Distributed Kalman Filtering and Control Through Embedded Average Consensus Information Fusion

Sayed Pouria Talebi and Stefan Werner, *Senior Member, IEEE*

***Abstract*—This work presents a unified framework for distributed filtering and control of state-space processes. To this end, a distributed Kalman filtering algorithm is developed via decomposition of the optimal centralized Kalman filtering operations. This decomposition is orchestrated in a fashion so that each agent retains a Kalman style filtering operation and an estimate of the state vector. In this setting, the agents mirror the operations of the centralized Kalman filter in a distributed fashion through embedded average consensus fusion of local state vector estimates and their associated covariance information. For rigor, closed-form expressions for the mean and mean square error performance of the developed distributed Kalman filter are derived. More importantly, in contrast to current approaches, due to the comprehensive framework for fusion of the covariance information, a duality between the developed distributed Kalman filter and decentralized control is established. Thus, resulting in an effective and all inclusive distributed framework for filtering and control of state-space processes over a network of agents. The introduced theoretical concepts are validated using simulations that indicate a precise match between simulation results and the theoretical analysis. In addition, simulations indicate that performance levels comparable to that of the optimal centralized approaches are attainable.**

***Index Terms*—Sensor networks, consensus Kalman filtering, distributed adaptive sequential estimation, decentralized control.**

I. Introduction

In recent years, multi-agent networked systems have emerged as a feasible solution in a wide range of engineering applications [1]–[8]. Therefore, making the task of developing reliable distributed filtering and decentralized control techniques an attractive research topic [1,9,10]. Owing to the flexibility of the state-space representation for modeling real-world systems, optimality of the single-agent Kalman filter, and the duality of the Kalman filter with the linear quadratic regulator [11], there has been a concerted effort for developing Kalman filtering techniques for distributed filtering and decentralized control applications [1,12]–[15].

In order to present filtering solutions for sensor networks that are robust to link/agent failure and scalable with the size of the network, distributed filtering algorithms have been proposed in the context of consensus [12,16,17] and diffusion [9,15] literatures. These techniques incorporate consensus or diffusion steps where agents of the sensor network average their intermediate state vector estimates with their neighbors allowing the agents to establish a consensus on the estimate of the state vector [12,15,16,18]. However, the current literature on distributed Kalman filtering is mostly concerned with estimation problems; hence, the developed frameworks are not readily expandable to decentralized control applications. This issue is further elaborated upon in Section II.

In the control arena, decentralized control techniques have emerged as frameworks that are not only scalable to large networks and robust to link/agent failure, but also allow each agent to be reasonably aware of the network status; thus, providing for advantageous performance characteristics as compared to their single agent non-cooperative duals. This has made decentralized control techniques an essential part of multi-agent control systems [1,2,19]–[21]. However, a computationally efficient and analytically tractable framework directly linking decentralized control to both the centralized approach and distributed filtering, akin to what was performed in [11] for single agent systems, is still lacking. Establishing such a flexible framework is of paramount importance as it can pave the way for extending various concepts in control to the decentralized setting.

In this work, a comprehensive distributed Kalman filtering technique is derived. This is achieved through distributing the operations of the centralized Kalman filter over the network via embedded average consensus information fusion filters. Then, performance of the derived distributed Kalman filter is analyzed and closed-form expressions for the mean and mean square error are formulated. The analysis indicates that the introduced distributed Kalman filter operates in an unbiased fashion and can achieve mean square error performance levels on par with that of the centralized Kalman filter. In addition, using results from the conducted performance analysis in combination with the inherent duality between filtering and control, the framework of the derived distributed Kalman filter is expanded to decentralized control applications. The work is unique among its contemporaries in that it is simple to implement, mathematically tractable, and rigorously extends the duality between filtering and control problems to the distributed setting. Finally, the introduced concepts are verified in a number of simulations.

***Mathematical Notations*:** Scalars, column vectors, and matrices are denoted by lowercase, bold lowercase, and bold uppercase letters. The state vector at time instant n is denoted by $\mathbf{x}_n$, while $\mathbf{I}$ represents the identity matrix with the same number of rows as the state vector. The Kronecker product is denoted by $\otimes$. The transpose operator is denoted by $(\cdot)^{\mathsf{T}}$ with

Sayed Pouria Talebi is with the Department of Signal Processing and Acoustics, School of Electrical Engineering, Aalto University, Espoo FI-02150 Finland, E-mail: pouria.talebi@aalto.fi. Stefan Werner is with the Department of Electronic Systems, Faculty of Information Technology and Electrical Engineering, Norwegian University of Science and Technology, Trondheim NO-7491 Norway, E-mail: stefan.werner@ntnu.no. This work was supported in part by Academy of Finland under Grant 296849 and the Research Council of Norway.

$\mathbb{E}\{\cdot\}$ denoting the statistical expectation operator.

II. Problem Formulation

Consider a connected multi-agent network modeled as an undirected graph $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$, where the node set $\mathcal{N}$ denotes the agents of the network, with $|\mathcal{N}|$ representing the number of nodes in the network, while the edge set $\mathcal{E}$ represents bidirectional communication links between the agents of the network. The neighborhood of node l is defined as the set of nodes that can communicate with it, including self-communication. The neighborhood of node l is represented by the set $\mathcal{N}_l$ whose cardinality is denoted as $|\mathcal{N}_l|$.

The aim is to track a state vector sequence through local observations, which are related via the state-space model

$$\mathbf{x}_n = \mathbf{A}_n \mathbf{x}_{n-1} + \boldsymbol{\nu}_n \tag{1}$$

$$\mathbf{y}_{l,n} = \mathbf{H}_{l,n} \mathbf{x}_n + \boldsymbol{\omega}_{l,n} \tag{2}$$

where $\mathbf{x}_n$ and $\mathbf{A}_n$ denote the state vector and state evolution matrix at time instant n, whereas $\mathbf{y}_{l,n}$ and $\mathbf{H}_{l,n}$ represent the observation and observation function at node l at time instant n, while the process noise $\boldsymbol{\nu}_n$ and observation noise $\boldsymbol{\omega}_{l,n}$ are independent zero-mean white Gaussian noise processes with joint covariance matrix given by

$$\mathbb{E}\left\{ \begin{bmatrix} \boldsymbol{\nu}_n \\ \boldsymbol{\omega}_{l,n} \end{bmatrix} \begin{bmatrix} \boldsymbol{\nu}_m^\mathsf{T} & \boldsymbol{\omega}_{i,m}^\mathsf{T} \end{bmatrix} \right\} = \begin{bmatrix} \mathbf{C}_{\boldsymbol{\nu}_n} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\boldsymbol{\omega}_{l,n}} \delta_{l,i} \end{bmatrix} \delta_{n,m} \tag{3}$$

where $\delta_{n,m}$ denotes the Kronecker delta function.

The optimal solution to this problem, in the mean square error sense, comes in the form of the centralized Kalman filter [12], the operations of which are summarized in Algorithm 1 where $\hat{\mathbf{x}}_{n|n-1}$ and $\hat{\mathbf{x}}_{n|n}$ denote the *a priori* and *a posteriori* estimates of $\mathbf{x}_n$.

Algorithm 1. Centralized Kalman Filter [12]

Initialize with:

$$\hat{\mathbf{x}}_{0|0} = \mathbb{E}\{\mathbf{x}_0\}$$
$$\mathbf{M}_{0|0} = \mathbb{E}\left\{ (\mathbf{x}_0 - \mathbb{E}\{\mathbf{x}_0\}) (\mathbf{x}_0 - \mathbb{E}\{\mathbf{x}_0\})^\mathsf{T} \right\} \tag{4a}$$

Model update:

$$\hat{\mathbf{x}}_{n|n-1} = \mathbf{A}_n \hat{\mathbf{x}}_{n-1|n-1} \tag{4b}$$

$$\mathbf{M}_{n|n-1} = \mathbf{A}_n \mathbf{M}_{n-1|n-1} \mathbf{A}_n^\mathsf{T} + \mathbf{C}_{\boldsymbol{\nu}_n} \tag{4c}$$

Measurement update:

$$\mathbf{S}_n = \sum_{\forall l \in \mathcal{N}} \mathbf{H}_{l,n}^\mathsf{T} \mathbf{C}_{\boldsymbol{\omega}_{l,n}}^{-1} \mathbf{H}_{l,n} \tag{4d}$$

$$\mathbf{q}_n = \sum_{\forall l \in \mathcal{N}} \mathbf{H}_{l,n}^\mathsf{T} \mathbf{C}_{\boldsymbol{\omega}_{l,n}}^{-1} \mathbf{y}_{l,n} \tag{4e}$$

$$\mathbf{M}_{n|n}^{-1} = \mathbf{M}_{n|n-1}^{-1} + \mathbf{S}_n \tag{4f}$$

$$\hat{\mathbf{x}}_{n|n} = \hat{\mathbf{x}}_{n|n-1} + \mathbf{M}_{n|n} \left(\mathbf{q}_n - \mathbf{S}_n \hat{\mathbf{x}}_{n|n-1} \right) \tag{4g}$$

Although the centralized Kalman filter can be implemented using a central processing unit, this method makes the filtering operation vulnerable to failure of its processing unit. In addition, calculation of $\{\mathbf{S}_n, \mathbf{q}_n\}$ (see (4d) and (4e)) obligates the central processing unit to communicate with all agents of the network in a timely manner, requiring complex communication protocols [14,22]. Thus, distributed solutions are perused.

In the context of consensus Kalman filtering, each agent employs low- and band-pass consensus filters to approximate $\mathbf{q}_n$ and $\mathbf{S}_n$ in a distributed fashion [12,22,23]. Then, using these approximations, each agent implements the operations in Algorithm 1 in order to determine an intermediate estimate of the state vector. These intermediate state vector estimates are then fused using additional consensus filters in order to achieve a final estimate of the state vector [12,22,23].

In the context of diffusion Kalman filtering [9,15,24], each agent calculates an intermediate state vector estimate through implementing the operation in Algorithm 1 while replacing $\{\mathbf{S}_n, \mathbf{q}_n\}$ with their local approximations so that

$$\mathbf{S}_{l,n} = \sum_{\forall i \in \mathcal{N}_l} \mathbf{H}_{i,n}^\mathsf{T} \mathbf{C}_{\boldsymbol{\omega}_{i,n}}^{-1} \mathbf{H}_{i,n} \text{ and } \mathbf{q}_{l,n} = \sum_{\forall i \in \mathcal{N}_l} \mathbf{H}_{i,n}^\mathsf{T} \mathbf{C}_{\boldsymbol{\omega}_{i,n}}^{-1} \mathbf{y}_{i,n}$$

denote the approximations of $\mathbf{S}_n$ and $\mathbf{q}_n$ at node l. Then, these intermediate state vector estimates are diffused in order to formulate the final state vector estimate [9,15,24].

Remark 1. In comparison to diffusion and average consensus filters, implementing the required band-pass consensus filters in the consensus Kalman filtering approach is computationally demanding and generates a larger amount of communication traffic. Moreover, analysis of their behavior is not straightforward and is often reliant on the assumption that these filters have fully converged [23]. These factors hinder efforts for expanding such approaches to decentralized control applications.

Remark 2. Note that $\mathbf{S}_{l,n}$ (*cf.* $\mathbf{q}_{l,n}$) is an approximation of $\mathbf{S}_n$ (*cf.* $\mathbf{q}_n$) that only accounts for information available to an agent and its neighbors. This limits the use of diffusion-based approaches to decentralized control applications. Thus, a more comprehensive approach is developed in this work.

Remark 3. In both the consensus and diffusion approaches, the motivation for fusion of intermediate state vector estimates is to improve the accuracy of the state vector estimate at each node. However, a mathematically tractable link between the operations of the centralized Kalman filter in Algorithm 1 and these distributed Kalman filtering approaches is lacking.

III. The Proposed Distributed Kalman Filter

Using the operation of the centralized Kalman filter as the basis for our work, replacing (4d) and (4e) into (4g) yields

$$\begin{aligned} \hat{\mathbf{x}}_{n|n} = &\hat{\mathbf{x}}_{n|n-1} + \mathbf{M}_{n|n} \left(\sum_{\forall l \in \mathcal{N}} \mathbf{H}_{l,k}^\mathsf{T} \mathbf{C}_{\boldsymbol{\omega}_{l,n}}^{-1} \mathbf{y}_{l,n} \right) \\ &- \mathbf{M}_{n|n} \left(\sum_{\forall l \in \mathcal{N}} \mathbf{H}_{l,n}^\mathsf{T} \mathbf{C}_{\boldsymbol{\omega}_{l,n}}^{-1} \mathbf{H}_{l,n} \right) \hat{\mathbf{x}}_{n|n-1}. \end{aligned} \tag{5}$$

The expression in (5) is now rearranged to give

$$\begin{aligned} \hat{\mathbf{x}}_{n|n} &= \hat{\mathbf{x}}_{n|n-1} + \sum_{\forall l \in \mathcal{N}} \mathbf{M}_{n|n} \mathbf{H}_{l,k}^\mathsf{T} \mathbf{C}_{\boldsymbol{\omega}_{l,n}}^{-1} \left(\mathbf{y}_{l,n} - \mathbf{H}_{l,n} \hat{\mathbf{x}}_{n|n-1} \right) \\ &= \frac{1}{|\mathcal{N}|} \sum_{\forall l \in \mathcal{N}} \boldsymbol{\psi}_{l,n} \end{aligned} \tag{6}$$

where $\boldsymbol{\psi}_{l,n}$ denotes the intermediate state vector estimate at node l at time instant n and is given by

$$\boldsymbol{\psi}_{l,n} = \hat{\mathbf{x}}_{n|n-1} + \mathbf{G}_{l,n}\left(\mathbf{y}_{l,n} - \mathbf{H}_{l,n}\hat{\mathbf{x}}_{n|n-1}\right) \tag{7}$$

with $\mathbf{G}_{l,n}$ denoting the local gain matrix, that is given by

$$\mathbf{G}_{l,n} = |\mathcal{N}|\mathbf{M}_{n|n}\mathbf{H}_{l,n}^{\mathsf{T}}\mathbf{C}_{\boldsymbol{\omega}_{l,n}}^{-1}. \tag{8}$$

Substituting the expression in (4d) into (4f) and after some mathematical manipulations, we have

$$\begin{aligned}\mathbf{M}_{n,n}^{-1} =& \mathbf{M}_{n|n-1}^{-1} + \mathbf{S}_n \\ =& \mathbf{M}_{n|n-1}^{-1} + \sum_{\forall l\in\mathcal{N}} \mathbf{H}_{l,n}^{\mathsf{T}}\mathbf{C}_{\boldsymbol{\omega}_{l,n}}^{-1}\mathbf{H}_{l,n} = \frac{1}{|\mathcal{N}|}\sum_{\forall l\in\mathcal{N}}\boldsymbol{\Gamma}_{l,n}\end{aligned} \tag{9}$$

where the matrix $\boldsymbol{\Gamma}_{l,n}$ contains the local covariance information at node l at time instant n and is given by

$$\boldsymbol{\Gamma}_{l,n} = \mathbf{M}_{n|n-1}^{-1} + |\mathcal{N}|\mathbf{H}_{l,n}^{\mathsf{T}}\mathbf{C}_{\boldsymbol{\omega}_{l,n}}^{-1}\mathbf{H}_{l,n}. \tag{10}$$

Remark 4. In order to formulate (4f) in form of a network average of matrices $\{\boldsymbol{\Gamma}_{l,n} : \forall l \in \mathcal{N}\}$, in (10), the parameter $\mathbf{H}_{l,n}^{\mathsf{T}}\mathbf{C}_{\boldsymbol{\omega}_{l,n}}^{-1}\mathbf{H}_{l,n}$ is scaled by a factor of $|\mathcal{N}|$. In a similar fashion, the gain in (8) is scaled by a factor of $|\mathcal{N}|$ in order to preserve the equivalence between the implementation in (4g) and the implementation in (6)-(8).

Note that the operations of the centralized Kalman filter as given in Algorithm 1 can now be implemented in a distributed fashion through the averaging of the local covariance information, $\boldsymbol{\Gamma}_{l,n}$, as given in (10) and local state vector estimates, $\boldsymbol{\psi}_{l,n}$, given in (7). Thus, based on the work in [25,26], next a framework for distributed calculation of the averages in (9) and (6) is established. To this end, consider a set of matrices $\{\mathbf{F}_{j,(0)} : j = 1,\ldots,|\mathcal{N}|\}$ as inputs of the iterative consensus filter given by

$$\mathbf{F}_{i,(k)} = \mathbf{F}_{i,(k-1)} + \sum_{\forall j\in\mathcal{N}_i} w_{i,j}\left(\mathbf{F}_{j,(k-1)} - \mathbf{F}_{i,(k-1)}\right) \tag{11}$$

where $\mathbf{F}_{i,(k)}$ denotes the output of the iterative consensus filter at node i after k iterations, while $w_{i,j}$ denotes a positive real-valued weight. The iterations of the consensus filter in (11) can be expressed in a more comprehensive fashion as

$$\boldsymbol{\mathcal{F}}_{(k)} = (\boldsymbol{\mathcal{W}}\otimes\boldsymbol{\mathcal{I}})\boldsymbol{\mathcal{F}}_{(k-1)} = (\boldsymbol{\mathcal{M}}\otimes\boldsymbol{\mathcal{I}})\boldsymbol{\mathcal{F}}_{(0)} \tag{12}$$

where $\boldsymbol{\mathcal{M}} = \boldsymbol{\mathcal{W}}^k$, $\boldsymbol{\mathcal{I}}$ is an identity matrix of appropriate dimensions, $\boldsymbol{\mathcal{F}}_{(k)} = [\mathbf{F}_{1,(k)}^{\mathsf{T}}, \mathbf{F}_{2,(k)}^{\mathsf{T}}, \ldots, \mathbf{F}_{|\mathcal{N}|,(k)}^{\mathsf{T}}]^{\mathsf{T}}$, and the element on the i^{th} row and j^{th} column of $\boldsymbol{\mathcal{W}}$ is

$$\boldsymbol{\mathcal{W}}_{i,j} = \begin{cases} 1 + w_{i,i} - \sum_{\forall l\in\mathcal{N}_i} w_{i,l} & \text{if } i = j, \\ w_{i,j} & \text{if } i \in \mathcal{N}_j\backslash j \\ 0 & \text{otherwise.}\end{cases}$$

If the weights $w_{i,j}$ are selected so that $\boldsymbol{\mathcal{W}}$ is also doubly stochastic and meets conditions in [25]; then, it follows that

$$\forall i \in \mathcal{N}: \quad \lim_{k\to\infty}\mathbf{F}_{i,(k)} = \frac{1}{|\mathcal{N}|}\sum_{\forall j\in\mathcal{N}}\mathbf{F}_{j,(0)} \tag{13}$$

which is the average consensus filter (ACF) required to calculate the averages in (9) and (6). For the sake of simplicity in presentation, the operation of the ACF at node i after k iterations is represented via the schematic

$$\mathbf{F}_{i,(k)} \leftarrow \boxed{\text{ACF}} \leftarrow \{\forall j \in \mathcal{N} : \mathbf{F}_{j,(0)}\}$$

where $\{\forall j \in \mathcal{N} : \mathbf{F}_{j,(0)}\}$ are the network-wide inputs to the ACF, that is initial inputs at all nodes in the network, and $\mathbf{F}_{i,(k)}$ is the output at node i after k iterations.

The operations of the derived distributed Kalman filter are summarized in Algorithm 2, where $\hat{\mathbf{x}}_{l,n|n-1}$ and $\hat{\mathbf{x}}_{l,n|n}$ denote the *a priori* and *a posteriori* state vector estimates at node l at time instant n, while the ACF is iterated a predefined number of times in order to achieve consensus.

Algorithm 2. Distributed Kalman Filter Through Embedded Average Consensus Information Fusion

For nodes $l = \{1,\ldots,|\mathcal{N}|\}$:
Initialize with:

$$\hat{\mathbf{x}}_{l,0|0} = \mathbb{E}\{\mathbf{x}_0\} \tag{14a}$$

$$\mathbf{M}_{l,0|0} = \mathbb{E}\left\{\left(\mathbf{x}_0 - \mathbb{E}\{\mathbf{x}_0\}\right)\left(\mathbf{x}_0 - \mathbb{E}\{\mathbf{x}_0\}\right)^{\mathsf{T}}\right\} \tag{14b}$$

Model update:

$$\hat{\mathbf{x}}_{l,n|n-1} = \mathbf{A}_n\hat{\mathbf{x}}_{l,n-1|n-1} \tag{14c}$$

$$\mathbf{M}_{l,n|n-1} = \mathbf{A}_n\mathbf{M}_{l,n-1|n-1}\mathbf{A}_n^{\mathsf{T}} + \mathbf{C}_{\boldsymbol{\nu}_n} \tag{14d}$$

Measurement update:

$$\boldsymbol{\Gamma}_{l,n} = \mathbf{M}_{l,n|n-1}^{-1} + |\mathcal{N}|\mathbf{H}_{l,n}^{\mathsf{T}}\mathbf{C}_{\boldsymbol{\omega}_{l,n}}^{-1}\mathbf{H}_{l,n} \tag{14e}$$

$$\mathbf{M}_{l,n|n}^{-1} \leftarrow \boxed{\text{ACF}} \leftarrow \{\forall i \in \mathcal{N} : \boldsymbol{\Gamma}_{i,n}\} \tag{14f}$$

$$\mathbf{G}_{l,n} = |\mathcal{N}|\mathbf{M}_{l,n|n}\mathbf{H}_{l,n}^{\mathsf{T}}\mathbf{C}_{\boldsymbol{\omega}_{l,n}}^{-1} \tag{14g}$$

$$\boldsymbol{\psi}_{l,n} = \hat{\mathbf{x}}_{l,n|n-1} + \mathbf{G}_{l,n}\left(\mathbf{y}_{l,n} - \mathbf{H}_{l,n}\hat{\mathbf{x}}_{l,n|n-1}\right) \tag{14h}$$

$$\hat{\mathbf{x}}_{l,n|n} \leftarrow \boxed{\text{ACF}} \leftarrow \{\forall i \in \mathcal{N} : \boldsymbol{\psi}_{i,n}\} \tag{14i}$$

IV. Performance Evaluation

Throughout this section, the following standard assumptions in Kalman filtering analysis are held to be true [23,27]:

Assumption 1: The state evolution and observation function become time invariant, i.e.,

$$\lim_{n\to\infty}\mathbf{A}_n = \mathbf{A} \quad \text{and} \quad \forall l \in \mathcal{N}: \lim_{n\to\infty}\mathbf{H}_{l,n} = \mathbf{H}_l.$$

Assumption 2: The state evolution noise and observation noises become stationary, i.e.,

$$\lim_{n\to\infty}\mathbf{C}_{\boldsymbol{\nu}_n} = \mathbf{C}_{\boldsymbol{\nu}} \quad \text{and} \quad \forall l \in \mathcal{N}: \lim_{n\to\infty}\mathbf{C}_{\boldsymbol{\omega}_{l,n}} = \mathbf{C}_{\boldsymbol{\omega}_l}.$$

which consequently results in

$$\lim_{n\to\infty}\mathbf{C}_{\boldsymbol{\nu}_{col,n}} = \mathbf{C}_{\boldsymbol{\nu}_{col}} \quad \text{and} \quad \lim_{n\to\infty}\mathbf{C}_{\boldsymbol{\omega}_{col,n}} = \mathbf{C}_{\boldsymbol{\omega}_{col}}$$

where

$$\mathbf{C}_{\boldsymbol{\nu}_{col,n}} = \mathbb{E}\left\{\boldsymbol{\nu}_{col,n}\boldsymbol{\nu}_{col,n}^{\mathsf{T}}\right\} \text{ and } \mathbf{C}_{\boldsymbol{\omega}_{col,n}} = \mathbb{E}\left\{\boldsymbol{\omega}_{col,n}\boldsymbol{\omega}_{col,n}^{\mathsf{T}}\right\}$$

with

$$\boldsymbol{\nu}_{col,n} = [\boldsymbol{\nu}_n^{\mathsf{T}},\ldots,\boldsymbol{\nu}_n^{\mathsf{T}}]^{\mathsf{T}} \text{ and } \boldsymbol{\omega}_{col,n} = [\boldsymbol{\omega}_{1,n}^{\mathsf{T}},\ldots,\boldsymbol{\omega}_{|\mathcal{N}|,n}^{\mathsf{T}}]^{\mathsf{T}}$$

denoting column-block matrices composed of $|\mathcal{N}|$ repetitions of $\boldsymbol{\nu}_n$ and $\{\boldsymbol{\omega}_{l,n} : \forall l \in \mathcal{N}\}$ respectively.

Assumption 3: The matrix pairs $\forall l \in \mathcal{N} : \{\mathbf{A}, \mathbf{H}_l\}$ are jointly observable over the communication matrix $\boldsymbol{\mathcal{M}} = \boldsymbol{\mathcal{W}}^k$ (see (12)) and the matrix pair $\{\mathbf{A}, \mathbf{C}_{\boldsymbol{\nu}}^{\frac{1}{2}}\}$ is controllable.

Assumption 4: The matrices $\forall l \in \mathcal{N} : \{\boldsymbol{\Gamma}_{l,n}, \mathbf{M}_{l,n}\}$ remain positive definite at all time instances.[1]

First, it is shown that the matrix set $\{\mathbf{M}_{l,n|n-1}, \forall l \in \mathcal{N}\}$, and by extension $\{\mathbf{M}_{l,n|n}, \forall l \in \mathcal{N}\}$, converge to a unique stabilizing solution. To this end, substituting (10) and (11) into (14f) gives

$$\begin{aligned}\mathbf{M}_{l,n|n}^{-1} =& \sum_{\forall i \in \mathcal{N}} \boldsymbol{\mathcal{M}}_{l,i} \boldsymbol{\Gamma}_{i,n} \\ =& \mathbf{M}_{l,n|n-1}^{-1} + \sum_{\forall i \in \mathcal{N}} \boldsymbol{\mathcal{M}}_{l,i} \left(\mathbf{M}_{i,n|n-1}^{-1} - \mathbf{M}_{l,n|n-1}^{-1} \right) \\ &+ \sum_{\forall i \in \mathcal{N}} \boldsymbol{\mathcal{M}}_{l,i} |\mathcal{N}| \mathbf{H}_{i,n}^{\mathsf{T}} \mathbf{C}_{\boldsymbol{\omega}_{i,n}}^{-1} \mathbf{H}_{i,n}\end{aligned} \tag{15}$$

where $\boldsymbol{\mathcal{M}}_{l,i}$ is the element on the l^{th} row and i^{th} column of $\boldsymbol{\mathcal{M}} = \boldsymbol{\mathcal{W}}^k$. Replacing (14d) into (15) allows the dynamics of matrix set $\{\mathbf{M}_{l,n|n-1} : \forall l \in \mathcal{N}\}$ to be rearranged as

$$\forall l \in \mathcal{N} : \begin{cases} \mathbf{M}_{l,n|n-1} = \mathbf{A}_n \mathbf{M}_{l,n-1|n-1} \mathbf{A}_n^{\mathsf{T}} + \mathbf{C}_{\boldsymbol{\nu}_n} \\ \boldsymbol{\Phi}_{l,n}^{-1} \leftarrow \boxed{\text{ACF}} \leftarrow \{\forall i \in \mathcal{N} : \mathbf{M}_{i,n|n-1}^{-1}\} \\ \mathbf{M}_{l,n|n}^{-1} = \boldsymbol{\Phi}_{l,n}^{-1} + \boldsymbol{\mathcal{H}}_{l,n}^{\mathsf{T}} \mathbf{C}_{\boldsymbol{\omega}_{col,n}}^{-1} \boldsymbol{\mathcal{H}}_{l,n} \end{cases} \tag{16}$$

with

$$\boldsymbol{\mathcal{H}}_{l,n} = \sqrt{|\mathcal{N}|} \left[\sqrt{\boldsymbol{\mathcal{M}}_{l,1}} \mathbf{H}_{1,n}^{\mathsf{T}}, \ldots, \sqrt{\boldsymbol{\mathcal{M}}_{l,|\mathcal{N}|}} \mathbf{H}_{|\mathcal{N}|,n}^{\mathsf{T}} \right]^{\mathsf{T}}.$$

Alternatively (16) can be expressed as

$$\forall l \in \mathcal{N} : \begin{cases} \mathbf{M}_{l,n|n-1} = f_{l,n}(\boldsymbol{\Phi}_{l,n-1}) \\ \boldsymbol{\Phi}_{l,n}^{-1} \leftarrow \boxed{\text{ACF}} \leftarrow \{\forall i \in \mathcal{N} : \mathbf{M}_{i,n|n-1}^{-1}\} \end{cases} \tag{17}$$

where

$$f_{l,n}(\boldsymbol{\Phi}_{l,n}) = \mathbf{A}_n \left(\boldsymbol{\Phi}_{l,n}^{-1} + \boldsymbol{\mathcal{H}}_{l,n}^{\mathsf{T}} \mathbf{C}_{\boldsymbol{\omega}_{col,n}}^{-1} \boldsymbol{\mathcal{H}}_{l,n} \right)^{-1} \mathbf{A}_n^{\mathsf{T}} + \mathbf{C}_{\boldsymbol{\nu}_n}. \tag{18}$$

Focusing on the recursive functions

$$\forall l \in \mathcal{N} : \boldsymbol{\Lambda}_n = f_{l,n}(\boldsymbol{\Lambda}_{n-1}) \tag{19}$$

it follows from *Assumption* 1 and *Assumption* 2 that the function $f_{l,n}(\cdot)$ and matrices $\{\boldsymbol{\mathcal{H}}_{l,n}, \mathbf{C}_{\boldsymbol{\omega}_{col,n}}^{-1}\}$ become time invariant. Thus, (19) constitutes an algebraic Riccati equation that will converge to a unique stabilizing solution given that $\{\mathbf{A}, \boldsymbol{\mathcal{H}}_l\}$ are detectable and the matrices $\{\mathbf{A}, \mathbf{C}_{\boldsymbol{\nu}}^{\frac{1}{2}}\}$ are stabilizable, i.e., *Assumption* 3 and *Assumption* 4 are satisfied. Finally, from **Theorem** 1 and **Theroem** 2, it follows that this is a sufficient condition for matrices $\{\mathbf{M}_{l,n|n-1} : \forall l \in \mathcal{N}\}$ to converge to unique stabilizing solutions.

Theorem 1. *If the recursions in (19) converge to a unique set of stabilizing matrices; then, the recursion in (17) also converges to a set of stabilizing matrices.*

[1]This holds true as the matrices in question are calculated from the summation of positive and positive semi-definite matrices. Note that this ensures the matrices $\{\mathbf{M}_{l,n|n} : \forall l \in \mathcal{N}\}$ are invertible. In practice, if required, at each time instance these matrices can be substituted with their positive definite approximations through replacing of their negative eigenvalues with a sufficiently small positive valued variable.

Proof of Theorem 1: Consider the matrix set

$$\{\boldsymbol{\Delta}_{l,n} = \mathbf{M}_{l,n+1|n} - \mathbf{M}_{l,n|n-1} : \forall l \in \mathcal{N}\}.$$

After some mathematical manipulations and using the framework of Lemma 4.2 in [28], we have

$$\begin{aligned}\boldsymbol{\Delta}_{l,n} =& \mathbf{L}_{l,n}^{\mathsf{T}} \tilde{\boldsymbol{\Delta}}_{l,n} \mathbf{L}_{l,n} \\ &- \mathbf{L}_{l,n}^{\mathsf{T}} \tilde{\boldsymbol{\Delta}}_{l,n} \boldsymbol{\mathcal{H}}_{l,n}^{\mathsf{T}} \mathbf{Z}_{l,n}^{-1} \boldsymbol{\mathcal{H}}_{l,n} \tilde{\boldsymbol{\Delta}}_{l,n} \mathbf{L}_{l,n}\end{aligned} \tag{20}$$

where $\tilde{\boldsymbol{\Delta}}_{l,n} = \boldsymbol{\Phi}_{l,n} - \boldsymbol{\Phi}_{l,n-1}$ while

$$\begin{aligned}\mathbf{K}_{l,n} =& \left(\boldsymbol{\mathcal{H}}_{l,n} \boldsymbol{\Phi}_{l,n-1} \boldsymbol{\mathcal{H}}_{l,n}^{\mathsf{T}} + \mathbf{C}_{\boldsymbol{\omega}_{col,n}} \right)^{-1} \boldsymbol{\mathcal{H}}_{l,n} \boldsymbol{\Phi}_{l,n-1} \mathbf{A}_n^{\mathsf{T}} \\ \mathbf{L}_{l,n} =& \mathbf{A}_n^{\mathsf{T}} - \boldsymbol{\mathcal{H}}_{l,n}^{\mathsf{T}} \mathbf{K}_{l,n} \\ \mathbf{Z}_{l,n} =& \boldsymbol{\mathcal{H}}_{l,n} \tilde{\boldsymbol{\Delta}}_{l,n} \boldsymbol{\mathcal{H}}_{l,n}^{\mathsf{T}} + \boldsymbol{\mathcal{H}}_{l,n} \boldsymbol{\Phi}_{l,n-1} \boldsymbol{\mathcal{H}}_{l,n}^{\mathsf{T}} + \mathbf{C}_{\boldsymbol{\omega}_{col,n}}.\end{aligned}$$

Moreover, from the expression in (11)-(12) and (17), we have

$$\tilde{\boldsymbol{\Delta}}'_{l,n} \leftarrow \boxed{\text{ACF}} \leftarrow \{\forall i \in \mathcal{N} : \boldsymbol{\Delta}'_{i,n-1}\} \tag{21}$$

where $\forall i \in \mathcal{N} : \boldsymbol{\Delta}'_{i,n-1} = \mathbf{M}_{l,n|n-1}^{-1} - \mathbf{M}_{l,n-1|n-2}^{-1}$ and $\tilde{\boldsymbol{\Delta}}'_{l,n} = \boldsymbol{\Phi}_{l,n}^{-1} - \boldsymbol{\Phi}_{l,n-1}^{-1}$.

If *Assumption* 1-*Assumption* 4 hold true; then, from [28], the expression in (20) will be a contracting function (with regards to input $\tilde{\boldsymbol{\Delta}}_{l,n}$ and output $\boldsymbol{\Delta}_{l,n}$). In addition, taking into account that $\boldsymbol{\mathcal{W}}$ is doubly stochastic, the spectral radius of $\boldsymbol{\mathcal{M}} = \boldsymbol{\mathcal{W}}^k$ is equal to one, and hence, the ACF will have a non-expanding effect on recursions in (20)-(21). Finally, note that the recursions in (20)-(21) have a stable point at $\forall l \in \mathcal{N} : \boldsymbol{\Delta}_{l,n} = \mathbf{0}$. Therefore, $\forall l \in \mathcal{N} : \boldsymbol{\Delta}_{l,n} \to \mathbf{0}$ as $n \to \infty$. This indicates that matrices $\{\mathbf{M}_{l,n|n-1} : \forall l \in \mathcal{N}\}$ become time invariant, i.e., they converge to a stabilizing solution. ■

Theorem 2. *Under Assumption 1-Assumption 4, the stabilizing solution of (17) is unique.*

Proof of Theorem 2: Assume that as $n \to \infty$ both matrix sets $\{\mathbf{M}_{l,n+1|n}, \boldsymbol{\Phi}_{l,n} : \forall l \in \mathcal{N}\}$ and $\{\mathbf{M}'_{l,n+1|n}, \boldsymbol{\Phi}'_{l,n} : \forall l \in \mathcal{N}\}$ stabilize (17). Then, for $l \in \mathcal{N}$, the evolution of $\mathbf{M}_{l,n+1|n} - \mathbf{M}'_{l,n+1|n}$ is attainable from (21)-(20) where $\mathbf{M}_{l,n|n-1}$ and $\boldsymbol{\Phi}_{l,n-1}$ are to be replaced by $\mathbf{M}'_{l,n+1|n}$ and $\boldsymbol{\Phi}'_{l,n}$.[2] Once again, if *Assumption* 1-*Assumption* 4 hold; then, from [28], as $n \to \infty$, the recursions in question converge to their stabilizing solution at $\mathbf{M}_{l,n|n-1} = \mathbf{M}'_{l,n|n-1}$ indicating a unique solution. ■

The error of the intermediate state vector estimate at node l at time instant n is given by $\boldsymbol{\epsilon}_{l,n} = \mathbf{x}_n - \boldsymbol{\psi}_{l,n}$, which using the expression in (14h) can alternatively be expressed as

$$\boldsymbol{\epsilon}_{l,n} = \mathbf{x}_n - \hat{\mathbf{x}}_{l,n|n-1} - \mathbf{G}_{l,n} \left(\mathbf{y}_{l,n} - \mathbf{H}_{l,n} \hat{\mathbf{x}}_{n|n-1} \right). \tag{22}$$

Now, substituting (2) and $\boldsymbol{\epsilon}_{l,n|n-1} = \mathbf{x}_n - \hat{\mathbf{x}}_{l,n|n-1}$ into the expression in (22) yields

$$\boldsymbol{\epsilon}_{l,n} = \left(\mathbf{I} - \mathbf{G}_{l,n} \mathbf{H}_{l,n} \right) \boldsymbol{\epsilon}_{l,n|n-1} - \mathbf{G}_{l,n} \boldsymbol{\omega}_{l,n}. \tag{23}$$

Moreover, denoting $\boldsymbol{\epsilon}_{l,n|n} = \mathbf{x}_n - \hat{\mathbf{x}}_{l,n|n}$, from (1) we have

$$\boldsymbol{\epsilon}_{l,n|n-1} = \mathbf{A}_n \boldsymbol{\epsilon}_{l,n-1|n-1} + \boldsymbol{\nu}_n.$$

[2]Once more, this result is a direct implication of Lemma 4.2 in [28].

This allows (23) to be reformulated as

$$\begin{aligned}\boldsymbol{\epsilon}_{l,n} =& \left(\mathbf{I}-\mathbf{G}_{l,n}\mathbf{H}_{l,n}\right)\mathbf{A}_n\boldsymbol{\epsilon}_{l,n-1|n-1}\\ &+\left(\mathbf{I}-\mathbf{G}_{l,n}\mathbf{H}_{l,n}\right)\boldsymbol{\nu}_n-\mathbf{G}_{l,n}\boldsymbol{\omega}_{l,n}.\end{aligned} \tag{24}$$

In order to analyze the impact of the ACF and to express the state estimation error term in a regressive fashion, we need to consider the network-wide dual of the expression in (24) which can be formulated as

$$\boldsymbol{\mathcal{E}}_n = \boldsymbol{\mathcal{P}}_n\boldsymbol{\mathcal{E}}_{n-1|n-1}+\boldsymbol{\mathcal{Q}}_n\boldsymbol{\nu}_{col,n}-\boldsymbol{\mathcal{T}}_n\boldsymbol{\omega}_{col,n} \tag{25}$$

where $\boldsymbol{\mathcal{P}}_n$, $\boldsymbol{\mathcal{Q}}_n$, and $\boldsymbol{\mathcal{T}}_n$ are block-diagonal matrices with the following structures

$$\begin{aligned}\boldsymbol{\mathcal{P}}_n =& \text{block-diag}\{\left(\mathbf{I}-\mathbf{G}_{l,n}\mathbf{H}_{l,n}\right)\mathbf{A}_n, l=1,\ldots,|\mathcal{N}|\}\\ \boldsymbol{\mathcal{Q}}_n =& \text{block-diag}\{\left(\mathbf{I}-\mathbf{G}_{l,n}\mathbf{H}_{l,n}\right), l=1,\ldots,|\mathcal{N}|\}\\ \boldsymbol{\mathcal{T}}_n =& \text{block-diag}\{\mathbf{G}_{l,n}, l=1,\ldots,|\mathcal{N}|\}\end{aligned}$$

while

$$\boldsymbol{\mathcal{E}}_n=\begin{bmatrix}\boldsymbol{\epsilon}_{1,n}\\ \vdots\\ \boldsymbol{\epsilon}_{|\mathcal{N}|,n}\end{bmatrix} \quad \text{and} \quad \boldsymbol{\mathcal{E}}_{n-1|n-1}=\begin{bmatrix}\boldsymbol{\epsilon}_{1,n-1|n-1}\\ \vdots\\ \boldsymbol{\epsilon}_{|\mathcal{N}|,n-1|n-1}\end{bmatrix}.$$

Now, considering the ACF given in (11), we have

$$\boldsymbol{\epsilon}_{l,n|n} \leftarrow \boxed{\text{ACF}} \leftarrow \{\forall i\in\mathcal{N}:\boldsymbol{\epsilon}_{i,n}\}$$

which using the expressions in (12) and (25) can be formulated in a more mathematically rigorous manner as

$$\begin{aligned}\boldsymbol{\mathcal{E}}_{n|n} =& (\boldsymbol{\mathcal{W}}^k\otimes\boldsymbol{\mathcal{I}})\boldsymbol{\mathcal{E}}_n\\ =& \underbrace{(\boldsymbol{\mathcal{W}}^k\otimes\boldsymbol{\mathcal{I}})\boldsymbol{\mathcal{P}}_n}_{\boldsymbol{\mathcal{P}}_n}\boldsymbol{\mathcal{E}}_{n-1|n-1}+\underbrace{(\boldsymbol{\mathcal{W}}^k\otimes\boldsymbol{\mathcal{I}})\boldsymbol{\mathcal{Q}}_n}_{\boldsymbol{\mathcal{Q}}_n}\boldsymbol{\nu}_{col,n}\\ &-\underbrace{(\boldsymbol{\mathcal{W}}^k\otimes\mathcal{I})\boldsymbol{\mathcal{T}}_n}_{\boldsymbol{\mathcal{T}}_n}\boldsymbol{\omega}_{col,n}.\end{aligned} \tag{26}$$

Remark 5. Taking the statistical expectation of (26) and considering that $\boldsymbol{\nu}_n$ and $\{\boldsymbol{\omega}_{l,n}, l\in\mathcal{N}\}$ are zero-mean, yields

$$\mathbb{E}\left\{\boldsymbol{\mathcal{E}}_{n|n}\right\}=\boldsymbol{\mathcal{P}}_n\mathbb{E}\left\{\boldsymbol{\mathcal{E}}_{n-1|n-1}\right\}=\left(\prod_{i=1}^{n}\boldsymbol{\mathcal{P}}_i\right)\mathbb{E}\left\{\boldsymbol{\mathcal{E}}_{0|0}\right\}. \tag{27}$$

Moreover, from (14a) we have $\forall l\in\mathcal{N}:\hat{\mathbf{x}}_{l,0|0}=\mathbb{E}\{\mathbf{x}_0\}$ resulting in $\mathbb{E}\left\{\boldsymbol{\mathcal{E}}_{0|0}\right\}=0$. Therefore, the expression in (27) proves the proposed algorithm operates in an unbiased fashion.

From the regressive expression of the state vector estimation error derived in (26) we have

$$\begin{aligned}\boldsymbol{\mathcal{E}}_{n|n}\boldsymbol{\mathcal{E}}_{n|n}^{\mathsf{T}} =& \boldsymbol{\mathcal{P}}_n\boldsymbol{\mathcal{E}}_{n-1|n-1}\boldsymbol{\mathcal{E}}_{n-1|n-1}^{\mathsf{T}}\boldsymbol{\mathcal{P}}_n^{\mathsf{T}}\\ &+\boldsymbol{\mathcal{Q}}_n\boldsymbol{\nu}_{col,n}\boldsymbol{\nu}_{col,n}^{\mathsf{T}}\boldsymbol{\mathcal{Q}}_n^{\mathsf{T}}-\boldsymbol{\mathcal{P}}_n\boldsymbol{\mathcal{E}}_{n-1|n-1}\boldsymbol{\omega}_{col,n}^{\mathsf{T}}\boldsymbol{\mathcal{T}}_n^{\mathsf{T}}\\ &-\boldsymbol{\mathcal{Q}}_n\boldsymbol{\nu}_{col,n}\boldsymbol{\omega}_{col,n}^{\mathsf{T}}\boldsymbol{\mathcal{T}}_n^{\mathsf{T}}+\boldsymbol{\mathcal{P}}_n\boldsymbol{\mathcal{E}}_{n-1|n-1}\boldsymbol{\nu}_{col,n}^{\mathsf{T}}\boldsymbol{\mathcal{Q}}_n^{\mathsf{T}}\\ &-\boldsymbol{\mathcal{T}}_n\boldsymbol{\omega}_{col,n}\boldsymbol{\mathcal{E}}_{n-1|n-1}^{\mathsf{T}}\boldsymbol{\mathcal{P}}_n^{\mathsf{T}}+\boldsymbol{\mathcal{T}}_n\boldsymbol{\omega}_{col,n}\boldsymbol{\omega}_{col,n}^{\mathsf{T}}\boldsymbol{\mathcal{T}}_n^{\mathsf{T}}\\ &+\boldsymbol{\mathcal{Q}}_n\boldsymbol{\nu}_{col,n}\boldsymbol{\mathcal{E}}_{n-1|n-1}^{\mathsf{T}}\boldsymbol{\mathcal{P}}_n^{\mathsf{T}}-\boldsymbol{\mathcal{T}}_n\boldsymbol{\omega}_{col,n}\boldsymbol{\nu}_{col,n}^{\mathsf{T}}\boldsymbol{\mathcal{Q}}_n^{\mathsf{T}}.\end{aligned} \tag{28}$$

Given that noise sequences $\boldsymbol{\nu}_n$ and $\{\boldsymbol{\omega}_{l,n},\forall l\in\mathcal{N}\}$ are white zero-mean Gaussian processes with joint the covariance matrix given in (3), taking the statistical expectation of the expression in (28) allows the network state vector estimation error covariance matrix to be formulated as

$$\boldsymbol{\Sigma}_n=\boldsymbol{\mathcal{P}}_n\boldsymbol{\Sigma}_{n-1}\boldsymbol{\mathcal{P}}_n^{\mathsf{T}}+\boldsymbol{\mathcal{Q}}_n\mathbf{C}_{\boldsymbol{\nu}_{col,n}}\boldsymbol{\mathcal{Q}}_n^{\mathsf{T}}+\boldsymbol{\mathcal{T}}_n\mathbf{C}_{\boldsymbol{\omega}_{col,n}}\boldsymbol{\mathcal{T}}_n^{\mathsf{T}} \tag{29}$$

where $\boldsymbol{\Sigma}_n = \mathbb{E}\left\{\boldsymbol{\mathcal{E}}_{n|n}\boldsymbol{\mathcal{E}}_{n|n}^{\mathsf{T}}\right\}$. Moreover, the matrices $\{\mathbf{M}_{l,n|n}, l\in\mathcal{N}\}$ converge, i.e., $\forall l\in\mathcal{N}:\lim_{n\to\infty}\mathbf{M}_{l,n|n}=\mathbf{M}_l$; thus, resulting in matrices $\{\mathbf{G}_{l,n}, l\in\mathcal{N}\}$ becoming time invariant. From (24)-(26) and (28)-(29) it becomes clear that if the matrices $\{\mathbf{G}_{l,n}, l\in\mathcal{N}\}$ become time invariant; then, the matrices $\{\boldsymbol{\mathcal{P}}_n,\boldsymbol{\mathcal{Q}}_n,\boldsymbol{\mathcal{T}}_n\}$ also become time invariant, i.e.,

$$\lim_{n\to\infty}\boldsymbol{\mathcal{P}}_n=\boldsymbol{\mathcal{P}},\ \lim_{n\to\infty}\boldsymbol{\mathcal{Q}}_n=\boldsymbol{\mathcal{Q}}, \text{ and } \lim_{n\to\infty}\boldsymbol{\mathcal{T}}_n=\boldsymbol{\mathcal{T}}$$

and as a result $\boldsymbol{\Sigma}_n$ in the formulation given in (29) converges, that is, $\boldsymbol{\Sigma}_n\to\boldsymbol{\Sigma}$ as $n\to\infty$, where $\boldsymbol{\Sigma}$ is the solution of the discrete time Lyapunov equation given by

$$\boldsymbol{\Sigma}=\boldsymbol{\mathcal{P}}\boldsymbol{\Sigma}\boldsymbol{\mathcal{P}}^{\mathsf{T}}+\boldsymbol{\mathcal{Q}}\mathbf{C}_{\boldsymbol{\nu}_{col}}\boldsymbol{\mathcal{Q}}^{\mathsf{T}}+\boldsymbol{\mathcal{T}}\mathbf{C}_{\boldsymbol{\omega}_{col}}\boldsymbol{\mathcal{T}}^{\mathsf{T}}. \tag{30}$$

Remark 6. In the case for which ACF iterations are large enough so that $\forall i,j\in\mathcal{N}:\boldsymbol{\mathcal{M}}_{i,j}\approx 1\backslash|\mathcal{N}|$; then, from the derivation processes in Section III, it becomes clear that the matrices $\{\mathbf{M}_{l,n|n}:\forall l\in\mathcal{N}\}$ converge to the same value as their centralized counterparts in Algorithm 1, obtainable from replacing (4c) and (4d) into (4f). In a similar manner, it follows that the state vector estimates of the proposed distributed Kalman filter will be close to those obtained by the centralized Kalman filter, that is, $\forall l\in\mathcal{N}:\hat{\mathbf{x}}_{l,n|n}\approx\hat{\mathbf{x}}_{n|n}$.

Remark 7. In a steady-state situation, where the matrices $\{\mathbf{M}_{l,n|n}:\forall l\in\mathcal{N}\}$ have converged, from Algorithm 2, it becomes apparent that the need for implementing ACF operations with regards to $\{\boldsymbol{\Gamma}_{l,n}:\forall l\in\mathcal{N}\}$ is eliminated. In this setting, only the local state vector estimates, $\boldsymbol{\psi}_{l,n}$, are required to be communicated. This reduces the communication traffic and implementation complexity to a minimum. In comparison, other consensus and diffusion approaches (see [10,15,16,18,22,23]) that require the observation information, $\mathbf{H}_{i,n}^{\mathsf{T}}\mathbf{C}_{\boldsymbol{\omega}_{i,n}}^{-1}\mathbf{y}_{i,n}$, to also be shared (see Section II) introducing an extra level of complexity.

Remark 8. From (14e) and (14f) in Algorithm 2, note that the observation covariance information at node $l\in\mathcal{N}$, i.e., $\mathbf{H}_{l,n}^{\mathsf{T}}\mathbf{C}_{\boldsymbol{\omega}_{l,n}}\mathbf{H}_{l,n}$, is integrated into the local covariance information, $\boldsymbol{\Gamma}_{l,n}$, and then fused over the network via the ACF. This provides a comprehensive framework for fusion of the covariance information and sets this work apart from its peers. For example, current consensus Kalman filtering techniques require more computationally complex band-pass consensus filters for the fusion of their covariance information, while in the diffusion Kalman filtering framework each node only has access to the observation covariance information from its neighboring nodes (see Section II).

V. Control Applications

The comprehensive fusion of covariance information in the proposed distributed Kalman filtering algorithm and the duality established in [11], allows the framework to be expanded to solving linear quadratic regulator problems, commonly encountered in model predictive control applications [1,28]. To this end, we next consider the noise-free linear quadratic regulator problem in its distributed formulation [14].

Given the dynamic system

$$\mathbf{x}_{n+1}=\mathbf{A}\mathbf{x}_n+\sum_{\forall l\in\mathcal{N}}\mathbf{B}_l\mathbf{u}_{l,n} \tag{31}$$

with $\mathbf{B}_l$ representing the actuator matrix at node l, the goal is to find control vector sequences, $\{\mathbf{u}_{l,n} : \forall l \in \mathcal{N}, n = 1,\ldots,N\}$, that minimize the cost function

$$\mathbf{J} = \frac{1}{2}\mathbf{x}_N^{\mathsf{T}}\mathbf{T}\mathbf{x}_N + \frac{1}{2}\sum_{n=1}^{N}\left(\mathbf{x}_n^{\mathsf{T}}\mathbf{Q}\mathbf{x}_n + \mathbf{u}_{col,n}^{\mathsf{T}}\boldsymbol{\mathcal{R}}\mathbf{u}_{col,n}\right) \quad (32)$$

with $\boldsymbol{\mathcal{R}} = \text{block-diag}\{\mathbf{R}_l, l = 1,\ldots,|\mathcal{N}|\}$ and

$$\mathbf{u}_{col,n} = \left[\mathbf{u}_{1,n}^{\mathsf{T}},\ldots,\mathbf{u}_{|\mathcal{N}|,n}^{\mathsf{T}}\right]^{\mathsf{T}}$$

where N is referred to as the control horizon, while $\mathbf{T}$, $\mathbf{Q}$, and $\{\mathbf{R}_l : \forall l \in \mathcal{N}\}$ are transpose symmetric positive definite weighting matrices.

According to classical results in control theory, the optimal solution to this problem is given by [11,29]

$$\mathbf{u}_{col,n} = -\left(\boldsymbol{\mathcal{R}} + \mathbf{B}_{row}^{\mathsf{T}}\boldsymbol{\Upsilon}_{n+1}\mathbf{B}_{row}\right)^{-1}\mathbf{B}_{row}^{\mathsf{T}}\boldsymbol{\Upsilon}_{n+1}\mathbf{A}\mathbf{x}_n \quad (33)$$

where $\mathbf{B}_{row}$ is the row-block matrix given by

$$\mathbf{B}_{row} = \left[\mathbf{B}_{1,n},\ldots,\mathbf{B}_{|\mathcal{N}|,n}\right]$$

and the matrices $\{\boldsymbol{\Upsilon}_n : n = 1,\ldots,N\}$ are found from the backward propagation of the discrete time algebraic Riccati equation, given by

$$\begin{aligned}\boldsymbol{\Upsilon}_{n-1} =& \mathbf{A}^{\mathsf{T}}\boldsymbol{\Upsilon}_n\mathbf{A} + \mathbf{Q} \quad (34)\\ &- \mathbf{A}^{\mathsf{T}}\boldsymbol{\Upsilon}_n\mathbf{B}_{row}\left(\boldsymbol{\mathcal{R}} + \mathbf{B}_{row}^{\mathsf{T}}\boldsymbol{\Upsilon}_n\mathbf{B}_{row}\right)^{-1}\mathbf{B}_{row}^{\mathsf{T}}\boldsymbol{\Upsilon}_n\mathbf{A}\end{aligned}$$

with terminating condition $\boldsymbol{\Upsilon}_N = \mathbf{T}$.

The expression in (34) can be rearranged using the matrix inversion lemma in reverse to yield

$$\boldsymbol{\Theta}_{n-1} = \left(\left(\mathbf{A}^{\mathsf{T}}\boldsymbol{\Theta}_n\mathbf{A} + \mathbf{Q}\right)^{-1} + \mathbf{B}_{row}\boldsymbol{\mathcal{R}}^{-1}\mathbf{B}_{row}^{\mathsf{T}}\right)^{-1} \quad (35)$$

where $\boldsymbol{\Upsilon}_{n-1} = \mathbf{A}^{\mathsf{T}}\boldsymbol{\Theta}_n\mathbf{A} + \mathbf{Q}$. Considering the block-diagonal structure of $\boldsymbol{\mathcal{R}}$, the expression in (35) is rearranged to give

$$\boldsymbol{\Theta}_{n-1}^{-1} = \left(\mathbf{A}^{\mathsf{T}}\boldsymbol{\Theta}_n\mathbf{A} + \mathbf{Q}\right)^{-1} + \sum_{\forall l\in\mathcal{N}}\mathbf{B}_l\mathbf{R}_l^{-1}\mathbf{B}_l^{\mathsf{T}}. \quad (36)$$

On the other hand, replacing (4c) and (4d) into (4f) yields

$$\begin{aligned}\mathbf{M}_{n|n}^{-1} =& \left(\mathbf{A}_n\mathbf{M}_{n-1|n-1}\mathbf{A}_n^{\mathsf{T}} + \mathbf{C}_{\boldsymbol{\nu}_n}\right)^{-1}\\ &+ \sum_{\forall l\in\mathcal{N}}\mathbf{H}_{l,n}^{\mathsf{T}}\mathbf{C}_{\boldsymbol{\omega}_{l,n}}^{-1}\mathbf{H}_{l,n}.\end{aligned} \quad (37)$$

The duality between (36) and (37) allows the values of $\{\boldsymbol{\Upsilon}_n, \boldsymbol{\Theta}_n\}$ to be calculated through iterations of the centralized Kalman filtering operations. In turn, a distributed framework akin to the developed Kalman filtering framework can be formulated. In this setting, using the expression in (33) optimal control vector sequences can be calculated as

$$\begin{aligned}\mathbf{u}_{col,n} &= -\left(\boldsymbol{\mathcal{R}} + \mathbf{B}_{row}^{\mathsf{T}}\boldsymbol{\Upsilon}_{n+1}\mathbf{B}_{row}\right)^{-1}\mathbf{B}_{row}^{\mathsf{T}}\boldsymbol{\Upsilon}_{n+1}\mathbf{A}\mathbf{x}_n\\ &= \left(\boldsymbol{\mathcal{R}}^{-1}\mathbf{B}_{row}^{\mathsf{T}}\boldsymbol{\Upsilon}_{n+1}\mathbf{B}_{row}\boldsymbol{\mathcal{R}}^{-1} - \boldsymbol{\mathcal{R}}^{-1}\right)\mathbf{B}_{row}^{\mathsf{T}}\boldsymbol{\Upsilon}_{n+1}\mathbf{A}\mathbf{x}_n\\ &= -\boldsymbol{\mathcal{R}}^{-1}\mathbf{B}_{row}^{\mathsf{T}}\left(\boldsymbol{\Upsilon}_{n+1}^{-1} + \mathbf{B}_{row}\boldsymbol{\mathcal{R}}^{-1}\mathbf{B}_{row}^{\mathsf{T}}\right)^{-1}\mathbf{A}\mathbf{x}_n.\end{aligned} \quad (38)$$

Now, replacing (35) into (38) yields

$$\mathbf{u}_{col,n} = -\boldsymbol{\mathcal{R}}^{-1}\mathbf{B}_{row}^{\mathsf{T}}\boldsymbol{\Theta}_{n+1}\mathbf{A}\mathbf{x}_n. \quad (39)$$

Given the block-diagonal structure of $\boldsymbol{\mathcal{R}}$ and the expression in (39) it can be shown that

$$\forall l \in \mathcal{N} : \mathbf{u}_{l,n} = -\mathbf{R}_l^{-1}\mathbf{B}_l^{\mathsf{T}}\boldsymbol{\Theta}_{n+1}\mathbf{A}\mathbf{x}_n. \quad (40)$$

In addition, using the same framework as in (9)-(10) it can be shown that

$$\boldsymbol{\Theta}_n = \frac{1}{|\mathcal{N}|}\sum_{\forall l\in\mathcal{N}}\boldsymbol{\Psi}_{l,n} \quad (41)$$

with $\boldsymbol{\Psi}_{l,n} = \boldsymbol{\Upsilon}_{n-1}^{-1} + |\mathcal{N}|\mathbf{B}_l\mathbf{R}_l^{-1}\mathbf{B}_l^{\mathsf{T}}$.

In a similar fashion to what was described for the proposed distributed Kalman filtering framework, the operations of the noise-free linear quadrature regulator can now be approximated in a decentralized fashion through the distributed implementation of the average in (41). The operations of such a distributed linear quadrature regulator are summarized in Algorithm 3, where $\{\hat{\boldsymbol{\Upsilon}}_{l,n}, \hat{\boldsymbol{\Theta}}_{l,n}\}$ are local estimates of $\{\boldsymbol{\Upsilon}_n, \boldsymbol{\Theta}_n\}$ at node l.

Algorithm 3. Decentralized Linear Quadrature Regulator

For nodes $l = \{1,\ldots,|\mathcal{N}|\}$:
Initialize with:

$$\forall l \in \mathcal{N} : \hat{\boldsymbol{\Upsilon}}_{l,N} = \mathbf{T}$$

Estimate output of the Riccati equation:

$$\hat{\boldsymbol{\Psi}}_{l,n} = \hat{\boldsymbol{\Upsilon}}_{l,n}^{-1} + |\mathcal{N}|\mathbf{B}_l\mathbf{R}_l^{-1}\mathbf{B}_l^{\mathsf{T}} \quad (42a)$$

$$\hat{\boldsymbol{\Theta}}_{l,n}^{-1} \leftarrow \boxed{\text{ACF}} \leftarrow \{\forall i \in \mathcal{N} : \hat{\boldsymbol{\Psi}}_{i,n}\} \quad (42b)$$

$$\hat{\boldsymbol{\Upsilon}}_{l,n-1} = \mathbf{A}^{\mathsf{T}}\hat{\boldsymbol{\Theta}}_{l,n}\mathbf{A} + \mathbf{Q} \quad (42c)$$

Calculate control vector sequences:

$$\mathbf{u}_{l,n} = -\mathbf{R}_l^{-1}\mathbf{B}_l^{\mathsf{T}}\boldsymbol{\Theta}_{n+1}\mathbf{A}\mathbf{x}_n \quad (43a)$$

VI. Simulations

In this section, the performance of the proposed distributed Kalman filter and decentralized control framework is illustrated, where the network of 20 nodes shown in Figure 1 was used in all simulations.

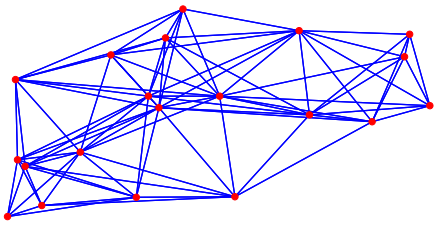

Fig. 1. The network with 20 nodes and 86 edges used in simulations.

A. Distributed Filtering and Tracking

In the first scenario, a classic target tracking problem is considered. In this setting, the state vector, $\mathbf{x}_n = [X_n, Y_n, \dot{X}_n, \dot{Y}_n]^{\mathsf{T}}$, consists of the positions, $\{X_n, Y_n\}$, and

velocities, $\{\dot{X}_n, \dot{Y}_n\}$, in the horizontal and vertical directions respectively. The state-space equations of such a dynamic system are given by

$$\mathbf{x}_n = \begin{bmatrix} 1 & 0 & \Delta T & 0 \\ 0 & 1 & 0 & \Delta T \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{x}_{n-1} + \begin{bmatrix} \frac{1}{2}(\Delta T)^2 & 0 \\ 0 & \frac{1}{2}(\Delta T)^2 \\ \Delta T & 0 \\ 0 & \Delta T \end{bmatrix} \boldsymbol{\nu}_n$$
$$\mathbf{y}_{l,n} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \mathbf{x}_n + \boldsymbol{\omega}_{l,n} \qquad (44)$$

where $\boldsymbol{\nu}_n = \left[\ddot{X}_n, \ddot{Y}_n\right]^{\mathsf{T}}$ represents unknown acceleration in the horizontal and vertical directions with $\Delta T = 0.04\,\mathrm{s}$ representing the sampling interval.

The acceleration was modeled as a white zero-mean Gaussian process with covariance matrix

$$\mathbb{E}\left\{ \begin{bmatrix} \ddot{X}_n \\ \ddot{Y}_n \end{bmatrix} \begin{bmatrix} \ddot{X}_n & \ddot{Y}_n \end{bmatrix} \right\} = \begin{bmatrix} 1.44 & 0 \\ 0 & 1.44 \end{bmatrix}$$

while the observational noise at each node was considered to be a zero-mean Gaussian vector with covariance matrix

$$\forall l \in \mathcal{N} : \mathbf{C}_{\boldsymbol{\omega}_{l,n}} = \begin{bmatrix} 4.16 & 0.8 \\ 0.8 & 4 \end{bmatrix} \times 10^{-2}.$$

Finally, at each time instant the ACF used for achieving consensus on the state estimate and its associated estimation error covariance matrix was iterated a total of 4 times.

In order to present a quantitative basis for evaluating the performance of the proposed distributed Kalman filter, we consider the mean square deviation (MSD) formulated as[3]

$$\mathrm{MSD}_l = \mathbb{E}\left\{ \boldsymbol{\epsilon}_{l,n|n} \boldsymbol{\epsilon}_{l,n|n}^{\mathsf{T}} \right\}.$$

In Figure 2, the steady-state MSD of the proposed distributed Kalman filter is benchmarked against that of the diffusion Kalman filter (Algorithm 2 in [15]), the consensus Kalman filter (Algorithm 3 in [22]), and the distributed Kalman filter in [30]. Note that the proposed algorithm achieved a steady-state MSD close to that of the centralized Kalman filter and outperformed the consensus and diffusion schemes.

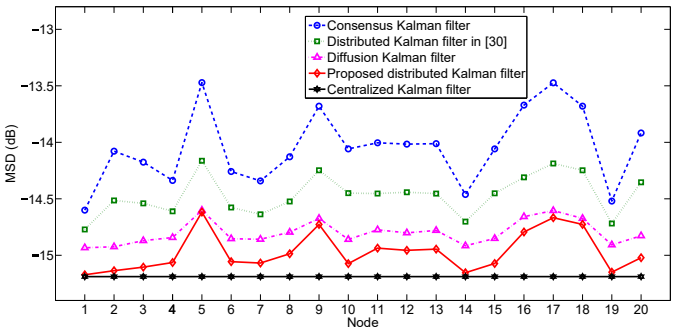

Fig. 2. Steady-state MSD performance of different distributed Kalman filtering techniques across all 20 nodes.

In the previous simulation, ACF of the proposed distributed Kalman filter was only iterated 4 times at each time instant.

[3]Note that the MSD at each node is the trace of its state vector estimation error covariance matrix obtainable in its transient and steady-state formulation from (29) and (30).

In Figure 3, the MSD performance of all nodes in the network for different number of iterations of the ACF obtained through both simulations and the theoretical framework in Section IV are shown. Note that the values obtained through the theoretical framework in Section IV precisely match those obtained through simulations, verifying the theoretical performance analysis in this work. In addition, it can be seen that the MSD performance of all nodes tend toward that of the centralized Kalman filter with increasing number of ACF iterations, where in the case that the ACF was iterated 12 times, the MSD performance of the proposed distributed Kalman filter only differs from that of the centralized Kalman filter by a maximum of 0.16 dB.

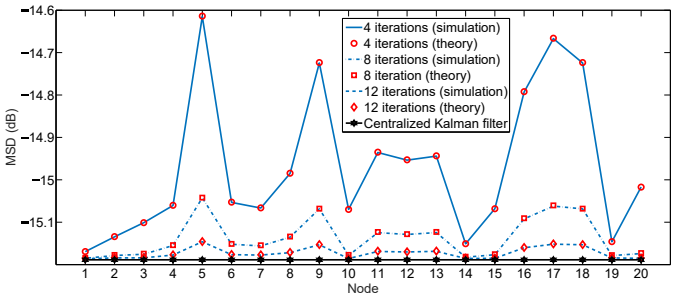

Fig. 3. Steady-state MSD performance of the proposed distributed Kalman filtering algorithm across all 20 nodes of the network, obtained with different number of iterations of the ACF. Performance of the centralized Kalman filter is provided as a benchmark.

B. Decentralized Control

The control dual of the target tracking scenario in the first simulation example was considered. Thus, in this setting, the state vector and state transition matrix remain the same as in (44) while

$$\forall l \in \mathcal{N} : \mathbf{B}_l = \begin{bmatrix} \frac{1}{2}(\Delta T)^2 & 0 & \Delta T & 0 \\ 0 & \frac{1}{2}(\Delta T)^2 & 0 & \Delta T \end{bmatrix}^{\mathsf{T}}$$

and control inputs of the agents are taken to be acceleration (or force per unit mass). The weighting matrices in the cost function in (32) were $\mathbf{Q} = \mathbf{I}$, $\mathbf{T} = 150 \times \mathbf{Q}$, and

$$\forall l \in \mathcal{N} : \mathbf{R}_l = l^2 \times \begin{bmatrix} 2.16 & 1.8 \\ 1.8 & 2 \end{bmatrix}.$$

The goal was to bring the object to a stand-still at the center of the coordinate system from an initial position and speed. To this end, the proposed decentralized linear quadrature regulator control procedure in Algorithm 3 and its centralized dual were implemented over the network in Figure 1. Control vector sequences were estimated for 0.8 s long segments. The first 0.4 s portion of the estimated control vector sequences were implemented; then, control vector sequences were re-estimated using the new position and speed of the object. This procedure was repeated to achieve the desired goal. The position and speed of the object are shown in Figure 4. Note that the developed decentralized control framework operated correctly. Furthermore, the object followed similar position and speed trajectories when using the centralized and developed decentralized control frameworks.

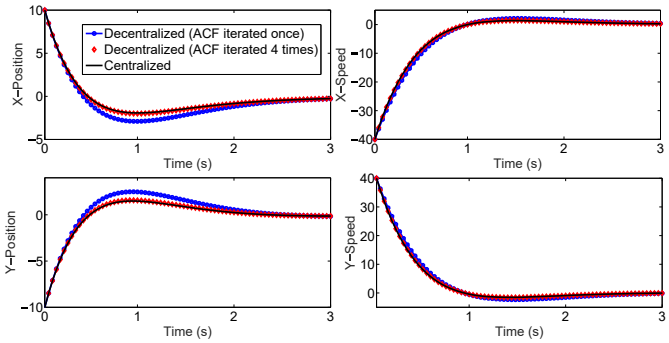

Fig. 4. Performance of the proposed decentralized linear quadrature regulator with one iteration on the ACF is shown alongside the performance of the proposed decentralized linear quadrature regulator with four iterations on the ACF and the performance of the centralized linear quadrature regulator.

VII. Conclusion

A distributed Kalman filtering algorithm for estimation and tracking applications over sensor networks has been developed. The developed algorithm has been realized through the decomposition of the operations of the centralized Kalman filter using embedded average consensus filters. The operations of the centralized Kalman filter have been decomposed in a fashion that allows each agent of the sensor network to retain an estimate of the state vector and its estimation error covariance matrix. In contrast to current distributed Kalman filtering frameworks, the developed algorithm does not require the exchange of observation data over the network, relaxing communication and computational requirements. The mean and mean square error performance of the developed distributed Kalman filter have been analyzed establishing that the developed algorithm operates in an unbiased fashion and can reach mean square error performance levels comparable to that of the centralized Kalman filter. Moreover, it has been shown that the comprehensive fusion of covariance information established for the developed distributed Kalman filter coupled with the duality between Kalman filters and linear quadrature regulators, allow the framework to also be used in decentralized control applications, providing a rigorous framework for extending various concepts in control to the distributed setting.

References

[1] J. S. Shamma, *Cooperative Control of Distributed Multi-Agent Systems*. John Wiley & Sons, 2007.

[2] R. Olfati-Saber, "Flocking for multi-agent dynamic systems: Algorithms and theory," *IEEE Transactions on Automatic Control*, vol. 51, no. 3, pp. 401–420, March 2006.

[3] S. P. Talebi, S. Kanna, and D. P. Mandic, "A distributed quaternion Kalman filter with applications to smart grid and target tracking," *IEEE Transactions on Signal and Information Processing over Networks*, vol. 2, no. 4, pp. 477–488, December 2016.

[4] S. Xu, K. Dogancay, and H. Hmam, "Distributed pseudolinear estimation and UAV path optimization for 3D AOA target tracking," *Signal Processing*, vol. 133, no. Supplement C, pp. 64–78, 2017.

[5] O. Hlinka, F. Hlawatsch, and P. M. Djuric, "Distributed particle filtering in agent networks: A survey, classification, and comparison," *IEEE Signal Processing Magazine*, vol. 30, no. 1, pp. 61–81, January 2013.

[6] A. G. O. Mutambara, *Decentralized Estimation and Control for Multisensor Systems*. CRC Press, 1998.

[7] E. Franco, R. Olfati-Saber, T. Parisini, and M. M. Polycarpou, "Distributed fault diagnosis using sensor networks and consensus-based filters," *In Proceedings of IEEE International Conference on Decision and Control*, pp. 386–391, December 2006.

[8] F. Cattivelli and A. H. Sayed, "Distributed nonlinear Kalman filtering with applications to wireless localization," *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 3522–3525, March 2010.

[9] A. H. Sayed, "Adaptation, learning, and optimization over networks," *Foundations and Trends® in Machine Learning*, vol. 7, no. 4-5, pp. 311–801, 2014.

[10] M. S. Mahmoud and H. M. Khalid, "Distributed Kalman filtering: A bibliographic review," *IET Control Theory Applications*, vol. 7, no. 4, pp. 483–501, March 2013.

[11] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Transactions of the ASME Journal of Basic Engineering*, vol. 82, Series D, pp. 35–45, 1960.

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

[13] Y. Zhang, C. Wang, N. Li, and J. Chambers, "Diffusion Kalman filter based on local estimate exchanges," *In Proceedings of IEEE International Conference on Digital Signal Processing*, pp. 828–832, July 2015.

[14] J. Speyer, "Computation and transmission requirements for a decentralized linear-quadratic-Gaussian control problem," *IEEE Transactions on Automatic Control*, vol. 24, no. 2, pp. 266–269, April 1979.

[15] 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, September 2010.

[16] M. Alighanbari and J. P. How, "An unbiased Kalman consensus algorithm," *In Proceedings of American Control Conference*, pp. 3519–3524, June 2006.

[17] W. Li, G. Wei, F. Han, and Y. Liu, "Weighted average consensus-based unscented Kalman filtering," *IEEE Transactions on Cybernetics*, vol. 46, no. 2, pp. 558–567, February 2016.

[18] R. Olfati-Saber, "Kalman-consensus filter: Optimality, stability, and performance," *In Proceedings of the 48th IEEE Conference on Decision and Control*, pp. 7036–7042, December 2009.

[19] F. Pasqualetti, S. Zampieri, and F. Bullo, "Controllability metrics, limitations and algorithms for complex networks," *IEEE Transactions on Control of Network Systems*, vol. 1, no. 1, pp. 40–52, March 2014.

[20] F. Dorfler, J. W. Simpson-Porco, and F. Bullo, "Breaking the hierarchy: Distributed control and economic optimality in microgrids," *IEEE Transactions on Control of Network Systems*, vol. 3, no. 3, pp. 241–253, September 2016.

[21] R. R. Negenborn and J. M. Maestre, "Distributed model predictive control: An overview and roadmap of future research opportunities," *IEEE Control Systems*, vol. 34, no. 4, pp. 87–97, August 2014.

[22] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," *In Proceedings of the 46th IEEE Conference on Decision and Control*, pp. 5492–5498, December 2007.

[23] M. Kamgarpour and C. Tomlin, "Convergence properties of a decentralized Kalman filter," *In Proceedings of the 47th IEEE Conference on Decision and Control*, pp. 3205–3210, December 2008.

[24] J. Hu, L. Xie, and C. Zhang, "Diffusion Kalman filtering based on covariance intersection," *IEEE Transactions on Signal Processing*, vol. 60, no. 2, pp. 891–902, February 2012.

[25] L. Xiao, S. Boyd, and S. Lall, "A scheme for robust distributed sensor fusion based on average consensus," *In Proceedings of Fourth International Symposium on Information Processing in Sensor Networks*, pp. 63–70, 2005.

[26] L. Xiao and S. Boyd, "Fast linear iterations for distributed averaging," *Systems & Control Letters*, vol. 53, no. 1, pp. 65–78, 2004.

[27] T. Kailath, A. H. Sayed, and B. Hassibi, *Linear Estimation*. Prentice Hall, 2000.

[28] R. R. Bitmead, M. Gevers, and V. Wertz, *Adaptive Optimal Control: The Thinking Man's GPC*. Prentice Hall, New Jersey, 1990.

[29] R. F. Stengel, *Optimal Estimation and Control*. Dover Publications, 1994.

[30] S. P. Talebi, S. Kanna, Y. Xia, and D. P. Mandic, "Cost-effective diffusion Kalman filtering with implicit measurement exchanges," *In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 4411–4415, 2017.