This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3181795

Numerical and Topological Conditions for Sub-optimal Distributed Kalman Filtering*

Osama Ennasr and Xiaobo Tan

***Abstract*—This paper considers the problem of distributed state estimation of a linear time-invariant (LTI) system by a network of sensors in a discrete-time setting. Specifically, we consider a consensus-based distributed Kalman Filter (KF) where each sensor updates its estimate in two steps: a consensus step dictated by a weighted and directed communication graph followed by a local Luenberger filtering step. For a given network, we show that the sub-optimal filtering gains that minimize an upper bound of a quadratic filtering cost can be computed by only exchanging state estimates and their covariance matrices among agents. The resulting dynamics for the network's covariance matrices are represented by a set of coupled algebraic Riccati equations (CARE's) that can be analyzed to ensure stability. Next, we connect some of the results from the Markovian jump linear systems (MJLS) literature regarding the underlying CARE dynamics to the distributed estimation problem, and provide separate necessary and sufficient conditions for successful estimation. We show that the notion of strong detectability (S-detectability) plays an important role in the stability of the filter matrices and ensures that the noise-free error dynamics exponentially converge to zero. We then utilize the notion of weak detectability (W-detectability) to provide necessary conditions in terms of the network topology and self-weights of the communication graph. We further demonstrate how these notions of detectability can be combined to search for feasible network topologies and weights that ensure this sub-optimal performance. Numerical examples are given to illustrate these results.**

I. Introduction

Sensor networks have broad applications in environmental surveillance and monitoring, collaborative information processing, and scientific data gathering from spatially distributed sources for environmental modeling and protection [1]–[3]. In many applications involving large-scale complex systems, the state of the system is monitored by a group of sensors spatially distributed over large networks where the communication between sensors is limited. To model such a scenario, consider the discrete-time linear time-invariant (LTI) dynamical system

$$x(k+1) = Ax(k) + \omega(k) \tag{1}$$

where k is the discrete-time index, $x(k) \in \mathbb{R}^n$ is the state vector, A is the system matrix, and $\omega(k) \in \mathbb{R}^n$ is the driving noise assumed to be Gaussian with zero mean and positive definite covariance Q. The state of the system is monitored by a network of N agents indexed by i, each of which is equipped with a sensor that receives a partial measurement of the state that is modeled by

$$y_i(k) = C_i x(k) + v_i(k), \qquad i = 1, \dots, N, \tag{2}$$

where $y_i(k) \in \mathbb{R}^{p_i}$ is the measurement made by sensor i with C_i as the measurement matrix, and $v_i \in \mathbb{R}^{p_i}$ is the measurement noise assumed to be Gaussian with zero mean and positive definite covariance R_i. An important problem in such networks is to develop distributed algorithms for state estimation of the process in (1), where the goal is for each agent to estimate the entire system state using its respective local measurements and the information obtained from its neighbors.

A. Related Work

This distributed filtering problem has received significant attention over the past two decades. Earlier work in [4], [5] considered the distributed estimation of scalar stochastic systems over general graphs. In these works, it is typically assumed that each node receives scalar local observations, leading to local observability at every node. For more general stochastic systems, a Kalman filter (KF)-based approach has been explored by several researchers. In [6]–[9] consensus filtering is conducted among agents to replicate a centralized KF. These approaches rely on a two-step strategy running on two time scales – a KF-based state estimate update rule on the system's timescale and a data fusion step based on average consensus at a much faster time scale. However, the convergence to the consensus in such approaches requires infinite communication steps between two consecutive local updates, which is far from realizable in any systems.

For that reason, researchers investigated single time-scale algorithms where the data fusion occurs once per time step of plant dynamics [10]–[18]. In [13] and [16], the authors develop single time-scale algorithms for the general case of a directed communication network, where local observability at every node is not necessarily satisfied. In 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. Recent work in [17] tackled the problem of distributed estimation of LTI systems under this general setting as well while alleviating the need for state augmentation and proposed a multi-sensor observable canonical decomposition of the networked system. Each node in the resulting distributed filtering strategy utilizes a Luenberger observer to estimate the modes that are observable using its own measurements, while relying on network-wide consensus to estimate the remaining unobservable modes. In general, to successfully implement this approach, the modes that each agent needs to perform consensus on should be determined during the design stage, which may be difficult to assess when the network topology changes. Furthermore, given a specific network topology and

*This work was supported by the National Science Foundation (ECCS 1446793, IIS 1848945, ECCS 2030556).

O. Ennasr and X. Tan are with the Department of Electrical and Computer Engineering, Michigan State University, East Lansing, MI 48824, USA. Email: {ennasros, xbtan}@msu.edu

weights, the filtering gains that can be obtained using this approach do not offer quantifiable measures to the optimality of the estimates.

The success of the KF in centralized systems, thanks to its optimality and its elegant algorithmic description, has made it a popular basis for distributed estimation algorithms. Several ideas have emerged to extend the Kalman filter to distributed estimation [7]–[11], [15], [18]. The single time-scale variations of the distributed KF, such as the ones proposed in [10], [11], [15], [18], have prompted further investigation of these filters. In [10], a diffusion-based KF is proposed where the covariance matrices from each sensor are updated incrementally by incorporating the observation matrices of the sensor's neighbors in a sequential manner such that only local information is required within the algorithm. However, the local estimates generated at each node do not offer any quantifiable optimality measure for the overall networked system. In [11], the authors propose a consensus-based distributed linear filter, where each sensor updates its estimate in two steps: a local update step using its own observations followed by a consensus step with its neighbors. The work done there offers a certain sub-optimal performance and the authors go on to provide sufficient conditions for the convergence of this distributed filter for a given network topology and weights in terms of the feasibility of a set of linear matrix inequalities (LMIs). While these LMI checks provide an effective way to numerically test whether or not a networked filter will successfully estimate the system state, they do not offer insight into the necessary network connectivity and weights needed for the stability of the networked filter's estimation error. In contrast, the authors in [15] employ structural analysis to derive the network topology requirements for successful estimation. While this structural analysis is useful in specifying the topological connections in the sensor network, it structural nature (i.e., the zero and non-zero structure of entries in the filter matrices) does not offer insight into the role of the consensus weights in distributed filtering. More recent work in distributed Kalman filtering is presented in [18] where the authors utilize covariance intersection for fusing the estimates and covariance matrices for each agent with those of their neighbors. The authors split the communication graph into an observation topology (over which the observations, or measurements, and their matrices are exchanged), and a covariance intersection topology (over which the covariance intersection procedure occurs).

In all of these distributed KF approaches, each agent needs to exchange their local measurements, measurement matrices, noise covariance matrices, or filtering gains in addition to their estimate and estimate covariance matrices with its neighbors in order to successfully estimate the system. This presents a drawback in networks with limited bandwidth as it increases the burden on the communication network to effectively exchange this data.

B. Summary of Contributions

In this work, we propose a consensus-based distributed KF for time-invariant topologies that only requires agents to exchange their estimates and their covariance matrices. For a given network topology and consensus weights, we derive the sub-optimal filtering gains that minimize an upper bound of a quadratic estimation error cost function. Our approach does not require additional information to be exchanged (such as local measurement matrices, gain matrices, noise matrices, or raw measurements) as in [11], [18], and does not require each node to determine which modes to observe directly and which ones to be estimated through consensus as was the case in [17]. We leverage existing results on the underlying CARE equations to derive separate necessary and sufficient conditions for our distributed KF. Combining these sufficient conditions, which can be efficiently checked numerically, and the necessary conditions, which can be verified by inspection, facilitates the search for viable communication topologies and consensus weights that ensure successful estimation.

The remainder of this paper is organized as follows. In Section II we provide the notation and preliminary results used throughout this paper, and discuss the error dynamics of our distributed filter. In Section III we derive the sub-optimal distributed KF gains for a given network that are computed using the CARE of the network. In Section IV we present the separate necessary and sufficient conditions for our distributed KF. A numerical example is given in Section V, where we utilize the aforementioned conditions to effectively search for consensus weights that ensure successful estimation. Concluding remarks are finally provided in Section VI. In the appendices, we provide a proof that was omitted to improve readability, and show that these results hold when the order of the filtering and consensus steps is reversed.

II. Background and Problem Setup

A. Notation

For normed spaces $\mathbb{X}$ and $\mathbb{Y}$, we set $\mathbb{B}(\mathbb{X},\mathbb{Y})$ as the space of all bounded linear operators of $\mathbb{X}$ into $\mathbb{Y}$, and use $\mathbb{B}(\mathbb{X}) \triangleq \mathbb{B}(\mathbb{X},\mathbb{X})$. We denote by $\mathbb{R}^n$ the n-dimensional real Euclidean spaces and $\mathbb{B}(\mathbb{R}^n,\mathbb{R}^m)$ the normed bounded linear space of $m \times n$ real matrices, with the uniform induced-norm represented by $\|.\|$, and use $\mathbb{B}(\mathbb{R}^n) \triangleq \mathbb{B}(\mathbb{R}^n,\mathbb{R}^n)$. We use the notation $null(X)$ to indicate the kernel of a matrix X. The spectral radius of an operator $\mathcal{T} \in \mathbb{B}(\mathbb{X})$ is denoted by $r_\sigma(\mathcal{T})$, while the superscript $'$ indicates the transpose of a matrix, and $\otimes$ represents the Kronecker product. We denote the expected value by $E\{\cdot\}$, and use $\mathrm{tr}(\cdot)$ to indicate the trace.

Throughout this paper, we will work with a collection of N matrices when deriving the sub-optimal filtering gains for our distributed filter. Therefore, it comes up naturally that a convenient space to be used is the one defined by $\mathbb{H}^{n,m}$, which is the linear space made up of all N-sequences of matrices $V = (V_1,\dots,V_N)$ with $V_i \in \mathbb{B}(\mathbb{R}^n,\mathbb{R}^m)$. For simplicity we set $\mathbb{H}^n \triangleq \mathbb{H}^{n,n}$. For $V \in \mathbb{H}^{n,m}$ we write $V' = (V_1',\dots,V_N') \in \mathbb{H}^{m,n}$ and say that $V \in \mathbb{H}^n$ is symmetric if $V = V'$. We set $\mathbb{H}^{n0}$ (respectively, $\mathbb{H}^{n+}$) as the space made up of all N-sequences of symmetric positive semidefinite (respectively, positive definite) matrices. For $V, S \in \mathbb{H}^n$, we write $V \geq S$ if $V - S = (V_1 - S_1,\dots,V_N - S_N) \in \mathbb{H}^{n0}$, and that $V > S$ if $V - S \in \mathbb{H}^{n+}$, i.e., $V_i - S_i > 0$. Finally, for a set of N matrices, $M_i \in \mathbb{B}(\mathbb{R}^n)$, we denote by $\mathrm{diag}[M_i]$ the $Nn \times Nn$ block-diagonal matrix with M_i in the diagonal.

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3181795

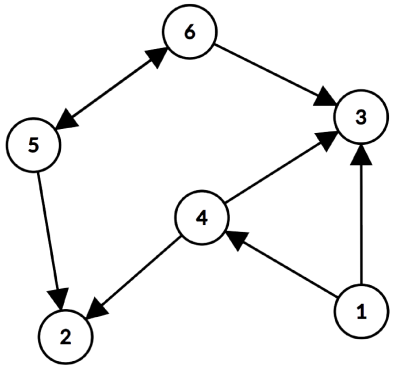


Fig. 1: Example network of 6 agents with source components $\mathbb{S}_1 = \{1\}$ and $\mathbb{S}_2 = \{5, 6\}$.

B. Communication Network

We represent the communication network and consensus weights by $W \in \mathbb{B}(\mathbb{R}^N)$ where elements w_{ij} are nonnegative scalars and the rows sum up to one, i.e.,

$$\sum_{j=1}^{N} w_{ij} = 1, \qquad i = 1, \ldots, N,$$

with $w_{ij} > 0$ if agent i can obtain information from agent j; otherwise, $w_{ij} = 0$.

For a weight matrix W, define recursively,

$$w_{ij}^{(l)} = \sum_{k=1}^{N} w_{ik}^{(l-1)} w_{kj}, \qquad w_{ij}^{(0)} = \delta_{ij}, \tag{3}$$

as the total weight of the directed paths of length l from node j to node i, where $\delta_{ij} = 1$ if $i = j$; otherwise $\delta_{ij} = 0$. We recall the following definition of a source component in the network [17].

Definition 1 (Source component): A set of agents $\mathbb{S} = \{s_1, \ldots, s_p\}$, $\mathbb{S} \subseteq \{1, \ldots, N\}$, is a source component of the network if there exists an integer $l > 0$, such that $w_{ij}^{(l)} > 0$ for any $i, j \in \mathbb{S}$, and for any $i \in \mathbb{S}$ and $j \notin \mathbb{S}$, $w_{ij}^{(l)} = 0$ for all $l > 0$.

Graphically, source components are connected components of the graph with no incoming edges. Figure 1 shows an example network of 6 agents that has two source components, $\mathbb{S}_1 = \{1\}$ and $\mathbb{S}_2 = \{5, 6\}$.

C. Coupled Riccati Equations

For any set of square matrices $S \in \mathbb{H}^n$, we define the operators $\mathcal{E}$ and $\mathcal{L}$ as

$$\mathcal{E}(S) = (\mathcal{E}_1(S), \ldots, \mathcal{E}_N(S)) \in \mathbb{B}(\mathbb{H}^n) \tag{4}$$
$$where \quad \mathcal{E}_i(S) \triangleq \sum_{j=1}^{N} w_{ij} S_j \in \mathbb{B}(\mathbb{R}^n),$$
$$\mathcal{L}(S) = (\mathcal{L}_1(S), \ldots, \mathcal{L}_N(S)) \in \mathbb{B}(\mathbb{H}^n) \tag{5}$$
$$where \quad \mathcal{L}_i(S) \triangleq \Gamma_i \mathcal{E}_i(S) \Gamma_i' \in \mathbb{B}(\mathbb{R}^n),$$

for some $\Gamma_i \in \mathbb{B}(\mathbb{R}^n), \forall i = 1, ..., N$.

Similarly, we define the operators $\mathcal{T}$, $\mathcal{V}$, and $\mathcal{J}$, as

$$\mathcal{T}(S) = (\mathcal{T}_1(S), \ldots, \mathcal{T}_N(S)) \in \mathbb{B}(\mathbb{H}^n) \tag{6}$$
$$where \quad \mathcal{T}_i(S) \triangleq \sum_{j=1}^{N} w_{ji} \Gamma_j' S_j \Gamma_j \in \mathbb{B}(\mathbb{R}^n),$$
$$\mathcal{V}(S) = (\mathcal{V}_1(S), \ldots, \mathcal{V}_N(S)) \in \mathbb{B}(\mathbb{H}^n) \tag{7}$$
$$where \quad \mathcal{V}_i(S) \triangleq \sum_{j=1}^{N} w_{ji} \Gamma_j' S_i \Gamma_j \in \mathbb{B}(\mathbb{R}^n),$$
$$\mathcal{J}(S) = (\mathcal{J}_1(S), \ldots, \mathcal{J}_N(S)) \in \mathbb{B}(\mathbb{H}^n) \tag{8}$$
$$where \quad \mathcal{J}_i(S) \triangleq \sum_{j=1}^{N} w_{ij} \Gamma_j S_j \Gamma_j' \in \mathbb{B}(\mathbb{R}^n).$$

The following result is important in the analysis of the coupled Riccati equations.

Lemma 1 (Proof in [19]): Consider the operators in (5)–(8). We have that

$$r_\sigma(\mathcal{L}) = r_\sigma(\mathcal{V}) = r_\sigma(\mathcal{T}) = r_\sigma(\mathcal{J}) = r_\sigma(\mathcal{A}),$$

where $\mathcal{A} = \text{diag}[\Gamma_i \otimes \Gamma_i](W \otimes I_{n^2}) \in \mathbb{B}\left(\mathbb{R}^{Nn^2}\right)$. Furthermore, the following statements are equivalent and hold when $\mathcal{L}$ is replaced with $\mathcal{T}, \mathcal{V}$, or $\mathcal{J}$:

- $r_\sigma(\mathcal{A}) < 1$.
- There exists some $P \in \mathbb{H}^{n+}$ such that $P - \mathcal{L}(P) > 0$.
- For any $S \in \mathbb{H}^{n+}$, there exists a unique $P \in \mathbb{H}^{n+}$, such that $P - \mathcal{L}(P) = S$.

The CARE's can be expressed in different ways using the operators defined in (4)–(8). In this paper, we will focus on the ones expressed as

$$\begin{aligned} P_i(k+1) =& A\mathcal{E}_i(P(k))A' + Q - A\mathcal{E}_i(P(k))C_i' \times \\ & (C_i \mathcal{E}_i(P(k))C_i' + R_i)^{-1} C_i \mathcal{E}_i(P(k))A'. \end{aligned} \tag{9}$$

These CARE's in (9) have received significant attention in optimal filtering and control of MJLS [19]–[21]. We note the following result from [19] regarding the convergence of these equations.

Lemma 2 (Proof in [19]): The CARE's in (9) converge to the *stabilizing solution*

$$\begin{aligned} P_i =& A\mathcal{E}_i(P)A' + Q - A\mathcal{E}_i(P)C_i' \times \\ & (C_i \mathcal{E}_i(P)C_i' + R_i)^{-1} C_i \mathcal{E}_i(P)A' \end{aligned}$$

only if there exist matrices $L_1, ..., L_N$ such that $r_\sigma(\mathcal{L}) < 1$ with $\Gamma_i = A - L_i C_i$. The stabilizing gains in this case are expressed as

$$L_i = A\mathcal{E}_i(P)C_i'(R_i + C_i \mathcal{E}_i(P)C_i')^{-1}.$$

D. Problem Setup

To solve the problem of estimating (1) by a network of sensors with measurements modeled by (2), we propose a two-step approach to the distributed estimation problem. In the first step, each agent produces an intermediate estimate of the state, $\xi_i(k)$, by performing a convex combination of its own

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3181795

estimate, $\hat{x}_i(k)$, and the state estimates by all other agents within its communication range, i.e.,

$$\xi_i(k) = \sum_{j=1}^{N} w_{ij}\hat{x}_j(k), \quad i = 1, \ldots, N. \tag{10}$$

where w_{ij} is the element of W representing the consensus weight associated with the time-invariant network topology. After all agents complete the consensus step, each agent then updates its estimate by performing a local filtering step based on the intermediate estimate $\xi(k)$, i.e.,

$$\hat{x}_i(k+1) = A\xi_i(k) + L_i\Big(y_i(k) - C_i\xi_i(k)\Big). \tag{11}$$

where L_i is the filter gain. Combining (10) and (11), we obtain the dynamic equation for each agent's estimate

$$\begin{aligned}\hat{x}_i(k+1) =& A\left(\sum_{j=1}^{N} w_{ij}\hat{x}_j(k)\right) \\ &+ L_i\left[y_i(k) - C_i\left(\sum_{j=1}^{N} w_{ij}\hat{x}_j(k)\right)\right].\end{aligned} \tag{12}$$

Letting $\epsilon_i(k) = x(k) - \hat{x}_i(k)$ denote the estimation error for agent i at time k, we obtain the following one-step formulation for the estimation error for each agent

$$\epsilon_i(k+1) = (A - L_iC_i)\sum_{j=1}^{N} w_{ij}\epsilon_j(k) + \omega(k) - L_iv_i(k). \tag{13}$$

Focusing on the noise-free dynamics of the estimation errors in (13) and denoting the noise-free error for agent i by $\varepsilon_i(k)$, we have

$$\varepsilon_i(k+1) = (A - L_iC_i)\sum_{j=1}^{N} w_{ij}\varepsilon_j(k). \tag{14}$$

Defining the network-wide noise-free estimation error as $\varepsilon(k) = [\varepsilon_1(k)' \cdots \varepsilon_N(k)']'$, one can show that the network-wide dynamics for $\varepsilon(k)$ are given by

$$\varepsilon(k+1) = \mathcal{B}\varepsilon(k), \tag{15}$$

where $\mathcal{B} \triangleq \text{diag}[A - L_iC_i](W \otimes I_n) \in \mathbb{B}\left(\mathbb{R}^{Nn}\right)$. Clearly, the system in (15) is asymptotically stable if $r_\sigma(\mathcal{B}) < 1$.

Definition 2 (Distributed detectability): For the system in (1)–(2) and the consensus weights W, we say (A, C, W) is detectable in the distributed sense if there exist gain matrices L_i, $i = 1, ..., N$, such that $r_\sigma(\mathcal{B}) < 1$.

III. Sub-optimal Distributed Kalman Filtering

To find the optimal filtering gains for this distributed problem, we introduce the following finite horizon quadratic filtering cost function

$$J(k) = \sum_{t=0}^{k}\sum_{i=1}^{N} E\{\|\epsilon_i(t)\|^2\}. \tag{16}$$

Ideally, we would like to obtain the optimal filtering gains that minimize the cost function in (16). However, due to the complexity of this optimization problem, we focus on providing filtering gains that minimize an upper bound of the cost function in (16). This, in turn, leads us to a sub-optimal distributed filtering scheme with quantifiable performance.

Lemma 3: Define $\Xi(k) = (\Xi_1(k), \ldots, \Xi_N(k)) \in \mathbb{H}^{n0}$, where $\Xi_i(k) = \varepsilon_i(k)\varepsilon_i(k)'$. Let $P(k) \in \mathbb{H}^{n0}$ be defined recursively as

$$P(k+1) = \mathcal{L}(P(k)), \qquad P(0) = \Xi(0), \tag{17}$$

with $\Gamma_i = (A - L_iC_i)$. Then,

$$\Xi(k) \le P(k), \qquad \forall k \ge 0.$$

Furthermore, define $\Sigma(k) \in \mathbb{H}^{n+}$ and $P(k) \in \mathbb{H}^{n+}$, where $\Sigma_i(k) = E\{\epsilon_i(k)\epsilon_i(k)'\}$ denotes the covariance of agent i's estimation error with noise in (13), and $P_i(k)$ is defined recursively as

$$P_i(k+1) = \mathcal{L}_i(P(k)) + Q + L_iR_iL_i', \quad P(0) = \Sigma(0), \tag{18}$$

with $\Gamma_i = (A - L_iC_i)$, R_i being the covariance matrix of the measurement noise for sensor i, and Q being the covariance matrix of the process noise. Then,

$$\Sigma(k) \le P(k), \qquad \forall k \ge 0.$$

Proof: See Appendix A. ■

Lemma 3 provides an upper bound for each agent's covariance matrix. This in turn allows us to establish an upper bound, $\bar{J}(k)$, for the cost function $J(k)$ in (16), where

$$\bar{J}(k) = \sum_{t=0}^{k}\sum_{i=1}^{N} \text{tr}(P_i(t)), \tag{19}$$

with $P_i(\cdot)$ defined in (18). From Lemma 3, it immediately follows that

$$J(k) \le \bar{J}(k).$$

The following theorem establishes the filtering gains that minimize the upper bound shown in (19) for our distributed filtering scheme.

Theorem 1: For a given matrix of consensus weights, W, the filtering gains that minimize the upper bound in (19) can be computed using the same network topology and weights in a two-step approach. During the consensus step, each agent computes an intermediate covariance matrix by performing a convex combination of their own $P_i(k)$ and those of their neighbors, i.e.,

$$\Pi_i(k) = \sum_{j=1}^{N} w_{ij}P_j(k). \tag{20}$$

Then, during the filtering step, each agent computes the local filtering gain to be used in (11) using the intermediate covariance

$$L_i(k) = A\Pi_i(k)C_i'(R_i + C_i\Pi_i(k)C_i')^{-1}, \tag{21}$$

and updates its covariance upper bound according to

$$\begin{aligned}P_i(k+1) =& A\Pi_i(k)A' + Q - A\Pi_i(k)C_i' \times \\ &(C_i\Pi_i(k)C_i' + R_i)^{-1}C_i\Pi_i(k)A'.\end{aligned} \tag{22}$$

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3181795

Proof: First, we show that the optimal cost $\bar{J}^*(k)$ is obtained using the gains

$$L_i^*(k) = A\mathcal{E}_i(P^*(k))C_i'(R_i + C_i\mathcal{E}_i(P^*(k))C_i')^{-1}, \quad (23)$$

with $P^*(0) = \Sigma(0)$, and $P^*(k) \in \mathbb{H}^{n+}$ is computed using (18) with the gains $L_i^*(k)$.

Let $\bar{J}(k)$ denote the cost when an arbitrary set of gains $L_i(k)$ is used, and let $P(k) \in \mathbb{H}^{n+}$ denote the matrices obtained when these arbitrary gains are used in (18). To show the desired result, we need to show that

$$\bar{J}(k) - \bar{J}^*(k) = \sum_{t=0}^{k}\sum_{i=1}^{N} \mathrm{tr}(P_i(t) - P_i^*(t)) \geq 0.$$

We do that by induction. For $P(0) \geq P^*(0)$, it can be shown through matrix manipulation (see [19]) that

$$\begin{aligned} &{P_i}^*(1) - P_i(1) \\ &\quad = (A - L_i^*C_i)\mathcal{E}_i(P^*(0) - P(0))(A - L_i^*C_i)' \\ &\quad\quad - (L_i - L_i^*)(R_i + C_i\mathcal{E}_i(P(0))C_i')(L_i - L_i^*). \end{aligned}$$

Since $P(0) \geq P^*(0)$ and $(R_i + C_i\mathcal{E}_i(P(0))C_i') > 0$, it immediately follows that $P_i(1) \geq P_i^*(1)$ and $\bar{J}(k) - \bar{J}^*(k) \geq 0$. Now, let $P(k) \geq P^*(k)$. Utilizing the same matrix manipulation as above, since $(R_i + C_i\mathcal{E}_i(P(k))C_i') > 0$, it follows that $P_i(k+1) \geq P_i^*(k+1)$ for all $k \geq 0$. Finally, we show that these sub-optimal gains and the matrices $P^*(k)$ can be computed in the distributed manner shown in (20)–(22). Indeed, substituting $L_i^*(k)$ in (18), and after performing additional matrix manipulation, we can write

$$\begin{aligned} P_i^*(k+1) &= A\mathcal{E}_i(P^*(k))A' + Q - A\mathcal{E}_i(P^*(k))C_i' \times \\ &\quad (C_i\mathcal{E}_i(P^*(k))C_i' + R_i)^{-1}C_i\mathcal{E}_i(P^*(k))A', \quad (24) \end{aligned}$$

which is the same expression obtained when substituting (20) into (21) and (22). ■

Substituting $\Pi_i(k)$ from (20) into (22) yields the set of CARE's in (9). We refer to the filter resulting from (10)–(11) and (20)–(22) as the *distributed KF* – due to its similarity to the centralized Kalman filter – and it presents an attractive solution to the distributed filtering problem as it provides a sub-optimal scheme with quantifiable performance.

IV. Conditions for Stability of Estimation Error

For a centralized estimation problem, the detectability of the pair (A, C) plays an important role in the convergence the centralized KF. Therefore, it is natural for us to seek a similar notion for the distributed problem. In this section, we employ some of the results reported in [19]–[21] to provide sufficient and necessary conditions for the stability of the estimation errors under this distributed KF. Specifically, we make use of the notions of *mean square detectability (MS-detectability)*, i.e., relating to the second moment of the underlying random variable, and *weak detectability (W-detectability)* [19], [21], to introduce similar notions for our distributed filtering problem. For a complete discussion on MS-detectability and relevant results in MJLS, we refer the reader to [19] and references therein. Roughly speaking, MS-detectability requires the existence of filtering gains that result in a contraction of an operator on the space of a *set of matrices*, and can be verified using LMI-based techniques [19]. On the other hand, W-detectability does not require the contraction property imposed by the MS-detectability [20], and instead deals with the observable subspaces for each agent, resulting in topological conditions that can be verified by inspection.

In what follows, we introduce the notion of *strong detectability (S-detectability)* for our distributed filtering problem, an analogy to the MS-detectability in MJLS systems. S-detectability ensures that the upper bounds of the error covariance matrices can be stabilized and remain bounded, which we use to derive conditions on some of the network's consensus weights. We then consider the W-detectability property for our distributed system, and utilize the results in [21] and [20] to arrive at the network topologies necessary for the success of this distributed filter.

A. Sufficient Conditions

Theorem 2: If there exist gain matrices $L_1, \ldots, L_N$, and symmetric positive definite matrices $P \in \mathbb{H}^{n+}$ and $S \in \mathbb{H}^{n+}$ such that

$$P - \mathcal{T}(P) = S, \quad (25)$$

with $\Gamma_i = A - L_iC_i$, then the noise-free error dynamics in (14) are exponentially stable. This holds when $\mathcal{T}$ is replaced by $\mathcal{L}, \mathcal{V}$, or $\mathcal{J}$.

Proof: To show the desired result, we define the Lyapunov function

$$V(k) = \sum_{i=1}^{N} \varepsilon_i(k)'P_i\varepsilon_i(k).$$

It follows that,

$$\begin{aligned} &V(k+1) - V(k) \\ &= \sum_{i=1}^{N}\left(\sum_{j=1}^{N} w_{ij}\varepsilon_j(k)'\right)\Gamma_i'P_i\Gamma_i\left(\sum_{j=1}^{N} w_{ij}\varepsilon_j(k)\right) \\ &\quad - \varepsilon_i(k)'P_i\varepsilon_i(k) \\ &\leq \sum_{i=1}^{N}\left(\left(\sum_{j=1}^{N} w_{ij}\varepsilon_j(k)'\Gamma_i'P_i\Gamma_i\varepsilon_j(k)\right) - \varepsilon_i(k)'P_i\varepsilon_i(k)\right). \end{aligned}$$

Reversing the order of the sums, we have

$$\begin{aligned} &V(k+1) - V(k) \\ &\quad\quad \leq \sum_{i=1}^{N} \varepsilon_i(k)'\left(\sum_{j=1}^{N} w_{ji}\Gamma_j'P_j\Gamma_j - P_i\right)\varepsilon_i(k), \end{aligned}$$

and using $\mathcal{T}$ in (6) yields

$$V(k+1) - V(k) \leq \sum_{i=1}^{N} \varepsilon_i(k)'\left(\mathcal{T}_i(P) - P_i\right)\varepsilon_i(k).$$

From (25), we have

$$P_i - \mathcal{T}_i(P) = S_i, \qquad i = 1, \ldots, N,$$

and thus

$$V(k+1)-V(k) \leq -\sum_{i=1}^{N} \varepsilon_i(k)' S_i \varepsilon_i(k).$$

From the fact that $S \in \mathbb{H}^{n+}$, we get $V(k+1)-V(k) < 0$ and the error dynamics in (14) are exponentially stable. From Lemma 1, it follows that this result holds when $\mathcal{T}$ is replaced by $\mathcal{L}, \mathcal{V}$, or $\mathcal{J}$. ■

From Lemma 1, it is clear that there exist matrices $P \in \mathbb{H}^{n+}$ and $S \in \mathbb{H}^{n+}$ such that (25) holds *iff* there exist gain matrices $L_1, \dots, L_N$ such that $r_\sigma(\mathcal{L}) < 1$ with $\Gamma_i = A - L_i C_i$. This gives rise to the following stronger notion of detectability for our distributed filter based on the operators defined in (5)–(8).

Definition 3 (S-detectability): For the system in (1)–(2) and the consensus weights W, we say (A, C, W) is S-detectable if there exists a set of gain matrices L_i, $i = 1, ..., N$, such that $r_\sigma(\mathcal{L}) < 1$ with $\Gamma_i = A - L_i C_i$.

Corollary 1: If (A, C, W) is S-detectable, then (A, C, W) is detectable in the sense of Definition 2.

Proof: The result follows immediately from Lemma 1 and Theorem 2. ■

Generally speaking, the reverse is not necessarily true. That is, finding gain matrices such that $r_\sigma(\mathcal{B}) < 1$ does not imply that $r_\sigma(\mathcal{L}) < 1$. Indeed, consider the example of a scalar system monitored by two agents, where $\Gamma_1 = 1.3$ and $\Gamma_2 = 0.75$, with consensus weights

$$W = \begin{bmatrix} 0.1 & 0.9 \\ 0.1 & 0.9 \end{bmatrix},$$

It can be verified that $r_\sigma(\mathcal{B}) = 0.975$ while $r_\sigma(\mathcal{L}) = 1.04$.

The interchangeability between the operators allows us to check for S-detectability using any of the given operators using LMI techniques such as the ones shown in [11] and [19]. Specifically, we show that the following numerical check can determine S-detectability of (A, C, W).

Proposition 1: (A, C, W) is S-detectable if there exist matrices $Y_1, ..., Y_N$ and $X \in \mathbb{H}^{n+}$ such that the following set of N LMIs are feasible for all $i = 1, \dots, N$

$$\begin{bmatrix} M_{i,11} & M_{i,12} \\ M_{i,12}' & M_{i,22} \end{bmatrix} > 0, \tag{26}$$

where

$$M_{i,11} = X_i, \qquad M_{i,22} = \text{diag}[X_i]$$

and

$$M_{i,12}' = \begin{bmatrix} \sqrt{w_{1i}}(X_1 A - Y_1 C_1) \\ \vdots \\ \sqrt{w_{ji}}(X_j A - Y_j C_j) \\ \vdots \\ \sqrt{w_{Ni}}(X_N A - Y_N C_N) \end{bmatrix}.$$

Proof: Applying the Schur complement lemma, the LMIs in (26) are feasible if for $i = 1, \dots, N$

$$X_i - \sum_{j=1}^{N} w_{ji}(X_j A - Y_j C_j)' X_j^{-1} (X_j A - Y_j C_j) > 0, \quad X_i > 0.$$

Grouping terms and using the definition of $\mathcal{T}$ in (6) yields

$$X - \mathcal{T}(X) > 0$$

with $\Gamma_i = A - X_i^{-1} Y_i C_i$. The result then follows from Theorem 2 and Lemma 1. ■

Note that while S-detectability is sufficient for exponential stability of the noise-free error dynamics in (14), it is only necessary for the convergence of the CARE's to a stabilizing solution. The full necessary and sufficient conditions for the convergence of the CARE's can be found in [19], but they are difficult to verify and are beyond the scope of this paper. Nevertheless, it has been shown that if (A, C, W) and $(A', Q^{\frac{1}{2}}, W')$ are S-detectable, then the CARE's converge to a unique stabilizing solution [19]. An LMI-based approach similar to the one used in Proposition 1 can be employed to check for S-detectability of $(A', Q^{\frac{1}{2}}, W')$.

B. Necessary Conditions

An important aspect of this distributed filtering problem is to find a matrix of consensus weights W that satisfies the S-detectability condition, as the LMI-based approach does not offer much insight into the topological conditions of the network. It is important, therefore, to investigate the conditions required for S-detectability, as they are required for the filter covariance matrices to remain bounded and ensure that the noise-free dynamics are exponentially stable. The mathematical description of S-detectability mimics the notion of MS-detectability in MJLS, allowing us to leverage some of the results there in analyzing the conditions on the consensus weights and network connectivity needed for the stability of the estimation errors. The remainder of this section is aimed at providing the necessary conditions for S-detectability of (A, C, W) in terms of network topology and communication weights.

Before presenting the main result, however, we review some of the relevant results from MJLS. Specifically, we make use of the notion of W-detectability, which has been shown to play an important role in the LQR control problem of MJLS [20], [21]. W-detectability is a weaker condition than S-detectability in that it does not require the existence of gain matrices that result in $r_\sigma(\mathcal{L}) < 1$. It has been shown that W-detectability is necessary for MS-detectability in MJLS [20], and is therefore necessary for S-detectability in this distributed estimation problem. The appeal for W-detectability is that it offers an easier way to check, compared to S-detectability, and allows us to consider the minimum network topology needed for S-detectability. Some of the results listed in this subsection require many of the auxiliary results developed in [21] and [20]. To keep the discussion concise, we refer the reader to those references for the full proofs of such results.

To formally define W-detectability in the context of distributed filtering, we consider the system

$$X(k+1) = \mathcal{T}(X(k)), \quad X_i(0) = \mu_i x(0) x(0)', \tag{27}$$

This article has been accepted for publication in IEEE Transactions on Control of Network Systems. This is the author's version which has not been fully edited and content may change prior to final publication. Citation information: DOI 10.1109/TCNS.2022.3181795

where μ_i are nonnegative scalars such that $\sum_i \mu_i = 1$ with $\Gamma_i = A$, and introduce the functional

$$\mathcal{W}_i^k(X) = \sum_{t=0}^{k-1} \text{tr}(X_i(t)C_i'C_i), \tag{28}$$

and define

$$\mathcal{W}^k(X) = \sum_{i=1}^{N} \mathcal{W}_i^k(X). \tag{29}$$

The following definition of W-detectability is adapted from the one presented in [20].

Definition 4 (W-detectability): Consider the system in (27). We say (A, C, W) is W-detectable when there exist integers $k_1, k_2 \geq 0$ and scalars $0 \leq \delta < 1, \gamma > 0$ such that $\mathcal{W}^{k_1}(X) \geq \gamma \|X(0)\|$ whenever $\|X(k_2)\| \geq \delta \|X(0)\|$, with $\|X\|$ defined as

$$\|X(k)\|^2 = \sum_{j=1}^{N} \text{tr}(X_i(k)'X_i(k)).$$

For the sequence $O(k) \in \mathbb{H}^{n0}$, defined recursively as

$$O_i(k+1) = C_i'C_i + A'\mathcal{E}_i(O(k))A, \quad O_i(0) = 0, \tag{30}$$

it has been shown (see [22]) that for the system in (27), $\mathcal{W}^k(X)$ can be expressed as

$$\mathcal{W}^k(X) = \sum_{i=1}^{N} \mu_i x(0)'O_i(k)x(0). \tag{31}$$

Note that the expression inside the sum in (31) can be expressed as

$$x(0)'O_i(k)x(0) = \sum_{t=0}^{k-1} \sum_{j=1}^{N} w_{ij}^{(t)} \|y_j\|^2,$$

which represents the total output energy available at node i at time k after getting diffused by the network. The following property for W-detectability is proven in [20].

Lemma 4: Consider the system in (27). (A, C, W) is W-detectable *iff* whenever $\mathcal{W}^{n^2N}(X) = 0$, one has that $\|X(k)\| \to 0$ as $k \to \infty$ for any scalars μ_i satisfying $\sum_i \mu_i = 1$.

We define the set of observability matrices $\mathcal{O} = \{\mathcal{O}_1, \cdots, \mathcal{O}_N\}$, where

$$\mathcal{O}_i = \begin{bmatrix} \mathcal{O}_i(0) & \cdots & \mathcal{O}_i(n^2N-1) \end{bmatrix}' \tag{32}$$

for each $i \in \{1, \ldots, N\}$, and the matrices $\mathcal{O}_i(k)$ are defined recursively as

$$\mathcal{O}_i(k+1) = A_i'\mathcal{E}_i(\mathcal{O}(k))A_i, \qquad \mathcal{O}_i(0) = C_i'C_i. \tag{33}$$

The following result relates the notion of W-detectability to that of the kernel of the observability matrices $\mathcal{O}_i$.

Lemma 5: (A, C, W) is W-detectable *iff* for some $x(0) \in null(\mathcal{O}_i)$, for any $i \in \{1, \ldots, N\}$, we have that $\lim_{k\to\infty} \|x(k)\|^2 = 0$, and the pair $(A, \mathcal{O}_i)$ is detectable.

Proof: First, we show that if $x(0) \in null(\mathcal{O}_i)$, for any $i \in \{1, \ldots, N\}$, then there exist scalars $\mu_i \geq 0$ satisfying $\sum_i \mu_i = 1$ such that $\mathcal{W}^{n^2N}(X) = 0$. Indeed, if $x(0) \in null(\mathcal{O}_s)$ for some $s \in \{1, \ldots, N\}$, it immediately follows from the definition of (30) and (32) that $x(0)'O_s(n^2N)x(0) = 0$. Finally, setting the the scalaras $\mu_i = \delta_{is}$ ($\delta_{is} = 1$ if $i = s$; otherwise $\delta_{is} = 0$) results in $\mathcal{W}^{n^2N}(X) = 0$. The proof is completed by applying Lemma 4 and noting that if $\|X(k)\| \to 0$ then $\|x(k)\| \to 0$. ■

We also utilize the following lemma that was presented in [20] to establish the relationship between S-detectability and W-detectability. Since the proof requires many of the preliminary results developed in [20], we encourage the reader to explore this reference for a complete proof of Lemma 6.

Lemma 6: If (A, C, W) is S-detectable, then (A, C, W) is W-detectable.

Proof: The statement immediately follows from [20], after noting the similarity between the notions of S-detectability in this distributed filtering problem and the MS-detectability in MJLS. ■

Finally, we provide the following definition regarding source components of the network.

Definition 5 (Detectable source component): A source component $\mathbb{S} = \{s_1, \ldots, s_p\}$ is detectable if the pair $(A, C_\mathbb{S})$ is detectable, where $C_\mathbb{S} = \begin{bmatrix} C_{s_1}' & \ldots & C_{s_p}' \end{bmatrix}'$.

In the example graph shown in Figure 1, the source components $\mathbb{S}_1$ and $\mathbb{S}_2$ are detectable if (A, C_1) and $(A, [C_5'C_6']')$ are detectable, respectively.

We now present the following necessary conditions for S-detectability.

Theorem 3: If (A, C, W) is S-detectable, then:

1) Every source component in the network is detectable.
2) $w_{ii} < (1/r_\sigma(A_i^u))^2$ for each $i = 1, \ldots, N$, where A_i^u is the unobservable partition of A using C_i.

Proof: First, we show that the first statement holds. To that end, if (A, C, W) is S-detectable, then it is necessarily W-detectable, and from Lemma 5 we have that $(A, \mathcal{O}_i)$ is detectable for each i. It is easy to check that $\mathcal{O}_i(k)$ in (33) can be written as

$$\mathcal{O}_i(k) = \sum_{j=1}^{N} w_{ij}^{(k)} A'^k C_j' C_j A^k.$$

Let $\mathbb{S}$ be a source component, and denote by $\mathcal{O}_\mathbb{S}$ the observability matrix corresponding to the pair $(A, \mathbb{C}_S)$. It is known that each agent $i \in \{1, \ldots, N\}$ is either part of a source component, or there exists a directed path to i from some agent j in a source component.

On the one hand, for some agent $i \in \mathbb{S}$, it follows from Definition 1 that there exists an integer $l > 0$ such that $w_{ij}^{(l)} > 0$ for every $j \in \mathbb{S}$. It can then be checked that for any $i \in \mathbb{S}$, $null(\mathcal{O}_i) = null(\mathcal{O}_\mathbb{S})$, and $\mathbb{S}$ is a detectable source component since for any $x(0) \in null(\mathcal{O}_\mathbb{S})$, $\lim_{k\to\infty} \|x(k)\|^2 = 0$.

On the other hand, if $i \notin \mathbb{S}$, then there is no directed path from i to $j \in \mathbb{S}$, and $null(\mathcal{O}_i) \subseteq null(\mathcal{O}_\mathbb{S})$. However, since $(A, \mathcal{O}_i)$ and $(A, \mathcal{O}_j)$ are detectable, with $null(\mathcal{O}_j) = null(\mathcal{O}_\mathbb{S})$, then $\mathbb{S}$ is a detectable source component.

Now we show that the second statement holds. Indeed, if (A, C, W) is S-detectable, then there exist gain matrices L_i, $i = 1, \ldots, N$, such that $r_\sigma(\mathcal{L}) < 1$. From Lemma 1, this implies that $P - \mathcal{L}(P) > 0$ for some positive definite matrices

$P \in \mathbb{H}^{n+}$. Following the same approach as in [23], we have that

$$P_i - (A - L_iC_i)' \sum_{j=1}^{N} w_{ij} P_j (A - L_iC_i) > 0,$$
$$\Rightarrow P_i - (\sqrt{w_{ii}})(A - L_iC_i)' P_i (\sqrt{w_{ii}})(A - L_iC_i) > 0,$$

for $i = 1, \ldots, N$, and therefore $(\sqrt{w_{ii}}A, C_i)$ is detectable. We can then find a similarity transformation to transform the pair $(\sqrt{w_{ii}}A, C_i)$ into

$$\bar{A}_i = T_i\,(\sqrt{w_{ii}}A)\,T_i' = \sqrt{w_{ii}} \begin{bmatrix} A_i^o & 0 \\ A_i^c & A_i^u \end{bmatrix},$$
$$\bar{C}_i = C_i T_i' = \begin{bmatrix} C_i^o & 0 \end{bmatrix},$$

where $\sqrt{w_{ii}}A_i^u$ represents the unobservable modes. Clearly, if $(\sqrt{w_{ii}}A, C_i)$ is detectable then $\sqrt{w_{ii}}A_i^u$ must be stable, and $w_{ii} < (1/r_\sigma(A_i^u))^2$. ∎

These necessary conditions allow us to consider the weakest communication topologies that are needed for S-detectability, as well as the maximum values for the *self-weights* used in consensus. Note that Theorem 3 does not require the overall network to be connected as long as each source component of the network is detectable in the sense of Definition 5. Practically, these necessary conditions allow us to analyze whether or not a given network prohibits our distributed KF from successfully estimating the system by violating the requirements in Theorem 3. Moreover, these conditions facilitate the search for a viable W that renders (A, C, W) S-detectable, and ensures that the noise-free error dynamics are stable.

V. Numerical Examples

We consider a two-dimensional system (i.e., $n = 2$) with

$$A = \begin{bmatrix} 2 & 1 \\ -0.5 & 4 \end{bmatrix},$$

where the eigenvalues of A are $\lambda_1 = 2.2929$ and $\lambda_2 = 3.7071$, and the system is unstable. This system is monitored by $N = 10$ sensors, and we denote by N_{odd} (respectively, N_{even}) the set of odd (respectively, even) numbered agents. Let $C_i = v_1$ for sensors $i \in N_{odd}$ and $C_i = v_2$ for $i \in N_{even}$, where v_1 and v_2 are the left eigenvectors corresponding to λ_1 and λ_2.

Clearly, no pair (A, C_i) is detectable on its own, implying that all agents require additional information to satisfy the conditions for S-detectability. Note that the spectral radius of the unobservable partitions for odd-numbered agents is λ_2, and for even-numbered agents it is λ_1. Therefore, we require that $w_{ii} < 1/\lambda_2^2 = 0.073$ for odd i and $w_{ii} < 1/\lambda_1^2 = 0.19$ for even i.

In order to find a network topology and weights that renders the networked system S-detectable, we first identify the network topologies that ensure W-detectability for the network. For this example, we can find several different topologies that satisfy the conditions in Theorem 3. Figure 2 shows some of these possible configurations. Note that the overall network does not need to be connected as long as every source component is detectable.

To show the importance of the values of self weights w_{ii}, we consider the cycle topology $\mathcal{G}_1$ shown in Figure 2(a) since the diagonal entries w_{ii} indirectly impose that the remaining consensus weight for each agent is $1 - w_{ii}$. We simulate the response of a distributed Kalman filter using (20)–(22) with the filtering gains shown in (21) using the same cycle topology, but with different w_{ii} values. The off-diagonal entries of W representing the incoming edge for each agent in $\mathcal{G}_1$ are set to $1 - w_{ii}$. Specifically, we consider four sets of consensus weight matrices, where

$$W_a : \begin{cases} w_{ii} = 0.03, i \text{ odd} \\ w_{ii} = 0.05, i \text{ even} \end{cases}, \quad W_b : \begin{cases} w_{ii} = 0.06, i \text{ odd} \\ w_{ii} = 0.18, i \text{ even} \end{cases}$$
$$W_c : \begin{cases} w_{ii} = 0.08, i \text{ odd} \\ w_{ii} = 0.20, i \text{ even} \end{cases}, \quad W_d : \begin{cases} w_{ii} = 0.12, i \text{ odd} \\ w_{ii} = 0.25, i \text{ even} \end{cases}$$

The simulations were initialized with $x(0) = \begin{bmatrix} -15 & 15 \end{bmatrix}'$, $\hat{x}_i(0) = \begin{bmatrix} 0 & 0 \end{bmatrix}'$, $P_i(0) = \epsilon_i(0)\epsilon_i(0)'$, $R_i = 10^{-3}$, and $Q = 10^{-3}I$, where I is the 2-by-2 identity matrix. Figure 3 shows the sum of the traces of the filter's error covariance matrices, which are the upper bounds of the network-wide error norms, for various consensus weights. Figure 3 shows the convergence behavior of the covariance matrices as the self-weights vary under the same communication topology, and highlights the importance of the consensus weights when running the distributed KF. These results suggest a slower divergence or convergence behavior as the self weights approach their cutoff limits, with quickest divergence observed for W_d, and the fastest convergence observed with W_a.

Finally, we utilize Proposition 1 to search for a set of consensus weights consistent with the other topologies in Figure 2 that satisfy the S-detectability condition. For example, searching for consensus weights consistent with the topology in Figure 2(b), one can find that the following consensus weights

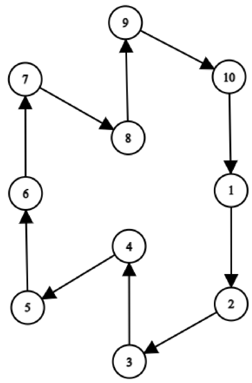
$$W = \begin{bmatrix} W_1 & 0 & 0 & 0 & 0 \\ 0 & W_2 & 0 & 0 & 0 \\ 0 & 0 & W_3 & 0 & 0 \\ 0 & 0 & 0 & W_4 & 0 \\ 0 & 0 & 0 & 0 & W_5 \end{bmatrix},$$

satisfy the S-detectability condition, where each W_i describes the weights for each source component, with

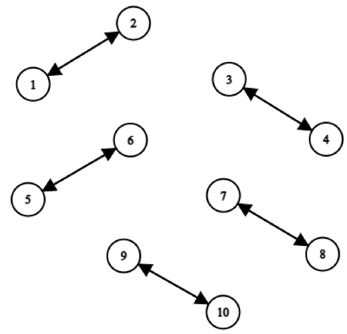
$$W_i = \begin{bmatrix} 0.02 & 0.98 \\ 0.9 & 0.1 \end{bmatrix}.$$

A similar search can be done for the topology shown in Figure 2(c), resulting in the consensus weights

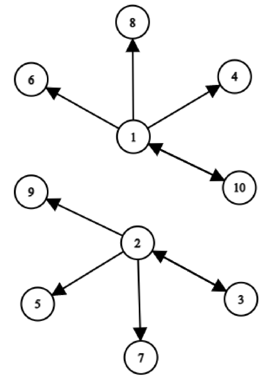
$$W = \begin{bmatrix} 0.05 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.95 \\ 0 & 0.1 & 0.9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.95 & 0.05 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.9 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.1 \end{bmatrix}.$$



(a) $\mathcal{G}_1$.



(b) $\mathcal{G}_2$.



(c) $\mathcal{G}_3$.

Fig. 2: Possible communication topologies that satisfy the W-detectability conditions.

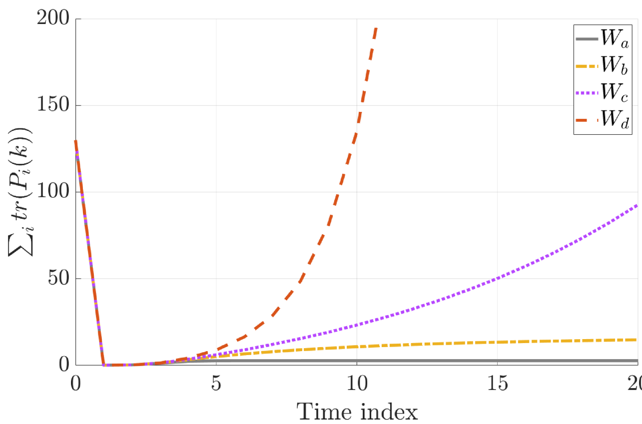


Fig. 3: Comparison of the sum of the filter covariance matrices for different consensus weights using the same network topology $\mathcal{G}_1$.

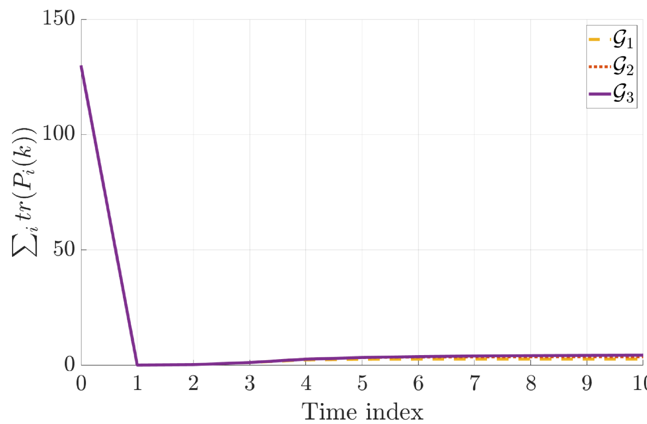


Fig. 4: Comparison of the convergence of the sum of the filter covariance matrices under different network topologies.

Figure 4 shows the traces of the covariance upper bounds for some of the found weights for the three topologies shown in Figure 2. Indeed, it is clear from inspecting the traces of the topology that the network does not need to be connected for our distributed KF to accurately estimate the system's state.

VI. Conclusion and Future Work

In this paper, we consider the problem of distributed estimation of an LTI system by a network of sensors, where the agents update their estimates by performing consensus followed by local filtering. We present a distributed Kalman filter, in which the filtering gains minimize an upper bound of a quadratic filtering cost. These filtering gains are computed using a set of coupled Riccati equations that, in turn, can be updated in a distributed manner under the same network, and the approach requires agents to only exchange their estimates and their covariance matrices. We then provide the notion of S-detectability, which ensures exponential stability of the noise-free estimation errors, and provide a numerical method for checking it. We then build on that notion to provide necessary conditions in terms of the minimum network connectivity and a limit on the self-weights used during the consensus step.

Future work will consider the numerical stability of the Riccati equations and quantify how close these upper bounds are to the true error values. We will also focus on the off-diagonal consensus weights for the network and analyze the consensus weights that guarantee S-detectability of the network to complement the numerical checks. Finally, we will consider extending these results to time-varying network topologies.

Appendix

In this section, we provide the proof for Lemma 3 and show that our results hold when the order of the filtering steps is reversed. We make use of the following remark.

Remark 1 (From [11]): Given a set of N nonnegative scalars s_i summing up to one, a set of N vectors x_i, and a set of N matrices A_i, the following holds

$$\left(\sum_{i=1}^{N} s_i A_i x_i\right)\left(\sum_{i=1}^{N} s_i A_i x_i\right)' \leq \sum_{i=1}^{N} s_i A_i x_i x_i' A_i'.$$

A. Proof of Lemma 3

For the noise-free case, we note that the matrix $\Xi_i(k+1)$ can be explicitly written as

$$\Xi_i(k+1) = \left((A - L_iC_i)\sum_{j=1}^{N} w_{ij}\epsilon_j\right) \times \left((A - L_iC_i)\sum_{j=1}^{N} w_{ij}\epsilon_j\right)'.$$

Using Remark 1, it follows that

$$\Xi_i(k+1) \leq (A - L_iC_i)\sum_{j=1}^{N} w_{ij}\Xi_j(k)(A - L_iC_i)'.$$

Defining $\Xi(k) \in \mathbb{H}^{n0}$, with $\Xi_i(k)$ shown above, we have

$$\Xi(k+1) \leq \mathcal{L}(\Xi(k))$$

with $\Gamma_i = (A - L_iC_i)$, and the first statement can then be proved by induction.

Assuming that the initial state $x(0)$, and the noises $v_i(k)$ and $\omega(k)$ are independent for all $k \geq 0$, the second statement can be proved in a similar manner. Since the noises have zero mean and are independent with respect to themselves and $x(0)$, using Remark 1 it can be checked that

$$\begin{aligned}\Sigma_i(k+1) &= E\left\{\epsilon_i(k+1)\epsilon_i(k+1)'\right\} \\ &\leq \mathcal{L}_i(\Sigma(k)) + Q + L_iR_iL_i',\end{aligned}$$

with $\Sigma(k) = (\Sigma_1(k), \ldots, \Sigma_N(k)) \in \mathbb{H}^{n+}$ and $\Gamma_i = (A - L_iC_i)$. The second statement can then be proved by induction.

B. Reversing Order of Consensus and Filtering Steps

It is possible to derive an alternate version of this distributed KF by reversing the order of which the consensus and filtering steps are taken, resulting in the distributed filter described in [11], and leading to a different form of the CARE's. In this case, the noise-free estimation error dynamics are given by

$$\varepsilon_i(k+1) = \sum_{j=1}^{N} w_{ij}\left(A - L_jC_j\right)\varepsilon_j(k). \tag{34}$$

First we show that the dynamics in (34) are asymptotically stable *iff* the dynamics in (13) are asymptotically stable. From (34), the network-wide error dynamics can be expressed as

$$\varepsilon(k+1) = \mathcal{B}_2\varepsilon(k), \tag{35}$$

where $\mathcal{B}_2 \triangleq (W \otimes I_n)\text{diag}[A - L_iC_i] \in \mathbb{B}\left(\mathbb{R}^{Nn}\right)$. Since $(W \otimes I_n) \in \mathbb{B}\left(\mathbb{R}^{Nn}\right)$ and $\text{diag}[A - L_iC_i] \in \mathbb{B}\left(\mathbb{R}^{Nn}\right)$, it follows that $r_\sigma(\mathcal{B}) = r_\sigma(\mathcal{B}_2)$ and (15) is asymptotically stable *iff* (35) is asymptotically stable.

The notion of S-detectability in Definition 3 implies that there exist certain gain matrices such that the dynamics of the upper bounds for the noise-free covariances in (17) are asymptotically stable. To show that notion holds regardless of the order of the consensus and filtering steps, we need to consider the upper bounds of the noise-free covariances for (35).

The upper bound for the noise-free covariance matrices $\Xi_i(k)$ can then be written as

$$P(k+1) = \mathcal{J}(P(k)), \qquad P(0) = \Xi(0), \tag{36}$$

with $\Gamma_i = (A - L_iC_i)$.

Lemma 1 implies that $r_\sigma(\mathcal{J}) = r_\sigma(\mathcal{L})$. Therefore, the dynamics in (17) are asymptotically stable *iff* (36) is asymptotically stable, and the notion of S-detectability in Definition 3 does not depend on the order in which the steps are taken. Furthermore, it has been shown in [19] that the coupled Riccati equations arising from reversing the order converge to a stabilizing solution *iff* the equations in (24) converge to a stabilizing solution. Therefore, the necessary and sufficient conditions we derive in Section IV do not depend on the order of the consensus and filtering steps. However, this realization of the filter requires agents to also share their local measurement and noise covariance matrices, making it less desirable than the distributed KF we propose here.

References

[1] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, “A survey on sensor networks,” *IEEE Communications Magazine*, vol. 40, no. 8, pp. 102–114, 2002.

[2] O. Ennasr and X. Tan, “Time-difference-of-arrival (TDOA)-based distributed target localization by a robotic network,” *IEEE Transactions on Control of Network Systems*, vol. 7, no. 3, pp. 1416–1427, 2020.

[3] M. S. Mahmoud and H. M. Khalid, “Distributed Kalman filtering: a bibliographic review,” *IET Control Theory & Applications*, vol. 7, no. 4, pp. 483–501, 2013.

[4] A. Speranzon, C. Fischione, and K. H. Johansson, “Distributed and collaborative estimation over wireless sensor networks,” in *Proceedings of the 45th IEEE Conference on Decision and Control*. IEEE, 2006, pp. 1025–1030.

[5] D. Acemoglu, A. Nedic, and A. Ozdaglar, “Convergence of rule-of-thumb learning rules in social networks,” in *2008 47th IEEE Conference on Decision and Control*. IEEE, 2008, pp. 1714–1720.

[6] R. Olfati-Saber and J. S. Shamma, “Consensus filters for sensor networks and distributed sensor fusion,” in *Proceedings of the 44th IEEE Conference on Decision and Control*. IEEE, 2005, pp. 6698–6703.

[7] R. Olfati-Saber, “Distributed Kalman filter with embedded consensus filters,” in *Proceedings of the 44th IEEE Conference on Decision and Control*. IEEE, 2005, pp. 8179–8184.

[8] R. Olfati-Saber, “Distributed Kalman filtering for sensor networks,” in *2007 46th IEEE Conference on Decision and Control*, Dec 2007, pp. 5492–5498.

[9] 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.

[10] F. S. Cattivelli and A. H. Sayed, “Diffusion strategies for distributed Kalman filtering and smoothing,” *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2069–2084, 2010.

[11] I. Matei and J. S. Baras, “Consensus-based linear distributed filtering,” *Automatica*, vol. 48, no. 8, pp. 1776–1782, 2012.

[12] D. Ding, Z. Wang, H. Dong, and H. Shu, “Distributed H_∞ state estimation with stochastic parameters and nonlinearities through sensor networks: the finite-horizon case,” *Automatica*, vol. 48, no. 8, pp. 1575–1585, 2012.

[13] V. Ugrinovskii, “Conditions for detectability in distributed consensus-based observer networks,” *IEEE Transactions on Automatic Control*, vol. 58, no. 10, pp. 2659–2664, 2013.

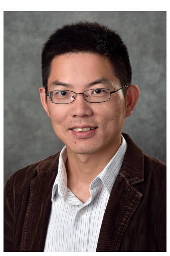
[14] ——, “Distributed robust estimation over randomly switching networks using H_∞ consensus,” *Automatica*, vol. 49, no. 1, pp. 160–168, 2013.

[15] M. Doostmohammadian and U. A. Khan, “Graph-theoretic distributed inference in social networks,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 4, pp. 613–623, 2014.

[16] S. Park and N. C. Martins, “Design of distributed LTI observers for state omniscience,” *IEEE Transactions on Automatic Control*, vol. 62, no. 2, pp. 561–576, Feb 2017.

[17] A. Mitra and S. Sundaram, "Distributed observers for LTI systems," *IEEE Transactions on Automatic Control*, vol. 63, no. 11, pp. 3689–3704, Nov 2018.
[18] T. Chang and A. Mehta, "Control-theoretical and topological analysis of covariance intersection based distributed Kalman filter," *IEEE Control Systems Letters*, vol. 2, no. 4, pp. 665–670, Oct 2018.
[19] O. L. V. Costa, M. D. Fragoso, and R. P. Marques, *Discrete-time Markov Jump Linear Systems*. Springer Science & Business Media, 2006.
[20] E. F. Costa and J. B. Do Val, "On the detectability and observability of discrete-time Markov jump linear systems," *Systems & Control Letters*, vol. 44, no. 2, pp. 135–145, 2001.
[21] ——, "Weak detectability and the linear-quadratic control problem of discrete-time Markov jump linear systems," *International Journal of Control*, vol. 75, no. 16-17, pp. 1282–1292, 2002.
[22] O. L. Costa and M. D. Fragoso, "Stability results for discrete-time linear systems with markovian jumping parameters," *Journal of Mathematical Analysis and Applications*, vol. 179, no. 1, pp. 154–178, 1993.
[23] J. B. do Val, J. C. Geromel, and O. L. Costa, "Uncoupled Riccati iterations for the linear quadratic control problem of discrete-time Markov jump linear systems," *IEEE Transactions on Automatic Control*, vol. 43, no. 12, pp. 1727–1733, 1998.

Osama Ennasr received the B.Sc. degree in Electrical Engineering from Michigan State University (East Lansing, MI, USA) in 2013, and he received the Ph.D. degree in Electrical and Computer Engineering Engineering from Michigan State University in 2020. His research interests include design, development, and applications of robotic platforms and networked systems to environmental sensing and mapping. Dr. Ennasr was the recipient of the Science Mathematics and Research for Transformation (SMART) scholarship in 2018, and is currently a robotics research scientist for the Geospatial Research Laboratory (GRL) and the U.S. Army Engineer Research and Development Center (ERDC), where he focuses on robotic applications to terrain mapping and exploration of unknown environments.



Xiaobo Tan (S'97 – M'02 – SM'11 – F'17) received the B.Eng. and M.Eng. degrees in automatic control from Tsinghua University, Beijing, China, in 1995 and 1998, respectively, and the Ph.D. degree in electrical and computer engineering from the University of Maryland, College Park, in 2002. He is currently an MSU Foundation Professor and the Richard M. Hong Endowed Chair in Electrical and Computer Engineering at Michigan State University (MSU). His research interests include modeling and control of systems with hysteresis, electroactive polymer sensors and actuators, soft robotics, and bio-inspired underwater robots and their application to environmental sensing. He has published over 300 papers and been awarded four US patents in these areas.

Dr. Tan currently serves as a Senior Editor for IEEE/ASME Transactions on Mechatronics. He is a Fellow of IEEE and ASME, and a recipient of NSF CAREER Award (2006), MSU Teacher-Scholar Award (2010), MSU College of Engineering Withrow Distinguished Scholar Award (2018), Distinguished Alumni Award from the Department of Electrical and Computer Engineering at University of Maryland (2018), and multiple best paper awards.