Average state estimation in presence of outliers

Ujjwal Pratap, Carlos Canudas de Wit, Federica Garin

Average state estimation in presence of outliers

Ujjwal Pratap*, Carlos Canudas-de-Wit† and Federica Garin*

***Abstract*— This paper addresses the problem of estimation of the average state of an LTI network system where measurements of only some of the states are available. We consider the case where there is an outlier among the unmeasured states, which is so different from the remaining states that it affects the average value significantly. The goal of this paper is to detect the outlier and to estimate the average of the unmeasured states excluding the outlier.**

I. Introduction

In recent years, there has been a lot of interest in the problem of estimation in network control systems due to its vast area of application such as power networks, communication networks, and traffic networks to name a few. In particular, in some networks, measurement of every state is not possible or the measurements of some sensors are unavailable. In such cases, measurements from dedicated sensors placed at some position in the network are used to estimate the remaining fraction of states or some function of those states. For instance, [1] investigates the possibility of estimating the average of the unmeasured nodes in the network. In [2] and [3], authors propose designs of average state observers. This approach is very beneficial in terms of reducing complexity. However, there might be an outlier in the unmeasured part which could be an error or an anomaly under some special circumstances. So, in presence of these outliers, the estimated average value obtained from these estimation techniques may not be the value one is interested in. Instead, one might look for an average value which excludes the outlier. For example, in a power distribution network, one might be interested in estimating the average household consumption of an area using measurements from sensors deployed at some dedicated position only. But in case, there is an "industry" or a "research facility" in the area, which is the major consumer, and its consumption is not measured directly, then one might not be able to get the desired value through average estimation since the "power consumption of the industry" can affect the estimated average household consumption significantly. Therefore, it is natural to ask some questions: what if there is an outlier in the unmeasured section of the network? How to apply an estimation technique such that it filters the outlier and identifies it simultaneously? How to estimate the desired function without the effect of such outliers?

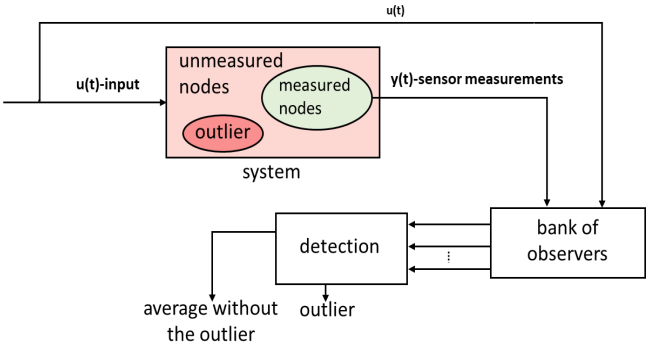

Fig. 1. Scenario considered in the paper and an outline of the proposed method.

Outlier analysis and detection is a very well studied problem in statistics. Some of the classical works are [4], [5]. The former defines the outlier as an observation which deviates so much from the other observations as to arouse suspicions that it was generated by a different mechanism and the latter proposes different regression based detection methods. Since then, research in this field has flourished and some of the recent works include [6], [7], [8], [9] which propose different detection methods such as detection by active learning, k-nearest neighbour method, depth-based and distance-based methods. However, the techniques used for outlier detection in this literature cannot be applied in our case because they consider only static data and also all the measurements are available to them. On the other hand, we have a dynamical system changing over time and have limited measurements.

In dynamical systems also, there have been works such as [10], [11], [12] which focus on outlier detection in sensor networks. In these cases, the outlier is assumed to be present in the measurements received from the sensors only. In particular, [12] proposes a robust Kalman filter to detect and exclude the outlier from the sensor measurements which is somehow similar to our goal. However, the main difference is that in our case the outlier is in the system itself, and is an unmeasured state which is so different from the other states that it affects their average significantly. Our goal is to estimate the average of the unmeasured states and detect the outlier simultaneously.

The scenario considered in this paper considers a continuous LTI system with dedicated sensors at some nodes with an outlier state corresponding to an unmeasured node.

∗ Ujjwal Pratap and Federica Garin are with Univ. Grenoble Alpes, CNRS, Inria, Grenoble INP, GIPSA-lab, Grenoble, France. Email: ujjwal.pratap@gipsa-lab.fr, federica.garin@inria.fr.

† C. Canudas-de-Wit is with CNRS, GIPSA-Lab 38000 Grenoble, France. Email: carlos.canudas-de-wit@gipsa-lab.fr.

This work is supported by European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC-Adg no. 694209, Scale-FreeBack, website: http://scale-freeback.eu/).

We propose a method to detect the outlier and estimate the average excluding it simultaneously. As illustrated in Figure 1, our approach is to run a bank of observers and then compare the estimates so obtained in order to detect the outlier. For this, at first, we provide a necessary and sufficient condition under which a bank of scalar and a tunable observer can be designed so as to estimate the average of the unmeasured states excluding an element at every possible position. Then, we define a distance-based dissimilarity criterion to differentiate between the average estimates so obtained. Finally, using a simple optimization we detect the outlier and obtain the average estimate excluding it.

II. Problem formulation

Consider a network represented by a weighted directed graph $(G) = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{1, 2, 3 \dots n\}$ denotes the set of the nodes and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ denotes the set of edges. We follow the convention that the edge $(i, j) \in \mathcal{E}$ is represented as $i \leftarrow j$, since this edge will correspond to the influence of state x_j on the dynamics of state x_i. Let $A = [a_{ij}]$ be the associated weighted adjacency matrix, where a_{ij} is the weight of the edge $(i, j) \in \mathcal{E}$.

The dynamics of the network is described by

$$\begin{cases} \dot{x} = & Ax + Bu \\ y = & Cx. \end{cases} \tag{1}$$

where $x \in \mathbb{R}^n, u \in \mathbb{R}^m$ and $y \in \mathbb{R}^{n_1}$ are the state vector, the input vector and the output vector respectively. The system matrices are $A \in \mathbb{R}^{n \times n}, B \in \mathbb{R}^{n \times m}$ and $C \in \mathbb{R}^{n_1 \times n}$. We assume to have bounded state trajectories.

We assume to have dedicated sensor measurements at n_1 nodes, i.e., the output y contains the values of n_1 states. Without loss of generality, we order the states starting with the measured ones, so that we have the state partition $x(t) = [x_1^T(t), x_2^T(t)]^T$, where the vector $y = x_1(t) \in \mathbb{R}^{n_1}$ contains the measured states and $x_2(t) \in \mathbb{R}^{n_2}$ contains the unmeasured states. Denoting by $I_s \in \mathbb{R}^{s \times s}$ the identity matrix of size s, the block structure of the matrices corresponding to the above-mentioned partition is

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \text{ and } C = \begin{bmatrix} I_{n_1} & 0 \end{bmatrix}. \tag{2}$$

With this partition, the system can be rewritten as

$$\begin{cases} \dot{x}_1 & = A_{11}x_1 + A_{12}x_2 + B_1 u \\ \dot{x}_2 & = A_{21}x_1 + A_{22}x_2 + B_2 u \\ y & = x_1. \end{cases} \tag{3}$$

Having given the system description, we define the outlier in consideration as follows.

Definition 1. *A state is called an outlier if its value differs from all the other states by such a large value that their average value changes significantly.*

In the next subsection we present an example to motivate the reader towards the problem. This example shows how an outlier affects the average value. The same example has been used in subsequent sections to illustrate the proposed outlier detection method.

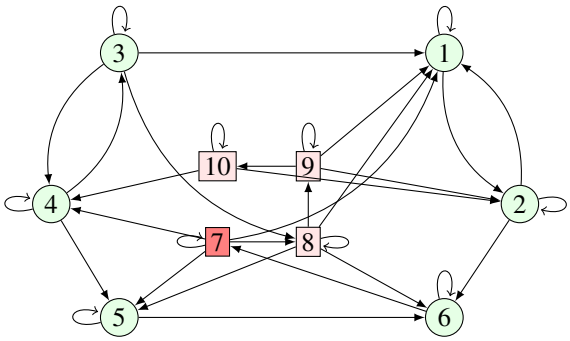

Fig. 2. A network system with an outlier at node 7. The nodes in green circles are the measured nodes and the nodes in red squares are the unmeasured nodes with the outlier in a darker red shade.

A. Motivating Example

Example 1. *Consider the network depicted in Figure 2. The dynamics of the network is described as in* (1)*. The input is given by* $u(t) = 0.2{*}\sin(0.1t)$*, and the corresponding system matrices* A, B, C *according to the partition in* (2) *are*

$$A_{11} = \begin{bmatrix} -3.253 & 0.9843 & 0.8439 & 0 & 0 & 0 \\ 0.6147 & -3.332 & 0 & 0 & 0 & 0 \\ 0 & 0 & -3.525 & 0.4813 & 0 & 0 \\ 0 & 0 & 0.2488 & -3.05 & 0 & 0 \\ 0 & 0 & 0 & 0.1325 & -1.689 & 0 \\ 0 & 0.5427 & 0 & 0 & 0.8452 & -2.178 \end{bmatrix}$$

$$A_{12} = \begin{bmatrix} 0.345 & 0.345 & 0.345 & 0 \\ 0 & 0 & 0.7607 & 0.7607 \\ 0 & 0.22 & 0 & 0 \\ 0.2288 & 0 & 0 & 0.2288 \\ 0.1647 & 0.1647 & 0 & 0 \\ 0 & 0.3452 & 0 & 0 \end{bmatrix},$$

$$A_{21} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0.020844 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$A_{22} = \begin{bmatrix} -0.02784 & 0 & 0 & 0 \\ 0.0234 & -1.4 & 0 & 0 \\ 0 & 0.639 & -1.159 & 0 \\ 0 & 0 & 0.3983 & -1.999 \end{bmatrix},$$

$B = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}^T$ *and* $C = \begin{bmatrix} I_6 & 0 \end{bmatrix}$.

Here, in this example, we have a network of $n = 10$ *nodes depicted in Figure 2. The sensor measurements are obtained from the nodes* $\{1, \dots, 6\}$ *denoted by the circle nodes and the nodes* $\{7, \dots, 10\}$ *are the unmeasured nodes denoted by*

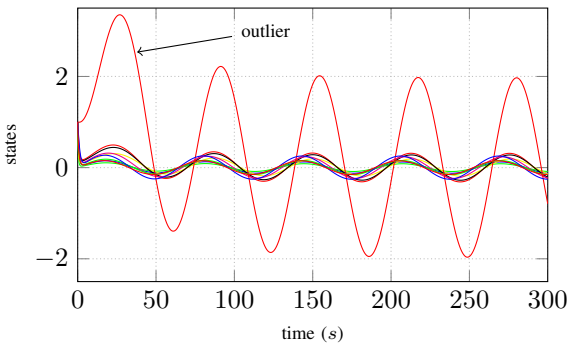

Fig. 3. State trajectories of the network in Figure 1 in response to $u(t) = 0.2{*}\sin(0.1t)$. Here, it can be seen that the outlier state (in red) is behaving very differently than the other states.

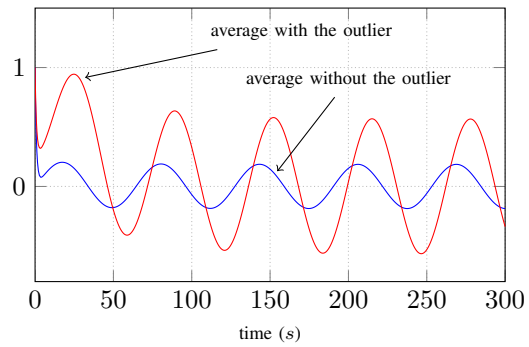

Fig. 4. Trajectories of the average of the unmeasured states with and without the outlier, demonstrating the effect of an outlier on the average.

the square nodes. Here, $n_1 = 6$ and $n_2 = 4$. We have an outlier at node 7 pointed as the shaded node. It can be seen in Figure 4, where the averages of the unmeasured nodes with and without the outlier have been computed, that there is a significant difference in the average value because of the outlier. Hence, we must find ways to estimate the average in such a way that it excludes the outlier.

B. Problem statement

Consider a network with the dynamics as (1) with the assumption that there exists an outlier in the unmeasured part of the system. Using only the measurements $y(t) = x_1(t)$, how is it possible to construct the average of the unmeasured states $x_2(t)$ without the outlier if its position may or may not be known? In addition, is it possible to construct a scalar observer to estimate such an average with arbitrary rate of convergence? Moreover, if the position of the outlier is unknown, how is it possible to detect it and compute the average excluding it simultaneously?

III. Outlier at a known position

In this section, we consider the case when the position of the existing outlier is known. We define j_o as the true position of the outlier in the set of unmeasured nodes. In the network, the position j_o refers to the node $n_1 + j_o$. At first, we recall the necessary and sufficient condition to design a scalar and tunable observer to estimate the average of all but one element j in the set of unmeasured nodes. Then, we provide an explicit construction of the observer. Such observer is then used in this section with $j = j_o$ (j_o being known), and will also be reused in the next section for all j, when the position j_o is to be found. In the end we illustrate the estimation using Example 1.

Let $c_j \in \mathbb{R}^{n_2}$, be a vector of all ones but zero at the $j-$th position. Let us define $q_j = \frac{1}{n_2-1} c_j$, so that the average state of the unmeasured nodes excluding the element at the $j-$th position be denoted as

$$x^{\text{av}}_{2,j}(t) = q_j^T x_2, \tag{4}$$

where x_2 is the vector of unmeasured states. Again, the node in $j-$th position among unmeasured nodes is the node $(n_1 + j)$.

To reconstruct the average $x^{\text{av}}_{2,j}(t)$, we consider a scalar observer, namely a system of the form

$$\begin{cases} \dot{w}_j(t) &= -\alpha w_j(t) + h_j^T y(t) + g_j^T u(t) \\ \hat{x}^{\text{av}}_{2,j}(t) &= w_j(t) + \ell_j^T y(t). \end{cases} \tag{5}$$

where $w_j(t) \in \mathbb{R}$ is the state of the observer, while $\alpha \in \mathbb{R}$, $\ell_j, h_j \in \mathbb{R}^{n_1}$ and $g_j \in \mathbb{R}^m$ will be suitably designed. Let $e_j(t) = x^{\text{av}}_{2,j}(t) - \hat{x}^{\text{av}}_{2,j}(t)$ be the estimation error. We say that (5) is an observer if the parameters α, h_j, g_j and ℓ_j can be chosen such that $e_j(t) \to 0$ as $t \to \infty$. Moreover, we are interested in designing an observer with an arbitrary rate namely that the error $e_j(t) \to 0$ as $t \to \infty$ with a desired rate of convergence. The condition under which this is possible is studied in [3].

Theorem 1. *[3, Theorem IV.1] A scalar and tunable observer of the form* (5) *to estimate $x^{\text{av}}_{2,j}(t)$ for a given j exists if and only if*

$$\operatorname{rank} \begin{bmatrix} A_{12} \\ q_j^T A_{22} \\ q_j^T \end{bmatrix} = \operatorname{rank}[A_{12}]. \tag{6}$$

Proof. For proof see [3]. □

Now, assuming that the condition (6) holds, we give the explicit design of the observer:

Choose an arbitrary $\alpha > 0 \in \mathbb{R}$ and compute the parameters ℓ_j^T, g_j^T and h_j^T as

$$\ell_j^T = q_j^T (A_{22} + \alpha I_{n_2}) A_{12}^{\dagger}, \tag{7a}$$
$$g_j^T = q_j^T B_2 - \ell_j^T B_1, \tag{7b}$$
$$h_j^T = q_j^T A_{21} - \ell_j^T (\alpha I_{n_1} + A_{11}), \tag{7c}$$

where $A_{12}^{\dagger}$ is the Moore-Penrose pseudo-inverse of A_{12}.

Now, we show that with above choice of the parameters, the observer (5) has the error dynamics $\dot{e}_j(t) = -\alpha e_j$. For that, let us consider the error dynamics $\dot{e}_j(t)$. From (3), (4)

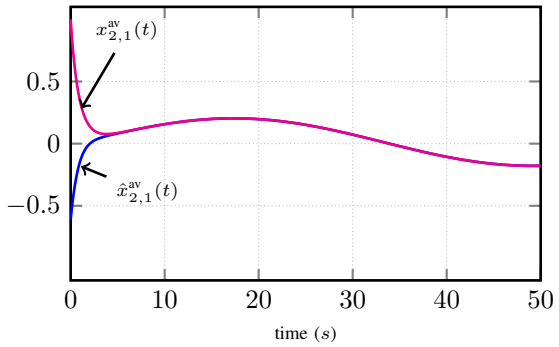

Fig. 5. For the network system in Example 1, where the outlier is node 7 (i.e., $j_o = 1$): estimated ($\hat{x}^{\text{av}}_{2,1}(t)$) and original ($x^{\text{av}}_{2,1}(t)$) average of all unmeasured states except the outlier's.

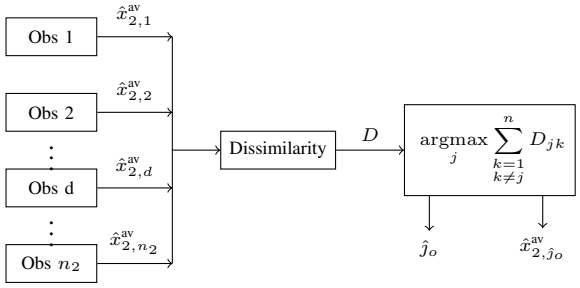

Fig. 6. Structure of the estimation and detection algorithm presented in Section IV for the identification of the outlier and estimation of the average excluding it. Here, obs stands for observer and D is the dissimilarity matrix. $\hat{j}_o$ and $\hat{x}^{\text{av}}_{2,\hat{j}_o}$ are the detected position of the outlier and the average estimate excluding the outlier respectively.

and (5) , we have

$$\begin{aligned}\dot{e}_j(t) = -\alpha e_j(t) + (-\alpha \ell_j^T + q_j^T A_{21} - \ell_j^T A_{11} - h_j^T)x_1(t) \\ +(q_j^T A_{22} - \ell_j^T A_{12} + \alpha q_j^T)x_2(t) \\ +(q_j^T B_2 - g_j^T - \ell_j^T B_1)u(t).\end{aligned} \tag{8}$$

Assuming that condition (6) holds, and with ℓ_j, h_j and g_j as in (7) we will now show that the following conditions are satisfied:

$$q_j^T A_{21} - \alpha \ell_j^T - \ell_j^T A_{11} - h_j^T = 0, \tag{9a}$$
$$q_j^T A_{22} + \alpha q_j^T - \ell_j^T A_{12} = 0, \text{ and} \tag{9b}$$
$$q_j^T B_2 - g_j^T - \ell_j^T B_1 = 0. \tag{9c}$$

and hence $\dot{e}_j(t) = -\alpha e_j(t)$. It can be seen that if the condition (6) holds, then the row vectors c_j and $c_j^T A_{22}$ lie in the row space of A_{12}. Then, the vector $\alpha q_j^T + q_j^T A_{22}$ also lies in the row space of A_{12} for any $\alpha \in \mathbb{R}$. Therefore, it holds that $(\alpha q_j^T + q_j^T A_{22})(I - A_{12}^\dagger A_{12}) = 0$. Hence,

$$\alpha q_j^T + q_j^T A_{22} = (\alpha q_j^T + q_j^T A_{22}) A_{12}^\dagger A_{12}. \tag{10}$$

It can be seen that with the choice of ℓ_j in (7a), (10) is equivalent to (9b). Finally, it can also be seen that with the choice of g_j and h_j in (7b) and (7c), the conditions (9a) and (9c) are satisfied.

Now, we illustrate the estimation using Example (1).

Example 1 (Continued)**.** *Here, we have the knowledge that the outlier is at node 7, which is the first node in the set of unmeasured nodes, i.e., $j_o = 1$. We want to estimate $x^{\text{av}}_{2,1}$. Note that the condition* (6) *with $j = 1$ is satisfied, as* $\text{rank}\begin{bmatrix} A_{12} \\ q_1^T A_{22} \\ q_1^T \end{bmatrix} = 4 = \text{rank}(A_{12})$*, where* $q_1^T = \begin{bmatrix} 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$*. Therefore, we can design an observer of the form* (5) *to estimate $x^{\text{av}}_{2,1}$ with an arbitrary rate of convergence. We fix $\alpha = 1.2$ for observer* (5) *and ℓ_1, g_1 and h_1 are computed by* (7)*. We initialize the observer with $w_1(0) = 0$. Notice that this gives $\hat{x}^{\text{av}}_{2,1}(0) = w_1(0) + \ell_1^T y(0) = -0.8491 \neq 0$, as it can be seen in Figure 5.*

It can be seen in Figure 5 *that our observer is able to estimate the average $x^{\text{av}}_{2,1}$, i.e., the average of unmeasured states excluding the outlier at node $n_1 + 1 = 7$.*

IV. Outlier at an unknown position

In this section, we consider the case when the position of the outlier i.e., j_o is unknown. We propose a method to estimate the average of the unmeasured states excluding the outlier and detecting it simultaneously. Our approach is to estimate the averages $x^{\text{av}}_{2,j}$ for all possible $j \in \{1, \ldots, n_2\}$ using a bank of scalar observers of the form (5) and then compare the estimates in order to detect the outlier. For the comparison, we propose a dissmilarity criterion inspired by the distance based dissimilarity used in signal processing. Figure 6 illustrates the process we use for the detection of the outlier which also gives us the required average estimate. At first, we provide a necessary and sufficient condition for the design of the observer of the form (5) for every possible $j \in \{1, \ldots, n_2\}$. Then we define a dissimilarity criterion to differentiate between the estimates. After that, we define a very general optimization problem to detect the outlier. At the end, we illustrate the method with the help of Example 1.

A. Existence condition for the bank of observers

In order to design observers of the form (5) to reconstruct $x^{\text{av}}_{2,j}$ for all j, the condition in Theorem 1 must be satisfied for all $j \in \{1, \ldots, n_2\}$. The resulting condition can be stated as the following theorem.

Theorem 2. *A tunable and scalar observer for $x^{\text{av}}_{2,j}$ exists for all $j \in \{1, \ldots, n_2\}$ if and only if*

$$\text{rank}(A_{12}) = n_2. \tag{11}$$

Proof. From Theorem 1, to estimate $x^{\text{av}}_{2,j}$ for all j, a tunable, scalar observer exists if and only if (6) holds for all j.

This is equivalent to

$$\operatorname{rank}\begin{bmatrix} A_{12} \\ q_1^T A_{22} \\ \vdots \\ q_{n_2}^T A_{22} \\ q_1^T \\ \vdots \\ q_{n_2}^T \end{bmatrix} = \operatorname{rank}[A_{12}]. \tag{12}$$

Now define $Q = (I_{n_2} - \mathbf{1}_{n_2}\mathbf{1}_{n_2}^T)$, where $\mathbf{1}_n \in \mathbb{R}^n$ denotes the vector of all ones, and $P = \begin{bmatrix} q_1^T \\ \vdots \\ q_{n_2}^T \end{bmatrix} = -\frac{1}{n_2-1}Q$, so that the left-hand side of (12) is equal to $\begin{bmatrix} A_{12} \\ PA_{22} \\ P \end{bmatrix}$. We can see that $\operatorname{rank}(P) = n_2$ by showing that $\operatorname{rank}(Q) = n_2$ and we show $\operatorname{rank}(Q) = n_2$ by showing that $\det Q \neq 0$.
For this, notice that

$$\begin{bmatrix} I_{n_2} & 0 \\ \mathbf{1}_{n_2}^T & 1 \end{bmatrix}\begin{bmatrix} Q & -\mathbf{1}_{n_2} \\ 0 & 1 \end{bmatrix}\begin{bmatrix} I_{n_2} & 0 \\ -\mathbf{1}_{n_2}^T & 1 \end{bmatrix} = \begin{bmatrix} I_{n_2} & -\mathbf{1}_{n_2} \\ 0 & 1 - \mathbf{1}_{n_2}^T\mathbf{1}_{n_2} \end{bmatrix}.$$

Taking determinant of matrices on both the sides, we have $\det(Q) = 1 - n_2 \neq 0$.
Therefore, $\operatorname{rank}(Q) = n_2$ and hence, $\operatorname{rank}(P) = n_2$.

Finally, since $\operatorname{rank}(P) = n_2$, we have $\operatorname{rank}\begin{bmatrix} A_{12} \\ PA_{22} \\ P \end{bmatrix} = n_2$. Therefore, (12) holds if and only if $\operatorname{rank}(A_{12}) = n_2$. □

Now we present some remarks on how restrictive the condition in Theorem 2 is.

Remark 1: It can be seen from (2), that $A_{12} \in \mathbb{R}^{n_1 \times n_2}$ and the condition (11) is $\operatorname{rank}(A_{12}) = n_2$. It implies $n_1 \geq n_2$, that is the number of measured nodes must be greater than or equal to the number of unmeasured nodes.

Remark 2: From the condition (11), A_{12} is full column rank so it cannot have an all zero column. Therefore, Theorem 2 requires that for every unmeasured node, there exists an edge pointing to some measured node, i.e., for every unmeasured node j there is an edge (i, j) with i a measured node (recall that the edge (i, j) is depicted as an arrow $j \to i$, representing an influence of state x_j on the dynamics of state x_i).

Now, we proceed towards the problem of detection of the outlier in the next section. For this, we propose to run a bank of observers and then compare the estimates we obtain.

B. Outlier detection

In this subsection, we define a dissimilarity matrix and an optimization problem in order to detect the outlier.

1) Dissimilarity criterion: Here we define a dissimilarity criterion in order to differentiate between the estimates obtained in the previous section with a goal in mind to pick the one which is without the outlier. Dissimilarity criteria of this kind are used in signal processing. For instance, [13] defines dissimilarity as pairwise Euclidean distance between two signals.

Consider the estimates $\hat{x}^{\text{av}}_{2,j}(t)\ \forall j \in \{1, \ldots, n_2\}$ obtained in the previous section, we define their dissimilarity at time t as

$$D_{jk}(t) = \int_0^t e^{-\beta(t-\tau)} \mid \hat{x}^{\text{av}}_{2,j}(\tau) - \hat{x}^{\text{av}}_{2,k}(\tau) \mid d\tau \quad \text{for } \beta > 0, \tag{13}$$

where $\hat{x}^{\text{av}}_{2,j}(\tau)$ is the estimated average of the unmeasured nodes except node j, at time τ. The matrix $D = [D_{jk}]$ so obtained is called the *Dissimilarity matrix*. Note that, D is a non-negative, symmetric matrix with zero diagonal elements. Here, the idea is to measure how far are the estimates from each another.

Note that the system is assumed to have an outlier at j_o, i.e., there is a significant difference between the average x^{av}_{2,j_o} excluding j_o and the average $x^{\text{av}}_{2,k}$ excluding any other node k. Moreover, the outlier is unique. Hence, for any j and k different from j_o, $|x^{\text{av}}_{2,j_o} - x^{\text{av}}_{2,k}|$ is large and $|x^{\text{av}}_{2,j} - x^{\text{av}}_{2,k}|$ is small (at least as an integral over time, as in the dissimilarity matrix). Recall from Section III that each estimate $\hat{x}^{\text{av}}_{2,j}$ converges to the corresponding correct average $x^{\text{av}}_{2,j}$, therefore we also have that $|\hat{x}^{\text{av}}_{2,j_o} - \hat{x}^{\text{av}}_{2,k}|$ is large and $|\hat{x}^{\text{av}}_{2,j} - \hat{x}^{\text{av}}_{2,k}|$ is small, except possibly for an initial transient. For this reason, we can say that $\hat{j}_o(t)$ is the detected position of the outlier at time t if $\hat{j}_o(t)-$th row sum of the dissimilarity matrix at time t is the largest. In particular,

$$\hat{j}_o(t) = \operatorname*{argmax}_j \sum_{\substack{k=1 \\ k\neq j}}^{n} D_{jk}(t). \tag{14}$$

The above processes as depicted in Figure 6 can be put altogether as follows:

1) Using a bank of n_2 observers of the form (5), obtain the average estimates $\hat{x}^{\text{av}}_{2,j}(t)$ for all j from 1 to n_2.
2) Compute the dissimilarity matrix $D_{jk}(t)$ defined in (13).
3) From the dissimilarity matrix, detect the outlier $\hat{j}_o(t)$ at time t given by (14). Then choose the corresponding average estimate $\hat{x}^{\text{av}}_{2,\hat{j}_o}(t)$ obtained from the bank of observers, which excludes $x_{n_1+\hat{j}_o}$.

Now, we illustrate the method with the Example 1. We will see that indeed $\hat{j}_o(t)$ converges to j_o, the true position of the outlier.

Example 1 (continued)**.** *The position of the outlier in Example* 1 *is* $j_o = 1$ *but here we assume that this information is unknown. We want to detect the outlier position and obtain the average estimate excluding the outlier. Note that the condition* (11) *is satisfied with* $\operatorname{rank}(A_{12}) = 4 = n_2$. *Therefore, we can design a bank of* n_2 *observers of the form* (5) *each of which estimates the average of all but one unmeasured state at a time. Here, each observer is designed with* $\alpha = 1.2$*, and the parameters* ℓ_j^T, g_j^T, h_j^T *computed with* (7) *and initialization* $w_j(0) = 0$*. The estimates obtained by this bank of observers are depicted in Figure 7. We compute the dissimilarity matrix as in* (13) *and then follow the process*

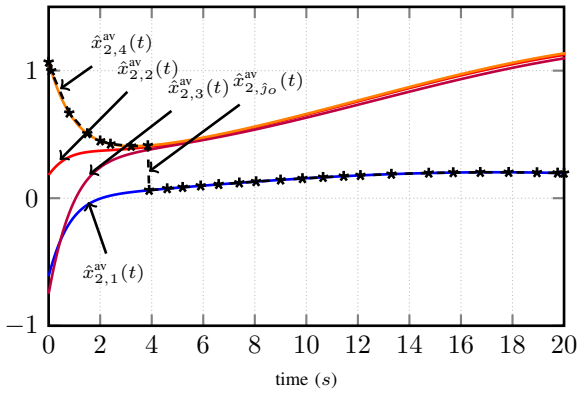

Fig. 7. Estimated trajectories of the averages of unmeasured states in Example 1 excluding one node at a time i.e. $\hat{x}^{\text{av}}_{2,j}$ $\forall j \in \{1,\ldots,n_2\}$, the output of the bank of observers. The dash-starred line is $\hat{x}^{\text{av}}_{2,\hat{j}_o}(t)$, the estimated average without the detected outlier at time t. It can be seen that $\hat{x}^{\text{av}}_{2,\hat{j}_o}(t)$ converges to $\hat{x}^{\text{av}}_{2,1}(t)$ as the outlier is at $j_o = 1$.

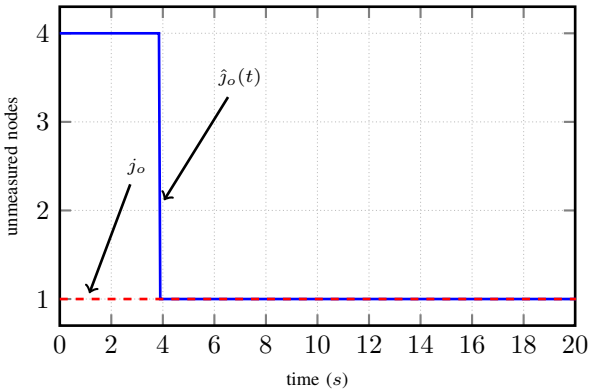

Fig. 8. Detected position of the outlier in the set of the unmeasured nodes at time t i.e. $\hat{j}_o(t)$ in Example 1. $j_o = 1$ is the true position of the outlier. At first, the method identifies unmeasured node 4 as the outlier but it converges quickly to the true position of the outlier that is $j_o = 1$.

described above to detect the outlier using (14)*. With* $\beta = 10$ *in* (13)*, we compute* D *and detect the outlier along with the estimation of averages using* (13) *and* (14)*. Figure* 8 *shows that indeed the proposed method is able to detect the outlier position* $j_o = 1$*, i.e., node* $n_1 + j_o = 7$ *in the network depicted in Figure 2. In Figure 8, it can be seen that at first the method identifies unmeasured node 4 (node 8 in the network) as the outlier but it converges quickly to the true position, i.e.,* $j_o = 1$ *(node* $n_1 + j_o = 7$ *in the network). One possible reason of this delay in detection could be the delay in convergence of the estimated value to the true value. Note that we already have the average estimates* $\hat{x}^{\text{av}}_{2,j}(t)$ $\forall j \in \{1,\ldots,n_2\}$ *obtained from the bank of observers. From them, we obtain* $\hat{x}^{\text{av}}_{2,\hat{j}_o}(t)$*, which is initially equal to* $\hat{x}^{\text{av}}_{2,4}$ *and then equal to* $\hat{x}^{\text{av}}_{2,1}$*, as illustrated by the dash-starred black line in Figure 7. This is consistent with the quantity we aim at reconstructing: the average estimate of the unmeasured states excluding the outlier, i.e.,* $\hat{x}^{\text{av}}_{2,1}$.

V. Concluding remarks

Average state reconstruction with the help of some sensor measurements can give unexpected results if there is an outlier among the unmeasured nodes. A method to estimate the average excluding the outlier has been proposed. For that, design of a scalar and tunable observer has been given along with the condition under which a bank of these observers can be designed to estimate the average of the unmeasured states while excluding an element at every possible position. Moreover, the problem of detection of the existing outlier has also been addressed by proposing a dissimilarity based matrix inspired from the euclidean distance based dissimilarity matrix used in signal processing.

Future works will be focused on cases where the system matrices are not fully known or there are multiple outliers. Moreover, sequential group testing methods can be investigated in order to reduce the number of observers.

Acknowledgement

This work is supported by European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (ERC-Adg no. 694209, Scale-FreeBack, website: `http://scale-freeback.eu`)

References

[1] M. U. B. Niazi, C. Canudas-de-Wit, and A. Y. Kibangou, "Average observability of large-scale network systems," *European Control Conference (ECC)*, 2019.

[2] T. Sadamoto, T. Ishizaki, and J. I. Imura, "Average state observers for large-scale network systems," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 4, pp. 761–769, 2017.

[3] M. U. B. Niazi, D. Deplano, C. Canudas-de-Wit, and A. Y. Kibangou, "Scale-free estimation of the average state in large-scale systems," *IEEE Control Systems Letters*, vol. 4, no. 1, pp. 211–216, 2020.

[4] D. M. Hawkins, *Identification of Outliers*. Chapman and Hall, 1980.

[5] P. J. Rousseeuw and A. M. Leroy, *Robust Regression and Outlier Detection*. USA: John Wiley & Sons, Inc., 1987.

[6] A. Arning, A. Rakesh, and P. Raghavan, "Method for deviation in large databases," *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, p. 164–169, 1996.

[7] E. M. Knorr, R. T. Ng, and V. Tucakov, "Distance-based outliers: algorithms and applications," *The VLDB Journal*, vol. 3, pp. 237–253, 2000.

[8] V. Hautamaki, I. Karkkainen, and P. Franti, "Outlier detection using k-nearest neighbour graph," in *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, vol. 3, 2004, pp. 430–433 Vol.3.

[9] N. Abe, B. Zadrozny, and J. Langford, "Outlier detection by active learning," *ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pp. 504–509, 2006.

[10] B. Sheng, Q. Li, W. Mao, and W. Jin, "Outlier detection in sensor networks," *Proceedings of the International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, pp. 219–228, 2007.

[11] J. Branch, B. Szymanski, C. Giannella, Ran Wolff, and H. Kargupta, "In-network outlier detection in wireless sensor networks," in *26th IEEE International Conference on Distributed Computing Systems (ICDCS'06)*, 2006, pp. 51–51.

[12] H. Wang, H. Li, J. Fang, and H. Wang, "Robust Gaussian Kalman filter with outlier detection," *IEEE Signal Processing Letters*, vol. 25, no. 8, pp. 1236–1240, 2018.

[13] Y. Xu, S. M. Salapaka, and C. L. Beck, "Aggregation of graph models and Markov chains by deterministic annealing," *IEEE Transactions on Automatic Control*, vol. 59, no. 10, pp. 2807–2812, 2014.