C = \begin{pmatrix} C_1^\top & \dots & C_N^\top \end{pmatrix}^\top$ is observable. The state estimate of node v_i is given by

$$\hat{\boldsymbol{x}}_i[k+1] = A\tilde{\boldsymbol{x}}_i[k] + L_i\Big(\boldsymbol{y}_i[k] - C_i\tilde{\boldsymbol{x}}_i[k]\Big), \tag{6}$$

where $\tilde{\boldsymbol{x}}_i[k]$ is the result of m-steps of ratio consensus over the strongly connected digraph $\mathcal{G}$, using $\hat{\boldsymbol{x}}_i[k]$ as the initial conditions, i.e,

$$\tilde{\boldsymbol{x}}_i[k] = \sum_{j=1}^{N} \Big(\Phi_S(m)\Big)_{ij} \hat{\boldsymbol{x}}_j[k]$$

is the resulting fused consensus estimate after m iterations of consensus[2]. In a compact form, defining $\tilde{\boldsymbol{x}}[k]$ and $\hat{\boldsymbol{x}}[k]$ as the stack of the vectors $\tilde{\boldsymbol{x}}_i[k]$ and $\hat{\boldsymbol{x}}_j[k]$, respectively, we have that

$$\tilde{\boldsymbol{x}}[k] = \underbrace{\Big(\Phi_S(m) \otimes I_d\Big)}_{\bar{\Phi}_S(m)} \hat{\boldsymbol{x}}[k].$$

In the remainder of this section, we define the dynamics of the error between the actual state of the system and the estimated one. In particular, the error dynamics can be seen as the result of two contributions, *i.e.,* the dynamics obtained assuming exact consensus among the agents at each estimation step and a term that accounts for the fact that perfect consensus is not reached. This will be beneficial for our convergence analysis, where we will interpret the second term as a perturbation.

A. Error Dynamics

Let us now define the local estimation error at node v_i as $\boldsymbol{e}_i[k] = \boldsymbol{x}[k] - \hat{\boldsymbol{x}}_i[k]$, and let $\boldsymbol{e}[k]$ denote the stack of the vectors $\boldsymbol{e}_i[k]$. Based on the above definition, the local estimation error evolves with time as follows:

$$\begin{aligned} \boldsymbol{e}_i[k+1] &= \boldsymbol{x}[k+1] - \hat{\boldsymbol{x}}_i[k+1] \\ &= A\boldsymbol{x}[k] - A\tilde{\boldsymbol{x}}[k] - L_iC_i\Big(\boldsymbol{x}[k] - \tilde{\boldsymbol{x}}_i[k]\Big) \\ &= (A - L_iC_i)\Big(\boldsymbol{x}[k] - \tilde{\boldsymbol{x}}_i[k]\Big) \end{aligned}$$

which, in stacked form, reads as follows

$$\boldsymbol{e}[k+1] = \underbrace{\begin{pmatrix} A - L_1C_1 & & 0 \\ & \ddots & \\ 0 & & A - L_NC_N \end{pmatrix}}_{\bar{\mathcal{A}}} \begin{pmatrix} \boldsymbol{x}[k] - \tilde{\boldsymbol{x}}_1[k] \\ \vdots \\ \boldsymbol{x}[k] - \tilde{\boldsymbol{x}}_N[k] \end{pmatrix}. \tag{7}$$

[2]Notice that the iterator k represents the estimation iterations and thus it does not change while computing $\tilde{\boldsymbol{x}}_i[k]$

Notably, since by construction $\Phi_S(m)\mathbf{1}_N = \mathbf{1}_N$, we have that $\bar{\Phi}_S(m)(\mathbf{1}_N \otimes \boldsymbol{x}[k]) = \mathbf{1}_N \otimes \boldsymbol{x}[k]$, and thus

$$\begin{aligned} \begin{pmatrix} \boldsymbol{x}[k] - \tilde{\boldsymbol{x}}_1[k] \\ \vdots \\ \boldsymbol{x}[k] - \tilde{\boldsymbol{x}}_N[k] \end{pmatrix} &= \begin{pmatrix} \boldsymbol{x}[k] \\ \vdots \\ \boldsymbol{x}[k] \end{pmatrix} - \bar{\Phi}_S(m) \begin{pmatrix} \hat{\boldsymbol{x}}_1[k] \\ \vdots \\ \hat{\boldsymbol{x}}_N[k] \end{pmatrix} \\ &= \bar{\Phi}_S(m) \begin{pmatrix} \boldsymbol{x}[k] - \hat{\boldsymbol{x}}_1[k] \\ \vdots \\ \boldsymbol{x}[k] - \hat{\boldsymbol{x}}_N[k] \end{pmatrix} = \bar{\Phi}_S(m)\boldsymbol{e}[k], \end{aligned}$$

which, plugged in Eq. (7), yields

$$\boldsymbol{e}[k+1] = \bar{\mathcal{A}}\,\bar{\Phi}_S(m)\ \boldsymbol{e}[k]. \tag{8}$$

Notice that, in the limit of m approaching infinity, the average consensus yields the exact average for all agents; in other words, we have that $\Phi_S(\infty) = \lim_{m\to\infty} \Phi_S(m) = \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top$. In general, for a finite value of m we have $\Phi_S(m) = \Phi_S(\infty) + \Delta(m)$, where $\Delta(m) \neq 0$ accounts for the error of the coefficients $(\Phi_S(m))_{ij}$ with respect to the asymptotic coefficients $1/N$. Based on the above definition, we have that

$$\begin{aligned} \bar{\Phi}_S(m) &:= \Phi_S(m) \otimes I_d = \big(\Phi_S(\infty) - \Delta(m)\big) \otimes I_d \\ &= \Phi_S(\infty) \otimes I_d - \Delta(m) \otimes I_d \end{aligned}$$

and we can express Eq. (8) as

$$\begin{aligned} \boldsymbol{e}[k+1] &= \bar{\mathcal{A}}\Big(\Phi_S(\infty) \otimes I_d - \Delta(m) \otimes I_d\Big)\boldsymbol{e}[k] && \text{(9a)} \\ &= \Big(\underbrace{\bar{\mathcal{A}}\,\Phi_S(\infty) \otimes I_d}_{\Gamma_\infty} - \underbrace{\bar{\mathcal{A}}\,\Delta(m) \otimes I_d}_{\Gamma_{\Delta}(m)}\Big)\boldsymbol{e}[k]. && \text{(9b)} \end{aligned}$$

III. Convergence Analysis

In this section, we prove that, for a sufficiently large number of consensus steps, the error dynamics is asymptotically convergent. In doing so, we develop a distributed initialization procedure that guarantees the selection of an adequate number of consensus steps.

A. Ideal Error Dynamics

The following lemma characterizes the stability of the ideal error dynamical matrix Γ_∞, which corresponds to an exact consensus among agents in-between estimation steps.

Lemma 1: Assume the system in Eq. (5) is jointly observable and let the local observer gains L_i be such that the spectral radius ρ satisfies $\rho\left(A^\dagger\right) < 1$, where $A^\dagger := A - \frac{1}{N}\sum_{i=1}^{N} L_iC_i$. Then, Γ_∞ is Schur stable.

Proof: In order to prove the statement, we observe that

$$\Gamma_\infty = \frac{1}{N}\begin{pmatrix} A - L_1C_1 & \cdots & A - L_1C_1 \\ \vdots & \ddots & \vdots \\ A - L_NC_N & \cdots & A - L_NC_N \end{pmatrix}.$$

At this point, let us consider the $Nd \times Nd$ matrix T, reported next, along with its inverse T^{-1}, i.e,

$$T = \begin{pmatrix} I_d & -I_d & \cdots & -I_d \\ 0 & I_d & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_d \end{pmatrix}, \; T^{-1} = \begin{pmatrix} I_d & I_d & \cdots & I_d \\ 0 & I_d & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_d \end{pmatrix}.$$

By some algebra, it can be noted that Γ_∞ is similar to

$$\widetilde{\Gamma}_\infty = T^{-1}\,\Gamma_\infty\,T = \begin{pmatrix} A^\dagger & 0 & \cdots & 0 \\ \frac{1}{N}\left(A - L_2C_2\right) & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{N}\left(A - L_NC_N\right) & 0 & \cdots & 0 \end{pmatrix}.$$

Hence, the eigenvalues of Γ_∞ are either zero or coincide with the eigenvalues of $A^\dagger$. Since we assumed $\rho(A^\dagger) < 1$, we conclude that Γ_∞ is Schur stable. This completes our proof. ■

Remark 1: In the above lemma, we assume that the gains are such that $A^\dagger$ is Schur stable. A possible distributed approach to select these gains is to resort to the token-passing procedure discussed in [11].

In view of the later developments in this section, it is convenient to explicitly identify the transform matrices that put Γ_∞ in diagonal form. These matrices will be essential in order to prove stability of the error dynamics.

Lemma 2: Assume $A^\dagger$ is diagonalizable and nonsingular[3] and let $\Pi_j = (A - L_jC_j)\left(A^\dagger\right)^{-1}$. Moreover, let Q be the matrix that diagonalizes $A^\dagger$, *i.e.,* such that $Q^{-1}A^\dagger Q$ is diagonal. Matrix Γ_∞ is diagonalizable, *i.e.,* $\Lambda_\infty = V^{-1}\Gamma_\infty V$ is diagonal with

$$V^{-1} = \begin{pmatrix} (I_d - \frac{1}{N}\sum_{i=2}^N \Pi_i)Q^{-1} & -I_d & \cdots & -I_d \\ \frac{1}{N}\Pi_2 Q^{-1} & I_d & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{N}\Pi_N Q^{-1} & 0 & \cdots & I_d \end{pmatrix},$$

and

$$V = \begin{pmatrix} Q & Q & \cdots & \cdots & Q \\ -\frac{1}{N}\Pi_2 & I_d - \frac{1}{N}\Pi_2 & -\frac{1}{N}\Pi_2 & \cdots & -\frac{1}{N}\Pi_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ -\frac{1}{N}\Pi_N & \cdots & \cdots & -\frac{1}{N}\Pi_N & I_d - \frac{1}{N}\Pi_N \end{pmatrix}.$$

Proof: Let us consider

$$W = \begin{pmatrix} Q^{-1} & 0 & \cdots & 0 \\ \frac{1}{N}\Pi_2 Q^{-1} & I_d & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{N}\Pi_N Q^{-1} & 0 & \cdots & I_d \end{pmatrix},$$

which is block triangular and, thus, nonsingular. We have that

$$W^{-1} = \begin{pmatrix} Q & 0 & \cdots & 0 \\ -\frac{1}{N}\Pi_2 & I_d & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{N}\Pi_N & 0 & \cdots & I_d \end{pmatrix}.$$

By some algebraic manipulation, we obtain

$$W^{-1}\widetilde{\Gamma}_\infty W = W^{-1}T^{-1}\Gamma_\infty T W = \Lambda_\infty,$$

with Λ_∞ diagonal. Noting that $V = TW$, the proof is complete. ■

B. Perturbation Analysis

In order to model the effect of the imperfect agreement among the agents in terms of a perturbation, we will resort on Bauer-Fike Theorem. To this end, let us now provide some ancillary results.

Lemma 3: Let the assumptions of Lemma 2 hold true. Then it holds $\|V\|_1 \le \xi$ and $\|V^{-1}\|_1 \le \theta$, where

$$\xi = 1 + \|Q\|_1 + \frac{N-1}{N}\max_{i=1,\dots,N}\|\Pi_i\|_1$$

and

$$\theta = \max\left\{2, (1 + 2\frac{N-1}{N}\max_{i=1,\dots,N}\|\Pi_i\|_1)\|Q^{-1}\|_1\right\}.$$

Proof: The proof follows noting that, given the structure of V and V^{-1} given in Lemma 2,

$$\begin{aligned} \|V\|_1 &\le \|Q\|_1 + \|I_d\|_1 + \frac{1}{N}\sum_{i=2}^N \|\Pi_i\|_1 \\ &\le 1 + \|Q\|_1 + \frac{N-1}{N}\max_{i=1,\dots,N}\|\Pi_i\|_1 \end{aligned}$$

and that $\|V^{-1}\|_1 \le \max\{2, \zeta\}$, with

$$\begin{aligned} \zeta &= \left\|(I_d - \frac{1}{N}\sum_{i=2}^N \Pi_i)Q^{-1}\right\|_1 + \frac{1}{N}\sum_{i=2}^N \|\Pi_i\|_1\|Q^{-1}\|_1 \\ &\le \left(1 + 2\frac{N-1}{N}\max_{i=1,\dots,N}\|\Pi_i\|_1\right)\|Q^{-1}\|_1. \end{aligned}$$

■

As a last ancillary result, let us now characterize an upper bound on the one-norm of $\Delta(m)$. Such a result will be the cornerstone for a distributed initialization procedure for choosing an adequately large $m \in \mathbb{N}$.

Lemma 4: It holds $\|\Delta(m)\|_1 \le \max_{i=1,\dots,N}\|\Phi_S(m)\boldsymbol{\eta}_i\|_\infty$, where

$$\boldsymbol{\eta}_i = \mathbb{e}_i - \frac{1}{N}\mathbf{1}_N, \tag{10}$$

with $\mathbb{e}_i$ being the i-th vector in the canonical basis in $\mathbb{R}^N$.

Proof: In order to prove the statement we observe that, being $\Phi_S(m)$ row stochastic, it holds

$$\frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top = \Phi_S(m)\left(\frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top\right);$$

[3] These assumptions are not restrictive. In fact, since we assumed joint observability, we have that the eigenvalues of $A^\dagger$ can be arbitrarily selected (e.g., via the token passing approach in [11]). Thus, by selecting distinct and nonzero eigenvalues for $A^\dagger$, we are guaranteed that it is diagonalizable and nonsingular.

hence, we have that

$$\Delta(m) = \Phi_S(m) - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top = \Phi_S(m)\left(I_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top\right).$$

At this point we observe that $\boldsymbol{\eta}_i$ is the i-th column of $I_N - \frac{1}{N}\mathbf{1}_N\mathbf{1}_N^\top$. Noting that the 1-norm of a matrix is the largest among the ∞-norms of its columns, we conclude that

$$\begin{aligned}\|\Delta(m)\|_1 &= \left\|\Phi_S(m)\begin{pmatrix}\boldsymbol{\eta}_1 & \dots & \boldsymbol{\eta}_N\end{pmatrix}\right\|_1 \\ &= \left\|\begin{pmatrix}\Phi_S(m)\boldsymbol{\eta}_1 & \dots & \Phi_S(m)\boldsymbol{\eta}_N\end{pmatrix}\right\|_1, \\ &= \max_{i=1,\dots,N}\|\Phi_S(m)\boldsymbol{\eta}_i\|_1.\end{aligned}$$

This completes our proof. ■

C. Proof of Convergence

We are now in position to prove convergence of the proposed distributed state estimation scheme.

Theorem 2: Let the assumptions of Lemma 2 hold true. The error dynamics is asymptotically stable if there is a number of ratio consensus steps $m^* \in \mathbb{N}$ such that

$$\max_{i=1,\dots,N}\|\Phi_S(m^*)\boldsymbol{\eta}_i\|_1 < \frac{1}{c}\left(1-\rho\left(A^\dagger\right)\right), \qquad (11)$$

with $\boldsymbol{\eta}_i$ defined in Eq. (10) and $c = \xi\,\theta\,\max_i\{\|A - L_iC_i\|_1\}$.

Proof: In order to prove the result we observe that, by using the Bauer-Fike theorem, the eigenvalue μ of $\bar{\mathcal{A}}\,\bar{\Phi}_P(m^*)$ with largest magnitude satisfies

$$\begin{aligned}|\mu-\lambda| &\le \|V\|_1\|V^{-1}\|_1\|\Gamma_\Delta(m^*)\|_1 \\ &\le \|V\|_1\|V^{-1}\|_1\|\bar{\mathcal{A}}\|_1\|\Delta(m^*)\otimes I_d\|_1 \\ &\le \|V\|_1\|V^{-1}\|_1\max_i\{\|A-L_iC_i\|_1\}\|\Delta(m^*)\|_1,\end{aligned}$$

where λ is one of the eigenvalues of Γ_∞ and V is the matrix of all eigenvectors of Γ_∞. Moreover, by Lemma 3, we have that $\|V\|_1\|V^{-1}\|_1 \le \xi\,\theta$, hence $\|V\|_1\|V^{-1}\|_1\max_i\{\|A-L_iC_i\|_1\} \le c$. Therefore, by Lemma 4, we have that

$$|\mu-\lambda| \le c\max_{i=1,\dots,N}\|\Phi_S(m^*)\boldsymbol{\eta}_i\|_1 < 1-\rho(A^\dagger).$$

Since $A^\dagger$ is Hurwitz stable, we have that $|\lambda| < 1$; hence, we conclude that $|\mu| < 1$, hence the real system is Schur stable. The proof is complete. ■

IV. A Distributed Initialization Procedure

Theorem 2 guarantees asymptotic stability of the proposed distributed estimation scheme as long as a suitably large m^* is identified. This section presents a distributed initialization procedure to select such an m^*.

Let us now show how to identify m_i^* such that $\|\Phi_S(m_i^*)\boldsymbol{\eta}_i\|_1 < \epsilon$ for a given $\epsilon > 0$. In order to accomplish this task, let us assume that the agents execute a ratio consensus procedure with initial condition[4] $\boldsymbol{z}[0] = \boldsymbol{\eta}_i$. Whenever k is an integer multiple of D we execute a max-consensus with initial conditions $\boldsymbol{r}[k] = |\boldsymbol{z}[k]|$, where $|\cdot|$ is the component-wise absolute value, so that at time $k + D$ all agents know

$$\max_i|z_i[k]| = \|\boldsymbol{z}[k]\|_\infty = \|\Phi_S(k)\boldsymbol{\eta}_i\|_\infty \ge \frac{1}{n}\|\Phi_S(k)\boldsymbol{\eta}_i\|_1.$$

Based on this knowledge, the agents stop the ratio consensus procedure at time hD, with h integer, if $\max_j|z_j[(h-1)]| < \frac{1}{cn}\left(1-\rho\left(A^\dagger\right)\right)$; then, it holds $\|\Phi_S(k)\boldsymbol{\eta}_i\|_1 < \frac{1}{c}\left(1-\rho\left(A^\dagger\right)\right)$, and the procedure returns $m_i^* = (h-1)D$. Let us now assume that the agents execute the above procedure for all $i \in \{1,\dots,N\}$, thus obtaining $m_1^*,\dots,m_N^*$. By setting $m^* = \max_i\{m_i^*\}$, the agents are guaranteed that m^* satisfies the condition in Theorem 2 and that the error dynamics is asymptotically convergent.

Remark 2: The constant c can be computed in a distributed fashion during the initialization. In fact, assuming all agents know $A^\dagger$ (e.g, by accumulating information in a token as done in [11]), $A - L_iC_i$ (as they compute it locally), and Π_i (also computed locally if the agents know $A^\dagger$), they can run max-consensus procedures [16] to compute $\max_{i=1,\dots,N}\|\Pi_i\|_1$, $\max_i\{\|A-L_iC_i\|_1\}$ and N, while they can compute Q locally.

V. Simulations

Consider a directed network $\mathcal{G} = \{\mathcal{V},\mathcal{E}\}$ comprised of four nodes, *i.e.,* $N = 4$, associated with the column-stochastic matrix P where the weights p_{li} are assigned by each node $v_i \in \mathcal{V}$ as described in Section I-C. Moreover, each node has access to an LTI system of $d = 8$ states, modeled by the dynamical matrix $A \in \mathbb{R}^{d\times d}$. Within this setup, the nodes aim at cooperatively estimating the state of the LTI system by exchanging and coordinating their local state estimates through $m^\star$-rounds of the ratio consensus algorithm. The local estimates are computed using (6) where the local estimation gains for node v_i is given by the i-th column of matrix L. Here it is important to note that, each node can determine the rounds of ratio consensus $m^\star$ that guarantee stability, through the distributed initialization procedure presented in Section IV.

$$A=\begin{bmatrix}0.94 & 0.5 & 0 & 0 & 0.39 & 0 & 0 & 0\\ 1.1 & 0 & 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 2.3 & 0.5 & 0 & 0 & 0 & 0\\ 0 & 0 & 0.9 & 0.9 & 0 & 0 & 0 & 0\\ 1.1 & 0 & 0 & 0 & -0.1 & 0.3 & 0 & 0\\ 1.1 & 0 & 0 & 0 & -0.1 & 0.2 & 0 & 0\\ -0.1 & 0 & 0 & 0 & 0 & 0 & 1.1 & 0.2\\ -2 & 0 & 0 & 0 & 0 & 0 & 2.99 & 0.2\end{bmatrix},\ L=\begin{bmatrix}0.4 & 0 & 0 & 0\\ 4 & 0 & 0 & 0\\ 0 & 8.4 & 0 & 0\\ 0 & 4 & 0 & 0\\ 4.4 & 0 & 0 & 0\\ 4 & 0 & 0 & 0\\ -0.4 & 0 & 0 & 4\\ -8 & 0 & 0 & 12\end{bmatrix}$$

$$P=\begin{bmatrix}1/2 & 0 & 1/3 & 1/3\\ 0 & 1/2 & 1/3 & 0\\ 0 & 1/2 & 0 & 1/3\\ 1/2 & 0 & 1/3 & 1/3\end{bmatrix}.$$

Furthermore, each node v_i in the network is associated with sensors and it has computational capability to run the ratio consensus algorithm and compute its local state estimates. In this particular example, each node is assumed to observe only one state of the system, *i.e.,* the local measurement matrices C_i are $C_1 = \begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0 & 0 & 0\end{bmatrix}$, $C_2 = \begin{bmatrix}0 & 0 & 1 & 0 & 0 & 0 & 0 & 0\end{bmatrix}$, $C_3 = \begin{bmatrix}0 & 0 & 0 & 0 & 1 & 0 & 0 & 0\end{bmatrix}$, and $C_4 = \begin{bmatrix}0 & 0 & 0 & 0 & 0 & 0 & 1 & 0\end{bmatrix}$.

[4]The agents need to know N to set their initial conditions to $\boldsymbol{\eta}_i$; to this end, they could resort to max-consensus [16].

In Fig. 1 we present the coordinated number of rounds of ratio consensus $m^\star$, that guarantees convergence of the estimation error to 0 as $k \to \infty$, for different predefined consensus errors $\max_{i=1,\dots,N} \|\Phi_P(m)\boldsymbol{\eta}_i\|_1$.

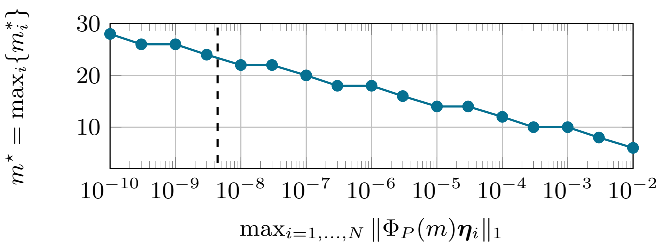

Figure 1: Number of sufficient ratio consensus iterations $m^\star$ for different predefined consensus errors $\max_{i=1,\dots,N} \|\Phi_P(m)\boldsymbol{\eta}_i\|_1$. The dashed vertical line provides an upper bound on $\max_{i=1,\dots,N} \|\Phi_S(m)\boldsymbol{\eta}_i\|_1$ associated to the minimum number of consensus iterations that satisfies the sufficient stability condition in Theorem 2.

Fig. 2 depicts the convergence rate of our proposed distributed estimation algorithm with respect to the Euclidean norm of the estimation error $\boldsymbol{e}[k]$, for different lengths of consensus iterations $m^\star = \{2, 5, 10, 25\}$. According to the upper bound $\max_{i=1,\dots,N} \|\Phi_P(m)\boldsymbol{\eta}_i\|_1$ shown in Fig. 1, with $m^\star \geq 24$ ratio consensus iterations, it is guaranteed that the estimation error will converge to 0 asymptotically, as depicted in Fig. 2. Notably, the distributed estimators achieve asymptotic convergence also for smaller values of m^*, e.g., for $m^\star = 10$; in fact, although allowing the agents to choose $m^\star$ that guarantees convergence, Theorem 2 provides only a sufficient condition. When m becomes too small (e.g., $m^\star \in \{2, 5\}$) we observe that the overall estimation process becomes unstable, and thus the estimation error grows unbounded.

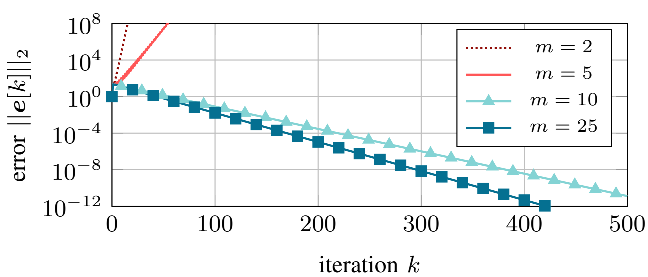

Figure 2: Convergence of the proposed distributed estimation algorithm within network $\mathcal{G}$ for $m^\star = \{2, 5, 10, 25\}$.

VI. Conclusions and Future Directions

This paper presents a scheme for distributed state estimation for LTI discrete-time systems, enabling agents to perform a specified number of average consensus steps between estimations across a directed and strongly connected graph topology, which may not necessarily be balanced. Unlike prior works, our approach eliminates the need for any centralized design and addresses numerical instabilities, tackling the issue of imperfect agreement's impact on estimation accuracy. Utilizing matrix perturbation theory, we define a condition that ensures the convergence of the estimation process. This condition serves as a foundation for a distributed initialization procedure, facilitating the collective determination by agents of an adequate number of consensus steps.

We envisage three main directions for future work: (i) characterizing the effect of the topology on the number of consensus steps, (ii) extending the approach to nonlinear estimation problems and to systems affected by noises and, (iii) handling delays and packet losses in the underlying consensus process, characterizing the overall effect on the estimation.

References

[1] F. F. Rego, A. M. Pascoal, A. P. Aguiar, and C. N. Jones, "Distributed state estimation for discrete-time linear time invariant systems: s survey," *Annual Reviews in Control*, vol. 48, pp. 36–56, 2019.

[2] C. Deng, C. Wen, J. Huang, X.-M. Zhang, and Y. Zou, "Distributed observer-based cooperative control approach for uncertain nonlinear mass under event-triggered communication," *IEEE Transactions on Automatic Control*, vol. 67, no. 5, pp. 2669–2676, 2021.

[3] B. Açıkmeşe, M. Mandić, and J. L. Speyer, "Decentralized observers with consensus filters for distributed discrete-time linear systems," *Automatica*, vol. 50, no. 4, pp. 1037–1052, 2014.

[4] A. Mitra and S. Sundaram, "Distributed observers for LTI systems," *IEEE Transactions on Automatic Control*, vol. 63, no. 11, pp. 3689–3704, 2018.

[5] L. Wang, J. Liu, A. S. Morse, and B. D. Anderson, "A distributed observer for a discrete-time linear system," in *IEEE Conference on Decision and Control (CDC)*, pp. 367–372, 2019.

[6] L. Wang, J. Liu, and A. S. Morse, "A distributed observer for a continuous-time linear system," in *IEEE American Control Conference (ACC)*, pp. 86–91, 2019.

[7] T. Liu and J. Huang, "Distributed exponential state estimation for discrete-time linear systems over jointly connected switching networks," *IEEE Transactions on Automatic Control*, vol. 68, no. 11, pp. 6836–6843 2023.

[8] A. J. Savas, S. Park, H. V. Poor, and N. E. Leonard, "On separation of distributed estimation and control for LTI systems," in *IEEE Conference on Decision and Control (CDC)*, pp. 963–968, 2022.

[9] P. Duan, Y. Lv, G. Wen, and M. Ogorzałek, "A framework on fully distributed state estimation and cooperative stabilization of LTI plants," *IEEE Transactions on Automatic Control*, pp. 1–16, 2024.

[10] P. Duan, T. Liu, Y. Lv, and G. Wen, "Cooperative control of multi-channel linear systems with self-organizing private agents," *IEEE Transactions on Control of Network Systems*, pp. 1–12, 2024.

[11] C. Fioravanti, E. Makridis, G. Oliva, M. Vrakopoulou, and T. Charalambous, "Distributed estimation and control for LTI systems under finite-time agreement," *IEEE Transactions on Automatic Control*, 2024 (early access).

[12] F. L. Bauer and C. T. Fike, "Norms and exclusion theorems," *Numerische mathematik*, vol. 2, no. 1, pp. 137–141, 1960.

[13] A. D. Dominguez-Garcia and C. N. Hadjicostis, "Distributed algorithms for control of demand response and distributed energy resources," in *IEEE Conference on Decision and Control (CDC) and European Control Conference (ECC)*, pp. 27–32, 2011.

[14] C. N. Hadjicostis, A. D. Domínguez-García, T. Charalambous *et al.*, "Distributed averaging and balancing in network systems: with applications to coordination and control," *Foundations and Trends® in Systems and Control*, vol. 5, no. 2-3, pp. 99–292, 2018.

[15] Y. Lin and J. Liu, "Subgradient-push is of the optimal convergence rate," in *IEEE Conference on Decision and Control (CDC)*, pp. 5849–5856, 2022.

[16] R. Olfati-Saber, J. A. Fax, and R. M. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.