Distributed and Decentralized State Estimation Applied to Complex Epidemic Networks

Anny Verly [*,**] **Bruno Otávio Soares Teixeira** [**]

[*] *Department of Electrical Engineering, Universidade Federal de Ouro Preto, Rua 36 n 115 ,35931-008, João Monlevade, MG, Brazil (e-mail: annyverly@ufop.edu.br)*
[**] *Graduate Program in Electrical Engineering, Universidade Federal de Minas Gerais, Av. Antônio Carlos 6627, 31270-901, Belo Horizonte, MG, Brazil (e-mail: brunoot@ufmg.br)*

Abstract: In epidemic models, the traditional approaches assume that each individual has an equal chance, per unit of time, to communicate with each other. In this regard, the use of complex networks can be considered a more realistic approach. Indeed, in epidemic complex networks the contact patterns are taken into account in the study of an epidemic spreading. However, due to the high dimensionality of the state and observation equations, the use of a classical centralized strategy for state estimation is a challenge. For this reason, as a preliminary study, we propose the use of distributed and decentralized information filter in order to overcome this issue. In a simulation example, for a susceptible-infected network, we show that the distributed and decentralized information filter is effective in state estimation as the centralized approach.

Keywords: State Estimation; Distributed Kalman Filter; Large-Scale Systems; Epidemics on Dynamical Networks; Complex Systems.

1. INTRODUCTION

The classical dynamical analysis of epidemic systems, known as *fully mixed models*, does not consider the spatial interactions among their elements. The accuracy of the epidemic spreading models can be improved considering the spatial interactions, when integrating large-scale data sets. For this reason, large-scale networks have received increasing attention for the community of epidemic process (Pastor-Satorras et al., 2015; Pinto et al., 2020).

The derivation of network models for epidemic spreading processes relies on different theoretical approaches. These approaches are based mainly on three field theories, known as the degree-based mean-field approach (DBMF), individual-based mean-field approach (IBMF), and generating function approach (GFA). According to Pastor-Satorras et al. (2015), the DBMF theory was the first theoretical approach proposed for the analysis of dynamical complex networks and does not explicitly contain neighborly relationships between individuals from a population. Thus, such relationships are expressed in the form of probability of contact between individuals (Barrat et al., 2008; Wang et al., 2020). The IBMF theory (Pastor-Satorras et al., 2015; Schaum and Jaquez, 2016) represents a model simplification of epidemic spreading in networks, in which the relation between the individuals are directly modeled by means of the adjacency matrix. In addition, the GFA (Pastor-Satorras et al., 2015) is used to model the late-time properties of epidemic outbreaks.

According to Newman (2010), a large-scale network can be understood as a high dimension system composed of a set of interconnected nodes by means of links, in which there is interaction between the connected nodes. Depending on the different topological structures from networks, there are differences in the dynamics of epidemic transmission (Wang et al., 2020). In regular networks or statistically homogeneous networks, each node has approximately the same number of links. However, most real-world networks share similar complex topological characteristics, which are statistically heterogeneous and described by *heavy-tailed* statistical distributions (Boccaletti et al., 2006; Barabási, 2009). This means that complex networks are heterogeneous in nature, that is, most nodes have few links and few nodes have many connections or links. In this last case, the large-scale networks are called *scale-free* and a classical graph to generate this kind of network is the Barabási and Albert (1999) (BA) model.

The classical state estimation, using Kalman filter (KF) (Kalman, 1960) or the extended Kalman filter (EKF) (Jazwinski, 1970) for nonlinear systems, yields better predictions than those produced by the free-running simulations of the model. When the interest for state estimation rises, most of the research efforts on epidemics spreading are concerned with fully mixed models (Yang et al., 2014; Younes and Hasan, 2020).

Recently, Wang et al. (2020) addressed the state estimation problem on homogeneous epidemic networks. They used nonlinear filters, such as EKF and unscented Kalman

filter (UKF) in susceptible infected susceptible (SIS), susceptible infected recovered (SIR), and susceptible infected recovered susceptible (SIRS) models, all these based on DBMF theory, and presented a performance comparison between the filters. The authors used a centralized state estimation architecture and the high dimensionality problem for large-scale networks was no directly attacked.

Schaum and Jaquez (2016) dealt with the aforementioned problem by reducing the dimensionality of the observation equation. The authors sought to find which of the states of the complex network must be monitored to ensure the convergence of the estimator. Thus, by using a centralized architecture, they proposed a detectability-based approach for state reconstruction, in which the number of monitored nodes is reduced based on the network topology. The detectability condition and the observer design were implemented over a complex SIS network based on the IBMF theory.

The great challenge in the development of state estimation algorithms for large-scale networks is the high dimensionality of the global model. In this case, the centralized state estimation using the Kalman filter does not provide robustness and scalability. Several approaches based on distributed state estimation methods have been proposed in the multisensor networks field (Viegas et al., 2018; Deshmukh et al., 2017; Cattivelli and Sayed, 2010), as an alternative to centralized state estimation. However, most of these researches are focused on large-scale sensor networks for monitoring low-dimension systems. Indeed, usually, a scalable solution is addressed based only on the observation equation and not on the state transition matrix (Cicala et al., 2020). In order to overcome this issue, an earlier study (Mutambara, 1998) inspired Cicala et al. (2020) to propose the distributed and decentralized information filter (DDEIF) algorithm combined with a consensus strategy using the internodal transformation theory on the state estimation. It is important to point out that the DDEIF is a variation of the distributed and decentralized extended Kalman filter (DDEKF) algorithm (Mutambara, 1998).

Given the aforementioned problems, the following question arises: *How the distributed state estimation problem can be adapted from the multisensor system theory for large-scale epidemic networks?* In this paper, we address this issue merging part of the epidemic network theory (Pastor-Satorras et al., 2015) with the internodal transformation theory (Mutambara, 1998) to obtain scalable solutions for a class of heterogeneous epidemic models based on the IBMF theory. In addition, with the aim of eliminating the need of monitoring all network nodes, the previous results (Schaum and Jaquez, 2016) have inspired us to make Monte Carlo simulations to determine a reduced number of observed nodes that do not affect the effectiveness of the state estimation methods. Thus, our contribution is to give a new direction on the state estimation of epidemic spreading in complex networks.

This paper is organized as follows: First, the problem formulation is given in Section 2. In Section 3, the extended information filter (EIF) algorithm and the DDEIF are revisited. In Section 4, we presented the susceptible infected (SI) network model in global and local or scalable version.

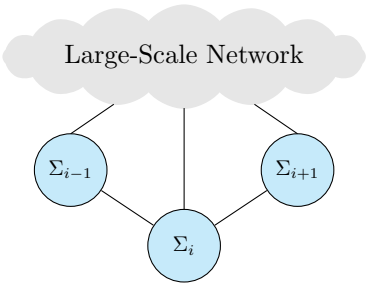

Figure 1. Illustration of a complex network.

In Section 5, we present the results of the employment of the EIF and the DDEIF in a new perspective. Finally, in Section 6, the conclusions are discussed.

2. PROBLEM FORMULATION

Consider an epidemic system modeled by a complex network of N subsystems, or nodes, that are interconnected as in Figure 1, where each node is denoted by Σ_i with $i = \{1, 2, \ldots, N\}$. Let $\mathcal{N}_i$ be the set of nodes connected to Σ_i, known as *neighborhood* of the i-th node. The *adjacency* matrix $\mathcal{A}$ represents how all N nodes are interconnected, and their elements are given by:

$$\mathcal{A}_{ij} \triangleq \begin{cases} 1 \text{ if } \mathrm{j} \in \mathcal{N}_\mathrm{i} \\ 0 \text{ if } \mathrm{j} \notin \mathcal{N}_\mathrm{i} \end{cases}. \tag{1}$$

To derive a generic epidemic network model the following considerations are necessary (Newman, 2010; Pastor-Satorras et al., 2015):

Assumption 1. The nodes are the individuals of a population, and the links represent the relationships among them.

Assumption 2. Each individual i has associated state probabilities, such as probabilities of being susceptible and infected. Furthermore, the network nodes have the same number of internal states.

Assumption 3. The dynamical function at each node are considered the same.

Assumption 4. The network is undirected, so the matrix $\mathcal{A}$ is symmetric.

The modeling equations of Σ_i could be written as (Newman, 2010):

$$\frac{dx_i(t)}{dt} = \mathcal{F}\left(x_i(t)\right) + \sum_{j \in \mathcal{N}_i} \mathcal{A}_{ij} \mathcal{G}\left(x_i(t), x_j(t)\right) + \omega_i(t), \quad i = 1, \ldots, N, \tag{2}$$

$$y_p(t) = \mathcal{C}_p x_{i_p}(t) + \upsilon_p(t), i_p \in \mathcal{M}, p = 1, \ldots, m, \tag{3}$$

where $x_i(t) \in \Re^n$ is a vector of n variables with $x_i(t) = [x_i^1(t)\ x_i^2(t) \ldots x_i^n(t)]^T$; $\mathcal{F}$ is the intrinsic dynamics of the i-th node, that is, it indicates how the vector $x_i(t)$ evolves without neighbors; $\mathcal{G}$ is the coupling function between variables from different nodes; $\omega_i(t) \in \Re^n$ is a zero-mean white Gaussian process noise; $y_p(t) \in \Re^l$ is a vector of the p-th observation which measures the state vector of the node i_p; l represents the measurements available per monitored node with $y_p(t) = [y_p^1(t)\ y_p^2(t)\ \ldots\ y_p^l(t)]^T$; $\mathcal{C}_p \in \Re^{l \times n}$ is the local measurement matrix; $\mathcal{M}$ is the set of all monitored nodes where $m \leq N$ and $\upsilon_p(t) \in \Re^l$ is

a zero-mean white Gaussian measurement noise. We can highlight that i_p represents the index of the monitored node corresponding to the observation p.

Remark 1. The dimensionality of an epidemic network scales with the increase of the number of nodes and their state dimension. For instance, let us consider a network defined by its adjacency matrix $\mathcal{A}$ and a general epidemic process with q-state compartmentalizations; see Figure 1. Recall that, each individual could be in a particular state, and these states are often called compartments. The SI model, for example, is called *two*-state ($q = 2$) model. The states of the i-th node at time t are specified by the random variable $X_i(t) \in \{0, 1, ..., q-1\}$, where $X_i(t)$ can take a value from 0 to $q-1$; that is, it means that i-th node belongs to one of these compartments at time t (Pastor-Satorras et al., 2015). Assuming that the dynamic state of every node is statistically independent of the state of its nearest neighbors (Pastor-Satorras et al., 2015) and that the network has a constant population (N nodes), the total number of states that need to be computed is equal to $(q-1)N$. If the initial infection probabilities are known, then the state probabilities $\Pr[X_i(t) = \alpha_j]$ at the time t, for each $i = 1, \ldots, N$ and $\alpha_j = 0, 1, \ldots, q-1$, can be obtained by solving $(q-1)N$ differential equations with a parameter (linearized) set of length $(q-1)N \times (q-1)N$. For notational simplicity, we consider that the state probabilities are represented by the vector $x_i(t)$ in (2).

The network presented in Figure 1 can be discretized with fixed step size t_s and written as a global function of the discrete-time instant (kt_s). Thus, the compact vector notation used in the filtering problem can be written as:

$$x_k = f_{k-1}(x_{k-1}) + \omega_{k-1}, \tag{4}$$

$$y_k = C_k x_k + \upsilon_k, \tag{5}$$

where the term t_s is omitted to simplify the global representation; f_{k-1} represents the global state transition function and $C_k \in \Re^{lm \times nN}$ is the global measurement matrix defined according to the set $\mathcal{M}$. The global state vector $x_k \in \Re^{nN}$ is related to (2) by means of a discretization and stacking procedure of the state vector of each node, as $x_k = \left[(x_{1,k})^T \ \ldots \ (x_{N,k})^T\right]^T$. Also, $y_k \in \Re^{lm}$ are the stacked vectors given by $y_k = \left[(y_{1,k})^T \ \ldots \ (y_{m,k})^T\right]^T$. Similarly, $\omega_{k-1} \in \Re^{nN}$ and $\upsilon_k \in \Re^{lm}$ are, respectively, the stacked vectors of process and measurement noise. We assume that these are zero-mean white Gaussian noises and satisfy the covariance matrices $E\left[\omega_a \omega_b^T\right] = \delta_{a,b} Q_a$, with $Q_a \in \Re^{nN \times nN}$, $E\left[\upsilon_a \upsilon_b^T\right] = \delta_{a,b} R_a$, with $R_a \in \Re^{lm \times lm}$, and $E\left[\omega_a \upsilon_b^T\right] = 0$, where $E[\bullet]$ is the expected value, δ is the Kronecker delta function and a, b time instants.

Our purpose is, given the measurements (3) (or equivalently (5)) to estimate the states of the node x^i from a network represented by (2) (or equivalently (4)). Regarding the state vector dimension, there is a gap in the state estimation of epidemic networks due to the high dimensionality of states. In addition, most of the research efforts in this issue deal with modeling problems. However, choosing an appropriate filtering method remains a challenge. To deal with the above-mentioned problems, we propose an alternative scalable solution by means of the DDEIF (Mutambara, 1998). In addition, differently from Schaum and Jaquez (2016), we propose to solve the problem of choosing m, i.e. the number of monitored nodes in epidemic networks, by means of Monte Carlo simulations.

Remark 2. The state vector dimension in a state estimation problem depends on the proposed filter architecture. If the state estimation performs a classical centralized solution, then the state vector is the global one, x_k (4). If we wish to obtain only $x_i(t)$ (2) in discrete-time by means of a scalable solution, then we need to use tools for decentralization and distribution of the global model, as described in Section 3.2.

3. METHODS FOR STATE ESTIMATION

We propose to solve the problem presented in Section 2 by using two types of state estimation architectures, namely: centralized and decentralized. Usually, the centralized filters use the global network model in a central processor with direct connections to all sensor devices. This approach, when applied to large-scale networks, has processing limitations. In contrast, the decentralized strategies allow all measurements and state estimates to be processed locally. Thus, each network node has its own processor, which can use both global and distributed models. In this paper, the term *distributed* refers to the use of scalable models, where the local models involve only locally relevant states to each node and the local filters perform a n_i reduced-order operation, with $n_i \ll n \times N$ (Mutambara, 1998; Cicala et al., 2020).

3.1 Centralized Estimation

The Extended Information Filter

According to Mutambara (1998), the EIF is derived from EKF in terms of information measures about the states rather than direct estimates about states and their associated covariances. The *information matrix*, Z, is defined as the inverse of the covariance matrix, $Z \triangleq P^{-1}$, and the *information state vector* as the product of the inverse of the covariance matrix and the state estimate, $\widehat{z} \triangleq P^{-1}\widehat{x}$.

Similar to the EKF that uses the global state transition (4) in the forecast step, the EIF provides the *one-step-ahead* forecast estimate, $\widehat{x}_{k|k-1}$, where $\widehat{x}_{k|k-1} = E[x_k | y_1, y_2, \ldots, y_{k-1}]$. Also, the covariance matrix is given by $Z^{-1}_{k|k-1} = P_{k|k-1} = E[(x_k - \widehat{x}_{k|k-1})(x_k - \widehat{x}_{k|k-1})^T | y_1, y_2, \ldots, y_{k-1}]$. In the *data-assimilation step* the EIF provides a recursive estimate $\widehat{x}_k$ at time k, by means of $\widehat{z}_k$, given all information up to time k, where $\widehat{x}_k = E[x_k | y_1, y_2, \ldots, y_k]$ and $Z_k^{-1} = P_k = E[(x_k - \widehat{x}_k)(x_k - \widehat{x}_k)^T | y_1, y_2, \ldots, y_k]$.

The information state contribution $\mathbf{i}_k$ obtained from an observation y_k, and its associated information matrix $\mathcal{I}_k$ are defined, respectively, as follows:

$$\mathbf{i}_k \triangleq C_k^T R_k^{-1} y_k, \tag{6}$$

$$\mathcal{I}_k \triangleq C_k^T (R_k)^{-1} C_k. \tag{7}$$

Let us consider F_{k-1} the Jacobian of f_{k-1} in (4) evaluated at $\widehat{x}_{k-1}$. The EIF runs the EKF in the information space, as described in the Algorithm 1. Although the EIF is algebraically equivalent to EKF, EIF has advantages in decentralization and distribution of state estimates. Due

to the high dimension of the network, such characteristics are relevant during the filtering procedure.

Algorithm 1 The Extended Information Filter

Consider the state space model given by (4)-(5). Define $\widehat{x}_0$ and Z_0, for every time instant k, compute:

Step 1: Forecast

$$Z_{k|k-1} = \left[F_{k-1}\left(Z_{k-1}\right)^{-1}\left(F_{k-1}\right)^{T} + Q_{k-1}\right]^{-1},$$
$$\widehat{x}_{k|k-1} = f_{k-1}\left(\widehat{x}_{k-1}\right),$$
$$\widehat{z}_{k|k-1} = Z_{k|k-1}\widehat{x}_{k|k-1},$$

Step 2: Data-Assimilation

$$\widehat{z}_k = \widehat{z}_{k|k-1} + \mathbf{i}_k,$$
$$Z_k = Z_{k|k-1} + \mathcal{I}_k,$$

where $\mathbf{i}_k$ is given by (6) and $\mathcal{I}_k$ is given by (7).

3.2 Distributed and Decentralized Estimation

To proceed with the distributed and decentralized state estimation strategy we must decentralize the information and distribute the global state vectors in local state vectors. So the states or the information states can be estimated by means of internodal transformation theory, as described as follows (Mutambara, 1998).

Decentralizing the Observer

The measurement equation (5) can be represented in a decentralized way as a function of x_k, as follows:

$$y_{p,k} = C_{p,k}x_k + \upsilon_{p,k},\ p = 1, \ldots, m, \tag{8}$$

where $y_{p,k} \in \Re^{l}$ and $\upsilon_{p,k} \in \Re^{l}$; $C_{p,k} \in \Re^{l\times nN}$ is related to C_k in (5) by means of a stacking procedure, as $C_k = \left[\left(C_{1,k}\right)^{T} \ldots \left(C_{m,k}\right)^{T}\right]^{T}$. We also assume that the observation noises are independent between the nodes, so the global covariance matrix of the observation noises can be written as $R_k = \text{diag}\left[R_{1,k}, R_{2,k}, \ldots, R_{m,k}\right]$, where diag is the operator of creation a diagonal matrix and $R_{p,k} \in \Re^{l\times l}$.

Thus, (6) and (7) can be expressed, respectively, by a linear combination of the local information state contributions $\mathbf{i}_p$ and the local information matrices associated $\mathcal{I}_p$, according to:

$$\mathbf{i}_k = \sum_{p=1}^{m}\mathbf{i}_{p,k} \triangleq \sum_{p=1}^{m}\left(C_{p,k}\right)^{T}\left(R_{p,k}\right)^{-1}y_{p,k}, \tag{9}$$

$$\mathcal{I}_k = \sum_{p=1}^{m}\mathcal{I}_{p,k} \triangleq \sum_{p=1}^{m}\left(C_{p,k}\right)^{T}\left(R_{p,k}\right)^{-1}C_{p,k}. \tag{10}$$

Remark 3. Usually, the network using this type of measurement architecture is a *fully connected network*. Where the observations from each measured node are shared among all local processors.

Model Distribution

Following some procedures (Berg and Durrant-Whyte, 1991), it is possible to obtain a local state vector with only locally relevant states. In this way, there is a dynamic equivalence between the local and global models (Mutambara, 1998). The local state vector from the i-th node, $x_{i,k}$, is related to global state vector, x_k, by means of:

$$x_{i,k} = \mathcal{T}_{i,k}x_k, \tag{11}$$

where $\mathcal{T}_{i,k} \in \Re^{n_i\times nN}$ is the nodal transformation matrix with n_i necessary states to estimate $\widehat{x}_{i,k}$. The idea is to identify which nodes should be locally known and then distribute the global model by means of $\mathcal{T}_{i,k}$. Thus, if the global state transition matrix F_{k-1} and the adjacency matrix $\mathcal{A}$ in the time k are known, the matrix $\mathcal{T}_{i,k}$ can be constructed by null and unitary indexes, where the unitary indexes correspond to the neighbors of the i-th node (Mutambara, 1998).

From (11) we get $x_k = \left(\mathcal{T}_{i,k}\right)^{\dagger}x_{i,k}$, where $(\bullet)^{\dagger}$ is the Moore-Penrose generalized inverse. $\left(\mathcal{T}_{i,k}\right)^{\dagger}$ reconstructs all states locally relevant to global space. Thus, by using (4) we can derive a reduced order state equation to the i-th node, as follows:

$$x_{i,k} = \mathcal{T}_{i,k}f_{k-1}\left(\left(\mathcal{T}_{i,k-1}\right)^{\dagger}x_{i,k-1}\right) + \omega_{i,k-1},\ i = 1, \ldots, N, \tag{12}$$

where $x_{i,k} \in \Re^{n_i}$ and $\omega_{i,k-1} = \mathcal{T}_{i,k}\omega_{k-1} \in \Re^{n_i}$. Also, $E\left[\omega_{i,a}\left(\omega_{j,b}\right)^{T}\right] = Q_{i,a}\delta_{a,b}\delta_{i,j}$ with $Q_{i,a} \in \Re^{n_i\times n_i}$.

Remark 4. Although each node has n states directly related to Σ_i, as described in (2), the scalability condition establishes that the state vector of the i-th node is composed by its directly connected neighbors. Thus, the local state vector $x_{i,k}$ in (12) is the local state vector x_i in (2), in discrete-time, expanded to contain the neighborhood of the i-th node, where $n_i > n$.

The matrix $\mathcal{C}_p$ in (3) is also expand to composed the neighborhood of the p-th measured node. Thus, the distributed observation equation, in the discrete-time version, can be written as:

$$y_{p,k} = \mathcal{C}_{p,k}x_{i_p,k} + \upsilon_{p,k},\ i_p \in \mathcal{M},\ p = 1, \ldots, m, \tag{13}$$

where $\mathcal{C}_{p,k} \in \Re^{l\times n_i}$. We must highlight that (13) is obtained from the local state vector $x_{i_p,k}$, the partitioned observation matrix $C_{p,k}$, in (8), is related to $\mathcal{C}_{p,k}$ by means of the transformation $C_{p,k} = \mathcal{C}_{p,k}\mathcal{T}_{i_p,k}$ (Mutambara, 1998).

Internodal Transformation

The local information matrix and the information state vector in the i-th node can be defined as (Mutambara, 1998):

$$Z_{i,k|k-1} \triangleq \left(P_{i,k|k-1}\right)^{-1}, \tag{14}$$

$$\widehat{z}_{i,k|k-1} \triangleq \left(P_{i,k|k-1}\right)^{-1}\widehat{x}_{i,k|k-1}. \tag{15}$$

Let us first deal with the state estimation problem of the i-th node based on the observations of the j-th node. The information contribution from the i-th node due to current observations from j-th node is defined by $\mathbf{i}_{i,k|y_{j,k}}$. Also, the global information contribution due to global observation is defined by $\mathbf{i}_{k|y_k} = \mathbf{i}_k$. Analogously, the associated local information matrix is given by $\mathcal{I}_{i,k|y_k^j}$ and the global associated information matrix is $\mathcal{I}_{k|y_k} = \mathcal{I}_k$. Thus, the local error covariance at the i-th node based only on current observations from the j-th node is defined as:

$$P_{i,k|y_{j,k}} \triangleq \left[\mathcal{I}_{i,k|y_{j,k}}\right]^{\dagger}. \tag{16}$$

In addition, the local state estimate at the i-th node based only on current observations from the j-th node can be defined as:

$$\widehat{x}_{i,k|y_{j,k}} \triangleq P_{i,k|y_{j,k}} \mathbf{i}_{i,k|y_{j,k}}. \tag{17}$$

In the information space, given only the observations from the j-th node, $y_{j,k}$, we need to transform the information contribution and its associated matrix from the j-th node to the corresponding information contribution and associated matrix at the i-th node, or even (Mutambara, 1998):

$$\mathbf{i}_{j,k|y_{j,k}} \longmapsto \mathbf{i}_{i,k|y_{j,k}}, \forall j \neq i, \tag{18}$$

$$\mathcal{I}_{j,k|y_{j,k}} \longmapsto \mathcal{I}_{i,k|y_{j,k}} \forall j \neq i. \tag{19}$$

Figure 2 shows the transformation steps performed in the information space to transform from the subspace information contribution at the j-th node , $\mathbf{i}_{j,k|y_{j,k}}$, to the subspace relative to the i-th node, $\mathbf{i}_{i,k|y_{j,k}}$, where:

$$\mathcal{T}_{ji,k} = \mathcal{I}_{i,k|y_{j,k}} \mathcal{V}_{ji,k} \left[\mathcal{I}_{j,k|y_{j,k}}\right]^{\dagger} \tag{20}$$

is the information space internodal transformation matrix that maps the relevant information from the j-th information subspace to the i-th information subspace, and

$$\mathcal{V}_{ji,k} = \mathcal{T}_{i,k} \left(\mathcal{T}_{j,k}\right)^{\dagger} \tag{21}$$

is the state space internodal transformation matrix that maps the estimated states from the j-th state subspace to the i-th state subspace.

Finally, the solutions for (18) and (19) can be derived to obtain (Mutambara, 1998; Berg and Durrant-Whyte, 1991):

$$\mathbf{i}_{i,k|y_{j,k}} = \mathcal{T}_{ji,k} \mathbf{i}_{j,k|y_{j,k}}, \tag{22}$$

and

$$\begin{aligned} \mathcal{I}_{i,k|y_{j,k}} &= \left[P_{i,k|y_{j,k}}\right]^{\dagger} \\ &= \left\{\mathcal{T}_{i,k}\left[\left(\mathcal{T}_{j,k}\right)^{T} \mathcal{I}_{j,k|y_{j,k}} \mathcal{T}_{j,k}\right]^{\dagger} \left(\mathcal{T}_{i,k}\right)^{T}\right\}^{\dagger}, \end{aligned} \tag{23}$$

where $\mathbf{i}_{j,k|y_{j,k}}$ and $\mathcal{I}_{j,k|y_{j,k}}$ are, respectively, the information contribution and its associated matrix calculated locally from local measures and given by:

$$\mathbf{i}_{j,k|y_{j,k}} \triangleq \left(C_{j,k}\right)^{T} \left(R_{j,k}\right)^{-1} y_{j,k}, \tag{24}$$

$$\mathcal{I}_{j,k|y_{j,k}} \triangleq \left(C_{j,k}\right)^{T} \left(R_{j,k}\right)^{-1} C_{j,k}. \tag{25}$$

The matrix $C_{j,k}$ is obtained according to the set $\mathcal{M}$.

The Distributed and Decentralized Extended Information Filter

Let us consider $F_{i,k-1}$ the Jacobian of f_{k-1} in (4) evaluated at $\widehat{x}_{i,k-1}$. By using the scalability and internodal transformation relationships, we can present the DDEIF described in the Algorithm 2 (Mutambara, 1998; Cicala et al., 2020).

Remark 5. In data-assimilation step, in order to perform the transformations (20)-(23), the i-th node must receive from the others j measured nodes the local information $\mathbf{i}_{j,k|y_{j,k}}$, $\mathcal{I}_{j,k|y_{j,k}}$ and $\mathcal{T}_{j,k}$. This means that all nodes that are neighbors, even if they are indirectly connected, must have a fully connected estimation architecture.

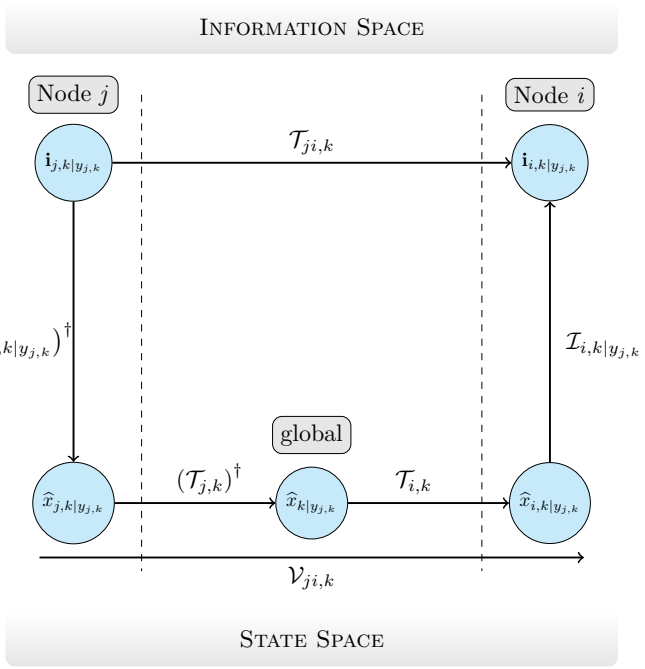

Figure 2. Information space transformation. Adapted from Mutambara (1998).

Algorithm 2 The Distributed and Decentralized Extended Information Filter

Consider the state transition equation for the i-th node given by (12) and the measured equations given by (8). Define $\widehat{x}_{i,0}$ and $Z_{i,0}$, for every time instant k, compute:

Step 1: Forecast

$$Z_{i,k|k-1} = \left[F_{i,k-1}\left(Z_{i,k-1}\right)^{-1}\left(F_{i,k-1}\right)^{T} + Q_{i,k-1}\right]^{-1},$$

$$\widehat{x}_{i,k|k-1} = \mathcal{T}_{i,k} f_{k-1}\left(\left(\mathcal{T}_{i,k-1}\right)^{\dagger} \widehat{x}_{i,k-1}\right),$$

$$\widehat{z}_{i,k|k-1} = Z_{i,k|k-1} \widehat{x}_{i,k|k-1},$$

Step 2: Data-Assimilation

$$\widehat{z}_{i,k} = \widehat{z}_{i,k|k-1} + \sum_{j=1}^{m} \mathbf{i}_{i,k|y_{j,k}},$$

$$Z_{i,k} = Z_{i,k|k-1} + \sum_{j=1}^{m} \mathcal{I}_{i,k|y_{j,k}},$$

where $\mathbf{i}_{i,k|y_{j,k}}$ and $\mathcal{I}_{i,k|y_{j,k}}$ are obtained by the transformations (20)-(23). When $m \neq N$, $\mathbf{i}_{j|y_j}$ and $\mathcal{I}_{j|y_j}$, given by (24) and (25), are respectively replaced by $\mathbf{i}_{i_j|y_j}$ and $\mathcal{I}_{i_j|y_j}$; in addition, $\mathcal{T}_{j,k} \equiv \mathcal{T}_{i_j,k}$ in (21) and (23), where $i_j \in \mathcal{M}$ for $j = 1, \ldots, m$, with i_j been the index of the monitored node corresponding to the observation j.

4. THE SI MODEL IN NETWORK SYSTEMS

In the simplest mathematical version of an epidemic, two states represent the spreading dynamic: *susceptible* and *infected*. An individual in a susceptible state is someone who does not yet have the disease but can become infected if he or she comes into contact with someone who does. An individual in the infected state has the disease and can transmit it if he or she contacts a susceptible individual (Newman, 2010).

As introduced in Section 2, in epidemic network models the dynamic analysis revolves around the time evolution of the probabilities for nodes to be in specific disease states. Thus, according to *Remark* 1, the states of the SI model are defined by a set of random variables $X_i(t)$, where $X_i(t) = 0$ for a susceptible node and $X_i(t) = 1$ for an infected node (Pastor-Satorras et al., 2015). Let us consider an infectious disease spreading through a population of N individuals. At time $t = 0$, most people are in the susceptible state and only a small fraction, or a single node, in the infected state. Let $\Pr[X_i = 1] \equiv \rho_i(t)$ be the probability that the i-th node becomes infected between times t and $t + dt$, and $\Pr[X_i = 0] \equiv \mathcal{S}_i(t)$ is the probability of the i-th node be susceptible in the same time interval, so $\rho_i(t) + \mathcal{S}_i(t) = 1$. To become infected, an individual must catch the disease from a neighbor j, who is infected. This means that the j-th node must be infected with probability $\rho_j(t) = 1 - \mathcal{S}_j(t)$, and must transmit the disease, during a given time interval, with probability βdt. β is the infection probability, i.e. the probability that a contact of an infected node with a susceptible one yields infection (Schaum and Jaquez, 2016). In addition, the i-th node is also required to be susceptible with probability $\mathcal{S}_i(t)$. So the total probability of i-th node getting infected is $\beta \mathcal{S}_i(t) \sum_j \mathcal{A}_{ij} \rho_j(t)$. Thus, $\rho_i(t)$ obeys the coupled set of N nonlinear differential equations, given by:

$$\frac{d\rho_i(t)}{dt} = \beta \mathcal{S}_i(t) \sum_{j \in \mathcal{N}_i} \mathcal{A}_{ij} \rho_j(t) = \beta \left(1 - \rho_i(t)\right) \sum_{j \in \mathcal{N}_i} \mathcal{A}_{ij} \rho_j(t), \quad i = 1, \ldots, N. \tag{26}$$

The model (26) is known as *first order approximation*. In such approximation, it is implicitly assumed that the product of the averages is equal to the average of the product of the quantities $\mathcal{S}_i(t)$ and $\rho_j(t)$. This is an approximation, as the odds are not independent (Newman, 2010).

Remark 6. Equation (26) is an example of the system described in (2), where $\mathcal{F} = 0$ and $\mathcal{G} = \beta(1 - \rho_i(t))\rho_j(t)$. Note that the state vector has one variable per node ($n = 1$).

4.1 The Discrete SI Global Model

In order to derive the global state equations, (26) can be discretized at t_s sample rate, and rewritten by stacking the state equations for all nodes, to obtain:

$$\rho_k = \rho_{k-1} + t_s \beta \left(\operatorname{diag}\left[1 - \rho_{k-1}\right] \mathcal{A} \rho_{k-1}\right) + \omega_{k-1}, \tag{27}$$

where $\rho_k \in \Re^N$ and $\omega_{k-1} \in \Re^N$. Let us assume a zero-mean white Gaussian process noise satisfying $E\left[\omega_a \omega_b^T\right] = \delta_{a,b} Q_a$, with $Q_a \in \Re^{N \times N}$.

The linearized model of (27) can be obtained by Taylor expansion. Thus, the terms (row= i and column= j) of the Jacobian matrix, associated with state transition matrix, are given by:

$$F_{k-1}^{ij} = \left(\delta_{ij} \left[1 - t_s \beta \left(\sum_{v \in \mathcal{N}_i} \mathcal{A}_{iv} \overline{\rho}_v \right) \right] + t_s \beta \left(1 - \overline{\rho}_i\right) \mathcal{A}_{ij} \right), \tag{28}$$

where $i = 1, \ldots, N$ and $j = 1, \ldots, N$. The variables $(\overline{\rho}_i, \overline{\rho}_v) = (\widehat{\rho}_{i,k-1}, \widehat{\rho}_{v,k-1})$ are the operation points where the Jacobian matrix $F_{k-1} \in \Re^{N \times N}$ is evaluated.

The SI network observation equation can be defined as a direct measure of the m observed states with a source of noise, given by:

$$y_k = C_k \rho_k + \nu_k, \tag{29}$$

where $C_k = \mathbf{1}_{m \times N}$ is a $m \times N$ matrix with 1's on the diagonal and zeros elsewhere and $\nu_k \in \Re^m$. Let us assume a zero-mean white Gaussian measurement noise satisfying $E\left[\upsilon_a \upsilon_b^T\right] = \delta_{a,b} R_a$, with $R_a \in \Re^{m \times m}$, and $E\left[\omega_a \upsilon_b^T\right] = 0$.

4.2 The Discrete SI Distributed Model

To proceed with the purpose of estimating the states of the i-th node by means of a scalable solution, as the problem formulated in Section 2, we can use the transformation relationship $\rho_k = \left(\mathcal{T}_{i,k}\right)^\dagger \rho_{i,k}$, given by (11), in (27), similar we proceeded to derive (12). Also, to make the analysis more simplified we consider that the network topology is constant over time, which means that $\mathcal{T}_{i,k-1} \equiv \mathcal{T}_{i,k} \equiv \mathcal{T}_i$. Thus, the local state transition equation is given by:

$$\rho_{i,k} = \rho_{i,k-1} + t_s \beta \{ \mathcal{T}_i \operatorname{diag}\left[\mathbf{1}_N - \left(\mathcal{T}_i\right)^\dagger \rho_{i,k-1}\right] \mathcal{A} \left(\mathcal{T}_i\right)^\dagger \times \rho_{i,k-1} \} + \omega_{i,k-1}, \; i = 1, \ldots N, \tag{30}$$

where $\rho_{i,k} \in \Re^{n_i}$ and $\omega_{i,k-1} \in \Re^{n_i}$. Furthermore, $\mathbf{1}_N$ indicates a unit vector of length N. Let us assume a zero-mean white Gaussian process noise satisfying $E\left[\omega_{i,a} \left(\omega_{j,b}\right)^T\right] = Q_{i,a} \delta_{a,b} \delta_{i,j}$, with $Q_{i,a} \in \Re^{n_i \times n_i}$.

The terms of the Jacobian matrix, associated with the local model (30), are given by:

$$F_{i,k-1}^{\text{row,col}} = \delta_{\text{row,col}} \left[1 - t_s \beta \left(\sum_{v \in \mathcal{N}_i} \mathcal{A}_{\text{row,v}} \overline{\rho}_v \right) \right] + t_s \beta \left(1 - \overline{\rho}_{\text{row}}\right) \mathcal{A}_{\text{row,col}}, \; i = 1, \ldots N, \tag{31}$$

where row $= 1, \ldots, n_i$ and col $= 1, \ldots, n_i$. The variables $(\overline{\rho}_{\text{row}}, \overline{\rho}_v) = (\widehat{\rho}_{\text{row,k-1}}, \widehat{\rho}_{v,k-1})$ are the values in which the Jacobian matrix $F_{i,k-1} \in \Re^{n_i \times n_i}$ is evaluated. In addition, the local observation equation can be written as:

$$y_{p,k} = \mathcal{C}_{p,k} \rho_{i_p,k} + \nu_{p,k}, i_p \in \mathcal{M}, p = 1, \ldots m, \tag{32}$$

where $\mathcal{C}_{p,k} = \mathbf{1}_{l \times n_i}$ with $l = 1$, means that the measured node i_p has one observation. We assume a zero-mean white Gaussian measurement noise satisfying $E\left[\upsilon_{p,a} \left(\upsilon_{j,b}\right)^T\right] = R_{p,a} \delta_{a,b} \delta_{p,j}$, with $R_{p,a} \in \Re^{l \times l}$.

5. RESULTS AND DISCUSSIONS

A scale-free network for the BA model (Barabási and Albert, 1999) is simulated with $N = 300$ nodes using the SFNG funcion for Matlab (George, 2020). Thus, the SI network was obtained for at $t_s = 0.01$ sampling rate, $\beta = 0.3$ and the initial distribution $\rho_{i,0} = 1$, with i set by a random number generator, indicates that node i is infected in $k = 0$.

We run Monte Carlo simulations for Algorithm (1) to investigate the effect that the number of sensors, chosen

to measure m randomly nodes, has on the estimation of network states. So, the EIF was performed on the global model (27)-(29) for each $m = 1, \dots 300$. For $Q_k = \left(10^{-3}\right)\mathbf{1}_{N\times N}$, $R_k = \left(10^{-1}\right)\mathbf{1}_{m\times m}$ and the initial matrix $Z_0 = \mathbf{1}_{N\times N}$ we obtain the mean value of the estimated state distribution in the superior window on Figure 3, where $\widehat{\rho}_k$ and ρ_k are defined by:

$$\widehat{\rho}_k = \frac{1}{10N}\sum_{M=1}^{10}\sum_{i=1}^{N}\widehat{\rho}_{i,k},\ \rho_k = \frac{1}{10N}\sum_{M=1}^{10}\sum_{i=1}^{N}\rho_{i,k}, \quad (33)$$

where M indicates the number of Monte Carlo simulations. The Root Mean Square Error (RMSE) from $\widehat{\rho}_k$ and ρ_k for a given number of sensors is shown in the inferior window of Figure 3. In this figure, we have shown the results for some m values, because for different values the convergence of the estimated states is too slow, and the RMSE is too large. According to our results, a good performance in state estimation is obtained from monitoring around half of the nodes in the network. Thus, for $m = 160$ and $\rho_{2,0} = 1$ a realization of the EIF algorithm provides the estimated state evolution for each node, as illustrated in Figure 4. As we discussed in Section (3.1), for this centralized algorithm, we must estimate all state network nodes to obtain an estimative about a single node.

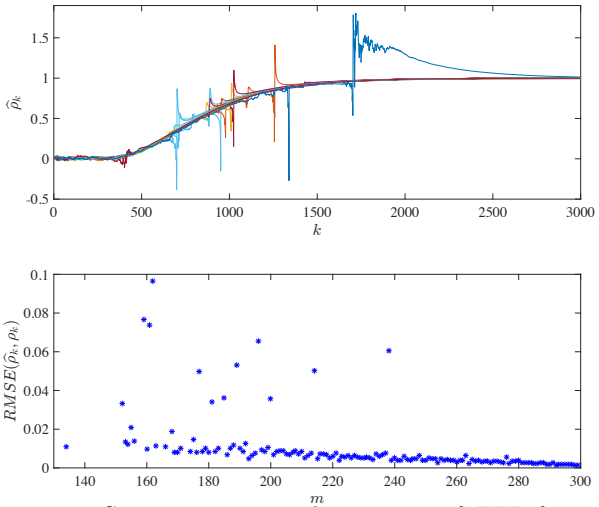

Figure 3. State estimation by means of EIF for a SI scale-free network with $N = 300$, $\beta = 0.3$ and $\Delta t = 0.01$, a mean result of 10 simulated networks with a random initial distribution. Estimated mean state for each number of sensors m (upper window) and corresponding estimated RMSE (lower window).

On the other hand, we can obtain a scalable solution by means of the DDEIF algorithm, in which the state of the i-th node is estimated taking into account, in the forecast step, only the states of its neighboring nodes including itself. For example, let us consider that we wish to estimate the states $\widehat{\rho}_{1,k} \subset \mathcal{N}_1$, $\widehat{\rho}_{6,k} \subset \mathcal{N}_6$, $\widehat{\rho}_{13,k} \subset \mathcal{N}_{13}$ and $\widehat{\rho}_{16,k} \subset \mathcal{N}_{16}$. In our simulations, such nodes are neighbors among themselves and they are not observed by the sensors. Thus, the DDEIF, Algorithm (2), can be applied to the distributed network models described by (30)-(32). We consider that $Q_{i,k} = 10^{-3}\mathbf{1}_{n_i\times n_i}$ and $Z_{i,0} = \mathbf{1}_{n_i\times n_i}$, for $i = 1, 6, 13, 16$. Also, to make a comparison we can do the same measurement conditions applied on EIF (results from Figure 4), i.e. $R_{i,k} = 10^{-1}$ with $i = 1, \dots, m$, for $m = 160$

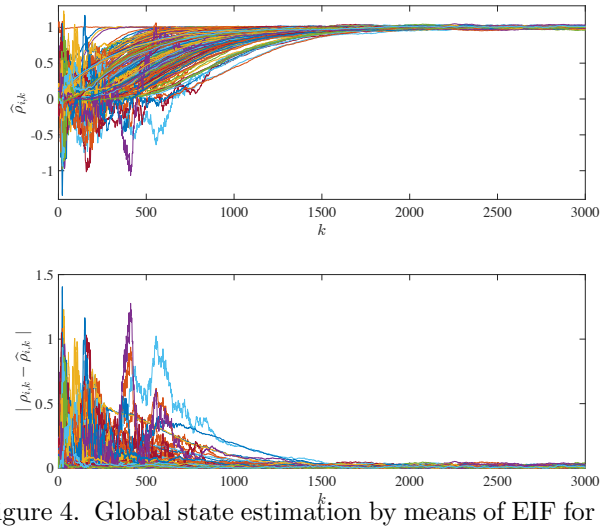

Figure 4. Global state estimation by means of EIF for a SI scale-free network with $N = 300$, $\beta = 0.3$, $\Delta t = 0.01$, $m = 160$ and $\rho_{2,0} = 1$. State distribution of each node (upper window) and estimation error module of each node (lower window).

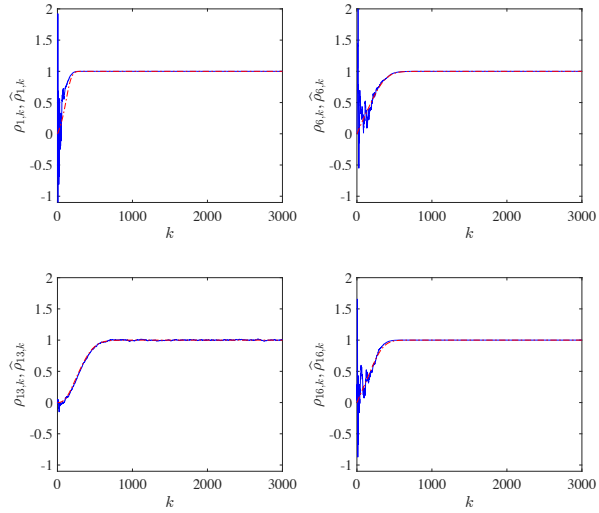

Figure 5. Comparison between the true states (-.) and the local estimated states (-) by means of DDEIF, for a SI scale-free network with $N = 300$, $\beta = 0.3$, $\Delta t = 0.01$, $m = 160$ and $\rho_{2,0} = 1$.

and $\rho_{2,0} = 1$, to obtain the results presented in Figure 5. This results show that the local estimated states (blue curves) are very close to the actual state of the epidemic system (red curves).

In addition, Figure 6 shows a comparison between the state estimation error module for the global and the scalable methods. These results indicate that local estimates (blue curves) are similar to the results obtained by the estimator based on the global network model (red curves). Beyond these qualitative analyses, Table 1 brings comparisons between the RMSE indexes of the EIF and the DDEIF. Differences in the RMSE values are due to numerical errors and nonlinear approximations. Indeed, as previously argued (Mutambara, 1998; Cicala et al., 2020), if the network is connected then the DDEIF is equivalent to the EIF centralized algorithm. Observe that a network is said

Table 1. RMSE index comparison for EIF and DDEIF.

	$\widehat{\rho}_1$	$\widehat{\rho}_6$	$\widehat{\rho}_{13}$	$\widehat{\rho}_{16}$
EIF	0.3838	0.2714	0.5111	0.2524
DDEIF	0.5042	0.5127	0.6581	0.3483

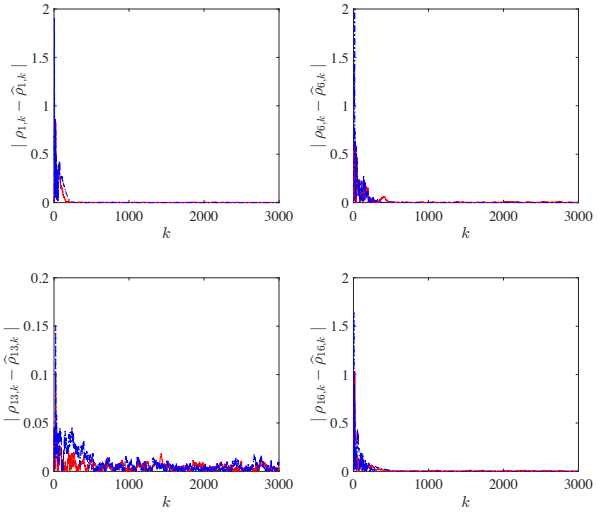

Figure 6. Comparison between estimation error module due to DDEIF (-) and EIF (-) for the nodes $i = 4, 5, 76, 276$.

to be connected if there is a path between every pair of nodes.

6. CONCLUSIONS

Most state estimation techniques in network systems are based on decentralized approaches which take into account the global state transition matrix in the forecast step (Viegas et al., 2018; Deshmukh et al., 2017; Cattivelli and Sayed, 2010). In this research, we seek to estimate the states for complex epidemic networks by means of scalable solutions. For this purpose, the internodal transformation theory is combined with distributed or local models. The distributed models are obtained by scaling the network into local subsystems, and these local subsystems can overlap each other. We can highlight that the local models are defined by the neighborhood of the node in which it is desired to estimate the states.

The results obtained from the simulation example of a SI network suggest that the distributed and decentralized approach is effective in state estimation. As reported in the literature (Mutambara, 1998; Cicala et al., 2020), if the network is connected, then the results of the centralized and decentralized approaches are equivalent. However, despite the order reduction in the global transition matrix, the decentralized and distributed approach described herein requires that all observations be shared between the local processors. In addition, we do not need to spread information among two non-connected nodes. In a scenario wherein the available observations are no shared between the local processors from a cluster (group of connected nodes), the decentralized estimation is not equivalent to the centralized one (optimal estimative).

For future researches, we will investigate the distributed and decentralized state estimation in epidemic complex networks with more than one variable, or state, per node, as the SIR network. Moreover,the establishment of mathematical conditions is of interest in order to reduce the number of required sensors of the connected network.

REFERENCES

Barabási, A.L. (2009). Scale-free networks: a decade and beyond. *science*, 325(5939), 412–413.

Barabási, A.L. and Albert, R. (1999). Emergence of scaling in random networks. *science*, 286(5439), 509–512.

Barrat, A., Barthmy, M., and Vespignani, A. (2008). *Dynamical Processes on Complex Networks*. Cambridge University Press. doi:10.1017/CBO9780511791383.

Berg, T.M. and Durrant-Whyte, H.F. (1991). Model distribution in decentralized multi-sensor data fusion. In *1991 American Control Conference*, 2292–2293. IEEE.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.U. (2006). Complex networks: Structure and dynamics. *Physics reports*, 424(4-5), 175–308.

Cattivelli, F.S. and Sayed, A.H. (2010). Diffusion strategies for distributed kalman filtering and smoothing. *IEEE Transactions on automatic control*, 55(9), 2069–2084.

Cicala, M., D'Amato, E., Notaro, I., and Mattei, M. (2020). Scalable distributed state estimation in utm context. *Sensors*, 20(9), 2682.

Deshmukh, R., Kwon, C., and Hwang, I. (2017). Optimal discrete-time kalman consensus filter. In *2017 American Control Conference (ACC)*, 5801–5806. IEEE.

George, M. (2020). BA scale-free network generation and visualization. MATLAB Central File Exchange. Retrieved March 5, 2021.

Jazwinski, A. (1970). Stochastic processes and filtering theory. new york, academic press, 1970.

Kalman, R.E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82(1), 35.

Mutambara, A.G. (1998). *Decentralized estimation and control for multisensor systems*. CRC press.

Newman, M. (2010). *Networks: An Introduction*. OUP Oxford.

Pastor-Satorras, R., Castellano, C., Van Mieghem, P., and Vespignani, A. (2015). Epidemic processes in complex networks. *Reviews of modern physics*, 87(3), 925.

Pinto, E.R., Nepomuceno, E.G., and Campanharo, A.S. (2020). Impact of network topology on the spread of infectious diseases. *TEMA (São Carlos)*, 21(1), 95–115.

Schaum, A. and Jaquez, R.B. (2016). Estimating the state probability distribution for epidemic spreading in complex networks. *Applied Mathematics and Computation*, 291, 197–206.

Viegas, D., Batista, P., Oliveira, P., and Silvestre, C. (2018). Discrete-time distributed kalman filter design for formations of autonomous vehicles. *Control Engineering Practice*, 75, 55–68.

Wang, W., Chi, K.T., and Wang, S. (2020). Performance comparison of nonlinear kalman filters in epidemic tracking on networks. *IEEE Systems Journal*.

Yang, W., Karspeck, A., and Shaman, J. (2014). Comparison of filtering methods for the modeling and retrospective forecasting of influenza epidemics. *PLoS Comput Biol*, 10(4), e1003583.

Younes, A.B. and Hasan, Z. (2020). Covid-19: Modeling, prediction, and control. *Applied Sciences*, 10(11), 3666.