Abstract: This work considers the problem of obtaining optimal estimates via distributed computation in a large scale system. The electric power system, the transportation system, and generally any computer or network system, are examples of large scale systems: a decentralized estimation of signals based on observations acquired by spatially distributed sensors is the basis for a wide range of important applications. In this work, we focus on the problem of reconstructing the initial state of a linear network in the presence of process and measurement noise. We consider a local model information setup, in which the entire dynamical and measurement model is nowhere available and cannot be reconstructed for the computation. Our estimation procedure relies upon a novel technique to solve a consistent system of linear equations, for which we prove correctness and convergence. In the second part of the paper we consider the problem of detecting anomalies in a large scale network driven by noise. Despite the theoretical advances in this field of research, the currently available procedures to enforce security in large scale systems are computationally inefficient and numerically unreliable. Using our optimal estimation scheme, we describe a distributed procedure with performance guarantees that only requires local knowledge of the system model.

1. INTRODUCTION

Large scale distributed systems, such as the electrical power network and the telecommunication network, are receiving increasing attention from researchers in different fields. The wide spatial distribution and the high dimensionality of network systems forbid the use of centralized solutions to tackle classical estimation, control, and fault detection problems, and they require, instead, the development of new decentralized techniques. One possibility to overcome these issues is to geographically deploy some leaders in the network, each one responsible for a different subpart of the whole system. Local estimation and control schemes can successively be used, together with an information exchange mechanism to recover the performance of a centralized scheme.

We focus here on the problem of reconstructing and computing a function of the initial state starting from local observations. This very basic estimation problem arises often in sensor networks and in multi-agent systems. For instance, many formation and coordination problems require the agents to compute the average of the initial states, or simply a consensus on a variable of interest, e.g., see Ren et al. (2005) and the references therein. The procedures proposed in these works require little knowledge of the network structure, and they can be implemented on agents with very limited computation capabilities. On the other hand, their convergence is usually very slow, and it depends upon the dimension of the network, e.g., see Olshevsky and Tsitsiklis (2009). A different approach is taken in Sundaram and Hadjicostis (2008), where, assuming that the network is observable in a system theoretic sense, a finite time algorithm is derived to both recover the initial state and evaluate a function of it. This method requires the knowledge by each agent of the entire network topology, and it suffers from numerical instabilities as the network size grows. Clearly, a strong tradeoff exists between the performance of a consensus or estimation algorithm and the work load on each single agent. In this paper, we further exploit this tradeoff by proposing a hierarchical structure, in which only a subset of the agents is in charge of the computational burden. By doing so, we maintain the performance and increase the stability of the finite time algorithms, while preserving, for the majority of the agents, the simplicity of the asymptotic procedures.

We allow for the presence of process and measurement noise in the network system. The state estimation problem in a distributed setting and in the presence of noise is receiving increasing attention in the last years, e.g., see Carli et al. (2008) and the references therein. A distributed version of the Kalman filter is often proposed, whose structure depends upon the assumptions on the noise covariance matrix and upon the knowledge each agent needs to have. Besides the fact that, under a general error model, no finite time distributed algorithm is available,
these solutions are better suited to track the evolution of the system, rather than computing a function of the initial state, and they constitute therefore a complementary area of research with respect to this work.

Because of the decentralized nature of network systems, sensor and actuator failures, as well as cyber attacks compromising the availability of resources, constitute an actual threat. As for the state estimation problem, the large dimension of the system under consideration prevents the use of a centralized detection system, hence ruling out classical solutions as presented in Ding (2008). Few techniques have been presented so far that assume only local knowledge of the system model. In Chung et al. (2001) a method based on decentralized estimation and game theoretic fault detection theory is proposed. In Li et al. (2009) an LMI based approach is used to design the structure of local residual generators, in a way that their behavior imitate the output of a centralized detection filter. Although appealing, these solutions have no performance guarantees, and it is a difficult matter to characterize the class of anomalies that can successfully be detected. A different approach is taken in Pasqualetti et al. (2010a), where, by exploiting the presence in the network of weakly interconnected subparts, a bound on the detection performance is characterized.

The main contributions of this work are threefold. First, we describe a novel technique to compute the minimum norm solution as well as the weighted least squares solution to a set of linear equations. The proposed linear system solver belongs to the class of row-projection methods, and it is shown to converge in a finite number of steps. Second, we introduce a hierarchical structure in the network by allowing for the presence of a subset of agents with better computation and communication capabilities. We describe a distributed procedure for the leaders to reconstruct the network initial state and to compute a linear function of it. Our procedure converges in a finite number of steps, which we characterize, and it requires local knowledge of the network topology. In the absence of noise in the system, our method reconstructs exactly the initial state of the system. In the presence of noise, instead, the linear minimum variance estimate of the initial state is computed. Third and finally, we use our estimation procedure to design a distributed and finite time detection scheme, in which each leader operates under local model information. A condition to avoid incorrect detection is given, and, consequently, the set of detectable anomalies is characterized. The proposed detection method is, to the best of our knowledge, the first decentralized detection method with performance guarantees that works on any network structure and in the presence of noise.

The rest of the paper is organized as follows. Section 2 contains the notation. Our linear system solver is in Section 3. Sections 4 and 5 contain our main results. Our numerical studies and conclusion are in Section 6 and 7.

2. PROBLEM SETUP

We introduce in this section the notation and the problems under consideration. Let $G$ denote a directed graph with vertex set $V = \{1, \ldots, n\}$ and edge set $E \subseteq V \times V$. The in-neighbor set of a node $i \in V$, i.e., all the nodes $j \in V$ such that the pair $(j, i) \in E$, is denoted with $N_i$. We let each vertex $j \in V$ denote an autonomous agent, and we associate a real number $x_j$ with each agent $j$. Let the vector $x$ contain the values $x_j$. A linear iteration over $G$ is an update rule for the state vector $x$ and it is described by the linear discrete time system

$$x(t + 1) = Ax(t).$$

We refer to the graph $G$ as to the communication graph associated with the network matrix $A$, in which the $(i, j)$-th entry is nonzero if and only if $(j, i) \in E$. An output matrix $C_j$ describes the information about the state of the network that is directly available to $j$. In particular, $y_j(t) = C_j x(t)$, $C_j = [c_{n_1}, \ldots, c_{n_p}]^T$, and $[n_1, \ldots, n_p] = N_j$. In this work, we consider the noisy network

$$x(t + 1) = Ax(t) + v(t),$$

$$y_j(t) = C_j x(t) + w_j(t),$$

where $v(t) \in \mathbb{R}^n$ and $w_j(t) \in \mathbb{R}^p$ for all $j \in \{1, \ldots, n\}$. We also make the standard assumptions that

$$E[v(t)v(t)] = 0, \quad E[w_i(t)w_j(t)] = 0,$$

and that, being $\delta$ the Kronecker delta function and $I$ the identity matrix of appropriate dimension,

$$E[v(t)] = 0, \quad E[(v(t)v(\tau)^T) = \sigma_v^2 \delta(t - \tau) I],$$

$$E[w_j(t)] = 0, \quad E[(w_j(t)w_j(\tau)^T) = \sigma_w^2 \delta(t - \tau) I],$$

for all $t, \tau \in \mathbb{R}$ and for all $i, j \in \{1, \ldots, n\}$ with $i \neq j$. We introduce a hierarchy in the network by assuming that some agents, denoted as leaders, have better computation and communication capabilities. Let $V^{(l)} = \{\ell_1, \ldots, \ell_m\} \subseteq V$ be the leader set, and let $G^{(l)} = (V^{(l)}, E^{(l)})$ be the leader graph, where $E^{(l)} \subseteq V^{(l)} \times V^{(l)}$ describes the feasible communications among the leaders. Let $N^{(l)}_i$ denote the set of the neighbors of the leader $\ell_i$ in $G^{(l)}$. The information of the state $x(t)$ directly available to the leader $\ell_i$ is given by $y_{\ell_i}(t) = C_{\ell_i} x(t)$, where $C_{\ell_i}$ is defined according to the neighbor set $N_{\ell_i}$ in $G$. We assume that the leader graph $G^{(l)}$ is strongly connected.

The problem of obtaining the linear minimum-variance estimate of the initial state of the network (1) is considered in this work. Observe that, if the agent $j$ has complete knowledge of the network matrix $A$, and if the pair $(A, C_j)$ is observable, see Ding (2008), then an optimal estimation of the initial state can be obtained by collecting a sufficient number of output measurements, and by solving a standard weighted least squares problem as described in Luenberger (1969). We consider in this paper the case in which every agent (resp. leader) has only local knowledge of the network model. In other words, we assume that the matrix $A$ is nowhere available and, furthermore, cannot be reconstructed, so that the centralized solution described above is inapplicable.

Problem 1. (Distributed minimum-variance estimate) Design a distributed algorithm for the leaders to obtain the linear minimum-variance estimate of $x(0)$.

In the second part of the paper we use our distributed estimation algorithm to address the problem of detecting failures and anomalies in the network. We model the presence of an anomaly as an unknown and unmeasurable exogenous input in the system (1). This model has been shown to be general, since it considers both sensor and
Algorithm 1  Sequential system solver

Input: $A_1, \ldots, A_m, y_1, \ldots, y_m$

Require: $y \in \text{Im}(A)$;

$x_0 := 0$, $K_0 := I_d$;

for $i = 1 : m$ do

$x_i := x_{i-1} + K_{i-1}(A_iK_{i-1})^T(y_i - A_i x_{i-1})$;

$K_i := \text{Basis}(K_{i-1} \text{Ker}(A_iK_{i-1}))$;

end for

return $x_m$;

---

unknown and unmeasurable $n$-dimensional random vector, with $\mathbb{E}[w] = 0$ and $\mathbb{E}[ww^T] = Q = Q^T > 0$. Let

$$y = Ax + w,$$  \hspace{1cm} (5)

where $\text{Ker}(A) = 0$, and recall from Luenberger (1969) that the linear minimum variance estimate of $x$ equals

$$x_{\text{wls}} = (A^TQ^{-1}A)^{-1}A^TQ^{-1}y.$$  

We aim at computing the vector $x_{\text{wls}}$ by means of Algorithm 1. Because the matrix $Q$ is symmetric and positive definite, there exists a full row rank matrix $B$ such that $Q = BB^T$. Notice that equation (5) can be rewritten as

$$y = [A \in B] x \hat{\bar{w}},$$  \hspace{1cm} (6)

where $\epsilon > 0$, $\mathbb{E}[\hat{w}] = 0$ and $\mathbb{E}[\hat{w}\hat{w}^T] = \epsilon^{-2}I$. Because $B$ has full row rank, the system (6) is underdetermined, i.e., $y \in \text{Im}([A \in B])$ and $\text{Ker}([A \in B]) \neq 0$. Let

$$\hat{x}(\epsilon) = [A \in B]^T y.$$  

Notice that the system (6) can be partitioned as in (4), and that Algorithm 1 can be used to compute the vectors $\hat{x}(\epsilon)$ and $\hat{w}$. We next show how the vector $\hat{x}(\epsilon)$ is related to the minimum variance estimate $x_{\text{wls}}$.

Theorem 2. (Asymptotic convergence) Consider the system of linear equations (5). Let $\mathbb{E}[w] = 0$ and $\mathbb{E}[ww^T] = Q = Q^T > 0$, for a full row rank matrix $B$. Let $C = \epsilon(I - AA^T)B$,

$$E = I - C^\dagger C,$$

$D = \epsilon E[I + \epsilon^2 EB^T(AA^T)^\dagger BE]^{-1}B^T(AA^T)^\dagger(I - \epsilon BC^\dagger).$  

Then

$$\lim_{\epsilon \to 0^+} A^\dagger - \epsilon A^\dagger B(C^\dagger + D) = C^\dagger + D,$$

and

$$\lim_{\epsilon \to 0^+} A^\dagger - \epsilon A^\dagger B(C^\dagger + D) = (A^TQ^{-1}A)^{-1}A^TQ^{-1}.$$  

Throughout the paper, let $\hat{x}(\epsilon)$ be the estimate of $x$ returned by Algorithm 1 applied to (6), and let

$$\hat{x} = \lim_{\epsilon \to 0^+} \hat{x}(\epsilon).$$

Theorem 2 guarantees that the minimum variance estimate $x_{\text{wls}}$ coincides with the limit of $\hat{x}(\epsilon)$ as $\epsilon \to 0^+$. When the parameter $\epsilon$ is fixed, the estimate $\hat{x}(\epsilon)$ differs from the minimum variance estimate $\hat{x}_{\text{wls}}$. We next characterize the approximation error $x_{\text{wls}} - \hat{x}(\epsilon)$.

Corollary 3. (Approximation error) Consider the system (5), and let $\mathbb{E}[ww^T] = BB^T > 0$. Then

$$x_{\text{wls}} - \hat{x}(\epsilon) = \epsilon A^\dagger BDy,$$

where $B$ and $D$ are as in Theorem 2.

---

1. Choose for instance $B = W A^T/2$, where $W$ is a basis of eigenvectors and $A$ is the corresponding diagonal matrix of the eigenvalues.
To conclude this section, we characterize the residual $y - Ax$. This quantity plays an important role in Section 5 for the synthesis of a distributed detection algorithm.

**Corollary 4. (Residual)** Consider the system (5), and let $\mathbb{E}[ww^{T}] = Q = Q^{T} > 0$. Then\(^2\)

$$\lim_{t \to +\infty} ||y - Ax|| \leq \|(I - AW)|| ||w||,$$

where $W = (ATQ^{-1}A)^{-1}ATQ^{-1}$.

### 4. OPTIMAL FUNCTION EVALUATION

The problem of obtaining the minimum variance estimate of the initial state of the noisy network (1) is considered in this section. Precisely, we exploit the hierarchic levels in the network, and we describe a distributed procedure for the leaders to recover the desired estimation under the assumption that no leader has knowledge of the entire network structure. We first introduce some notation. Let $V$ denote the set of agents, and let $\{\ell_{1}, \ldots, \ell_{m}\} \subseteq V$ denote the subset of leaders. Let $C_{\ell_{i}}$ be the output matrix associated with the leader $\ell_{i}$, and let $w_{\ell_{i}}$ be the corresponding measurement noise. The composite information available to the leaders can be conveniently described by the output matrix $C^{(i)} = [C^{(i)}_{1}, \ldots, C^{(i)}_{2}]$, so that the evolution of the network with the leaders observations becomes

$$x(t + 1) = Ax(t) + v(t),$$
$$y(t) = C^{(i)}x(t) + w(t),$$

where $w(t)^{T} = [w_{\ell_{1}}(t)^{T} \ldots w_{\ell_{m}}(t)^{T}]^{T}$, and $v(t), w(t)$ satisfy the assumptions (2) and (3) at all instants $t$. Let $I_{n}$ denote the identity matrix of dimension $n \times n$, and let $0_{m,n}$ denote the zero matrix of dimension $m \times n$. Let $|V| = n$, and let $|N_{\ell_{i}}| = p_{\ell_{i}}$. For $s \in \mathbb{Z}_{>0}$, let

$$O_{i}^{s} = \begin{bmatrix}
C_{\ell_{i}} & C_{\ell_{i}}A & \ldots & C_{\ell_{i}}A^{s-1} \\
0 & y_{\ell_{i}}(0) & \ldots & y_{\ell_{i}}(s-1)
\end{bmatrix}, 
Y_{i}^{s} = \begin{bmatrix}
y_{\ell_{i}}(0) \\
y_{\ell_{i}}(1) \\
y_{\ell_{i}}(2) \\
\vdots \\
y_{\ell_{i}}(s-1)
\end{bmatrix},$$

and

$$F_{i}^{s} = \begin{bmatrix}
D_{\ell_{i}} & 0 & \ldots & 0 \\
C_{\ell_{i}}B & D_{\ell_{i}} & \ldots & 0 \\
C_{\ell_{i}}AB & C_{\ell_{i}}B & \ldots & 0 \\
\vdots & \vdots & \ddots & \vdots \\
C_{\ell_{i}}A^{-2}B & C_{\ell_{i}}A^{-3}B & \ldots & C_{\ell_{i}}B & D_{\ell_{i}}
\end{bmatrix},$$

where $B = [I_{n} \ 0_{n,p_{\ell_{i}}} \ \ldots \ 0_{n,p_{\ell_{m}}}]$, $D_{\ell_{i}} = [0_{p_{\ell_{i}},n} \ \ldots \ 0_{p_{\ell_{i}},p_{\ell_{m-1}}} \ I_{p_{\ell_{i}}} \ 0_{p_{\ell_{i}},p_{\ell_{i+1}}} \ \ldots \ 0_{p_{\ell_{i}},p_{\ell_{m}}}]$.

Finally, let

$$O^{s} = \begin{bmatrix}
O_{1}^{s} \\
O_{2}^{s} \\
\vdots \\
O_{m}^{s}
\end{bmatrix}, 
Y^{s} = \begin{bmatrix}
Y_{1}^{s} \\
Y_{2}^{s} \\
\vdots \\
Y_{m}^{s}
\end{bmatrix}, 
F^{s} = \begin{bmatrix}
F_{1}^{s} \\
F_{2}^{s} \\
\vdots \\
F_{m}^{s}
\end{bmatrix}.$$\(^2\)

\(^2\) Given a vector $w$ and a matrix $A$, we denote by $||w||$ any vector norm, and by $||A||$ the corresponding induced matrix norm.

**Algorithm 2** Sequential state estimator (Leader $\ell_{i}$)

**Input:** $O_{i}^{s}, F_{i}^{s}A^{1/2}, y_{t}, \epsilon$;

- if $i = 1$ then
  - $\hat{x} := 0$ and $K := I$;
  - receive $\hat{x}$ and $K$ from leader $\ell_{i-1}$;
- else
  - $\hat{x} := \hat{x} + K([O_{i}^{s} \ \epsilon F_{i}^{s}A^{1/2}][y_{t} - [O_{i}^{s} \ \epsilon F_{i}^{s}A^{1/2}]] \hat{x})$;
  - $K := $ Basis($K$ Ker($[O_{i}^{s} \ \epsilon F_{i}^{s}A^{1/2}][y_{t} - [O_{i}^{s} \ \epsilon F_{i}^{s}A^{1/2}]])$);
  - if $i = m$ then
    - transmit $\hat{x}$ to every leader;
  - return $\hat{x}$;
  - else
    - transmit $\hat{x}$ and $K$ to leader $\ell_{i+1}$;
- end if

and, for $t \geq 0$, $U(t) = \begin{bmatrix}
v(t) \\
w_{t_{1}}(t) \\
\vdots \\
w_{t_{m}}(t)
\end{bmatrix}, 
U^{*} = \begin{bmatrix}
U(0) \\
U(1) \\
\vdots \\
U(s-1)
\end{bmatrix}$.

Observe that

$$Y^{s} = O^{s}x(0) + F^{s}U^{*}.$$  \(^8\)

Assume that Ker($O^{*}$) = 0, and notice that $F^{s}$ has full row rank. The minimum variance estimate of $x(0)$ according to equation (8) is given by

$$\hat{x}(0) = ((O^{s})^{T}Q^{-1}O^{s})^{-1}(O^{s}Q^{-1}Y^{s}),$$

where $Q = \mathbb{E}[(F^{s}U^{*})(F^{s}U^{*})^{T}] = (F^{s}A^{1/2})(A^{1/2}(F^{s})^{T})$, and $\Lambda = \mathbb{E}[U^{*}(U^{*})^{T}]$. Observe that $\Lambda$ is diagonal and positive definite. We now show how Algorithm 1 can be used to compute the minimum variance estimate of $x(0)$. Let $\epsilon \neq 0$, and consider the system of linear equations

$$Y^{s} = \begin{bmatrix}
O^{s} & \epsilon F^{s}A^{1/2}
\end{bmatrix}^{T} \begin{bmatrix}
x(0) \\
U^{*}
\end{bmatrix},$$

where $U^{*}$ is such that $U^{*} = \epsilon A^{1/2}U^{*}, \mathbb{E}[U^{*}(U^{*})^{T}] = \epsilon^{-2}I$.

Assume that an ordering of the leaders has been defined in a way that $\ell_{i} \in N_{i+1}$ for all $i \in \{1, \ldots, m-1\}$. A procedure for the leaders to estimate $x(0)$ is in Algorithm 2.

**Theorem 5. (Minimum variance estimate)** Consider the noisy network (7). Let $\epsilon > 0$, and assume that each leader knows the matrices $O^{s}$ and $F^{s}A^{1/2}$. Then Algorithm 2 returns an estimate $\hat{x}(\epsilon)$ such that $\|x_{\text{mns}} - \hat{x}(\epsilon)\| = \epsilon c$, where the constant $c$ depends on $O^{s}, F^{s}A^{1/2}$, and $Y^{s}$.

It should be noticed that if the parameter $\epsilon$ converges to zero, then the estimate obtained with Algorithm 2 coincides with the minimum variance estimate of the network initial state. Finally, it can be shown that a leader based procedure, such as Algorithm 2, is numerically more reliable with respect to classical centralized solutions.

**Remark 1.** (Function evaluation). The proposed method can be effectively employed to achieve the minimum variance estimate of a linear function of the network initial state. Indeed, suppose that the function $f(x(0))$ needs to be computed by all (or a subset of) the leaders, and recall...
from Luenberger (1969) that the minimum variance estimate of \( f(x(0)) \) is \( \hat{x} \), where \( \hat{x} \) is the minimum variance estimate of \( x(0) \). Hence, the function evaluation problem is a direct extension of the state estimation problem.

**Remark 2.** (Network knowledge). The computation of the matrices \( O_i^s \) and \( F_i^s \), \( i \in V(\ell) \), does not require the knowledge of the entire network model. Let a path be a sequence of vertices, such that any two consecutive vertices in the sequence are connected through an edge. Let the length of a path equal the number of its edges. Let \( A \) be the network matrix, and observe that the \( (i,j) \)-th entry of \( A^K \), with \( k \in \mathbb{Z} \), is nonzero if and only if there exists a path of length \( k \) connecting the agent \( j \) to \( i \). Let \( N_i^s \subseteq V \) denote the set of the neighbors connected to \( \ell_i \) through a path of length at most \( s \). It can be shown that the matrix \( O_i^s \) and \( F_i^s \), \( i \in V(\ell) \), can be computed by only knowing the sub-matrix of \( A \) and \( \Lambda^{1/2} \) with rows and columns in \( N_i^s \).

### 5. Anomaly Detection

Because of the absence of a centralized and omniscient entity that monitors the behavior of the network, the system (1) is prone to malfunctions and external attacks. It can be shown that the presence of sensor and actuator failures, as well as of external attacks, can be modeled by adding an unknown and unmeasurable input to the network, e.g., see Pasqualetti et al. (2010a). We say that the network (1) undergoes an anomaly if it is affected by an exogenous input that steers the network along state trajectories not belonging to (1). Precisely, the network model in the presence of an anomaly \( e \) becomes

\[
x(t + 1) = Ax(t) + Ee(t) + v(t),
\]

where the matrix \( E \) is referred to as the anomaly or failure signature. We show in this section how our estimation procedure can be used to allow the leaders to detect an anomaly in the network by only using local model information. It should be noticed that the presence of the noise \( v \) and \( w \) limit the class of detectable anomalies.

Because the failure signature is a priory unknown, the leaders needs to detect the anomaly by means of only the past observations. Let \( Y_i^s \) denote the vector containing the measurements taken by the leader \( \ell_i \) from time \( t \) up to time \( t+s-1 \). Analogously, let \( U_i^s \) and \( U_{i,e}^s \) contain, respectively, the noise value and the exogenous input from time \( t \) up to time \( t+s-1 \). Then, with the notation introduced in the previous section, we have

\[
\begin{bmatrix}
Y_{1,t}^s \\
Y_{2,t}^s \\
\vdots \\
Y_{m,t}^s
\end{bmatrix}
= \begin{bmatrix}
O_1^s & \cdots & 0 \\
O_2^s & \ddots & \vdots \\
\vdots & \ddots & \ddots \\
O_m^s & \cdots & O_m^s
\end{bmatrix} \begin{bmatrix}
x(t) \\
x(t^+1) \\
\vdots \\
x(t^+m)
\end{bmatrix}
+ \begin{bmatrix}
F_1^s(E) \\
F_2^s(E) \\
\vdots \\
F_m^s(E)
\end{bmatrix} \begin{bmatrix}
U_i^s \\
U_{i,e}^s \\
\vdots \\
U_{i,u}^s
\end{bmatrix},
\]

where, for \( i \in \{1, \ldots, m\} \),

\[
F_i^s(E) = 
\begin{bmatrix}
0 & 0 & \cdots & 0 \\
C_t. E & 0 & \cdots & 0 \\
C_t. E A & C_t. E & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots \\
C_t. E A^{m-2} E & C_t. E A^{m-1} E & \cdots & C_t. E
\end{bmatrix}
\]

### 6. SIMULATIONS

In this section two numerical studies are considered to demonstrate the effectiveness of our numerical procedures. First, following Theorem 5, we show that, as the parameter \( \varepsilon \) decreases, the estimation of the network initial state converges to the minimum variance estimate of the initial state. Second, we demonstrate the performance of our Anomaly detection procedure. Let \( G \) denote the graph of Fig. 1, and construct the matrix \( A = [a_{ij}] \) as follows:

(i) if the agents \( i \) and \( j \) are connected, then set the value of \( a_{ij} = a_{ji} \) according to an uniform distribution \(^3\) in the interval \((0, 1)\), and set \( a_{ij} = a_{ji} = 0 \) otherwise;

(ii) normalize the rows of \( A \) such that \( \|A\|_\infty = 1 \).

Let \( V(\ell) \) denote the leader set, and let \( C(\ell) \) be the composite observation matrix. The evolution of the network is described by (7), where the random vectors \( v(t) \) and \( w(t) \) satisfy the assumptions (2) and (3) at all instants

\(^3\) If the entries of \( A \) are chosen uniformly and independently in the interval \((0, 1)\), then the pair \((A, C_j)\) is observable with probability one for every output matrix \( C_j \).
The problems of reconstructing the initial state and of detecting anomalies in a linear network driven by process and measurement noise have been considered. Finite time algorithms that only assume local knowledge of the network model have been proposed and characterized.

REFERENCES