A sliding mode fault detection scheme for corrupted measurement data exchange in a network of dynamical systems

Prathyush P Menon and Christopher Edwards
University of Exeter, Exeter, EX4 4QF, United Kingdom
P.M.Prathyush@exeter.ac.uk, C.Edwards@exeter.ac.uk

Abstract—This paper considers the problem of detection of a specific class of faults/failures in a network of dynamical systems, whereby an agent communicates corrupted measurement data to its neighbouring agents, even though the local measurement information at the agent itself is not faulty/corrupted. A network of linear dynamical systems is considered in which the interconnections among the dynamical systems occur via information exchange due to the use of distributed consensus type control laws. An approach based on sliding mode observers is developed to address this problem. A distributed network of sliding mode observers is proposed to detect this specific class of faults. The proposed fault isolation scheme is however centralised in nature. The complexity associated with the design of the observers is of the order of the node level dynamical systems and not related to the size of the network. The approach is demonstrated on an example taken from the existing recent literature.

I. INTRODUCTION

Control of network of dynamical systems using distributed and consensus control methods [1] involve the exchange of output information among the agents of the network over cyber-space. Since these networks are becoming more prevalent and conspicuous, there is a significant potential danger arising from the corruption of the information exchanged [2]–[4]. Consequently there is an emerging requirement to develop efficient monitoring methods to address such issues.

Recently, there has been a growing interest in understanding and developing methods for identifying corruptions/faults in the measurement signals that are communicated among the agents in a network of linear dynamical systems [3], [6]–[8]. The agents in the network are controlled by a distributed or consensus type control law. As argued in [3], [6], a key issue is to detect when a communicated measurement signal has been corrupted, which is often described as a cyber data attack on the network [8]. There exist many different approaches to the problem of detection and isolation of faults [9]. However here, since the structure is distributed, and the fault is a corrupted communication of measurements, conventional observers might not be well suited to detect this type of fault: see for example [5]–[7]. Cumulative sum and classic sequential probability ratio tests have been successfully used in [5] to detect anomalous behaviour in the closed loop model responses due to different faults classified as stealthy, surge, bias and geometric attacks. In [8], methods based on Luemberger observers have been developed to detect cyber attacks, and the efficacy of the approach is demonstrated using a numerical example of the swing dynamics of a power network. A formulation based on the stochastic hypothesis testing methodology has been employed in [10] to detect so-called geometrical attacks, i.e., attacks on the topology of the network, which has also been demonstrated using the swing dynamics of a power network assuming stochastic power injections.

In [11], decentralised observers are designed locally and subsequently these node level estimates are combined to generate estimates of the entire network. In [12], full order observers of the entire network (in a completely different ideology to that in [11]), have been designed at node level, and then the exchanged information among the observers is used to realise a fault diagnosis scheme for the network. In [6], unknown input observers of the order of the network, have been developed at each node to detect the corrupted information and to isolate the corrupted nodes. Also, in [6], a distributed implementation of unknown input observers for networks of dynamical systems has been proposed and methods have been put forward to isolate the affected agents. In [7], the authors have extended the method from [6] to a specific class of imprecise network models. Specifically the authors aim to tackle the issues associated with the addition and removal of communication links, which infact is a fixed node graph with time varying Laplacian matrix structure. In all this work, the authors have given a clear message to the community that “distributed fault detection for systems comprised of a network of autonomous nodes is still in its infancy” [6]. This is the main motivation for the work proposed in this paper.

Our aim in this paper is to expand the use of sliding mode observers [13]–[15] to address the specific class of problems appearing in the recent literature [6], [7], related to the detection and isolation of data corruption in the measurements that the agents communicate over a fixed network. The detection of corrupted data can be achieved using a distributed structure, but currently isolation requires a monitoring coordination layer which is centralised.

Previously, sliding mode observers with a centralised architecture have been designed by the authors for a networked class of systems in [16], [17]. However, the structure of the observers in this paper is distinct from this previous work and explicitly addresses the problem of corruption in the exchange of information among the agents. In the proposed network of observers, each individual observer exchanges information with ‘neighbouring observers’ in a topology which is identical to that of the original network. For this reason the scheme is considered as distributed. In this way, the design of the distributed observers has a complexity which is of the order of the node level dynamics and is independent of the dimension of the network. A centralised coordination layer interacting with the distributed observers, carries out the isolation once corruption has been detected. The synthesis of the distributed observer uses concepts from the area of parameter varying systems [18].
II. SYSTEM DESCRIPTION

A network of \(N\) identical linear dynamical systems, interconnected over an information network, can be represented in terms of a graph \(G(N, E)\), where \(N\) denotes the number of nodes in the network, i.e., the dynamical systems, indexed as \(1, \ldots, N\), and \(E\) represents the edges, which are the interactions between agents that occur through distributed/consensus type control structures. Suppose each node has the following dynamics:

\[
\begin{align*}
\dot{x}_i(t) &= A x_i(t) + Bu_i(t) \\
\dot{w}_i(t) &= C x_i(t) \\
y_i(t) &= C x_i(t)
\end{align*}
\]

for \(i = 1, \ldots, N\), where \(x_i(t) \in \mathbb{R}^n\), \(u_i(t) \in \mathbb{R}^m\), \(w_i(t) \in \mathbb{R}^p\) and \(y_i(t) \in \mathbb{R}^p\) which represents the states, control inputs, output measurements available locally at each node, and the output measurements that the node communicates over the information network. Despite the fact that \(w_i(t)\) and \(y_i(t)\) are associated with fundamentally identical measured values obtained from a sensor located at \(i\)th node, the output measurements \(y_i(t)\) which are communicated over the information network may be prone to be corrupted by network attack. It is assumed that the source of the corruption is associated with the network exchange while the local measurement \(w_i(t)\), directly associated with node \(i\), is secure. This representation has been previously adopted in [6], [8]. Here it is assumed that the control input \(u_i(t)\) employed for the system in (1) is distributed in nature (of a consensus type) so that

\[
u_i(t) = -\alpha(t) \sum_{j \in N_i} a_{ij} (w_i(t) - y_j(t))
\]

where \(N_i \subset \{1, \ldots, N\}\{i\}\) represents a set constituting all the indices of neighboring agents of the \(i\)th node, i.e., all the agents with which the \(i\)th agent can communicate within the network. The term \(a_{ij} = 1\), if \(i \neq j\) and there exists information exchange between the \(i\)th and \(j\)th agents, otherwise \(a_{ij} = 0\). The diagonal element \(a_{ii} = 0\) for \(i = 1, \ldots, N\), i.e., there are no self loops. The resulting symmetric matrix \(A(G) = [a_{ij}] \in \mathbb{R}^{n \times n}\) defines the adjacency matrix of the graph \(G\). Since \(G\) is assumed to be undirected, the degree of a given node, \(k_i\) is the cardinality of the neighbourhood set \(N_i\), that is \(|N_i|\), and is equal to the number of nodes that the \(i\)th agent can communicate with among the \(N\) nodes of the graph \(G\). The Laplacian matrix \(\Delta(G)\) of graph \(G\) is a diagonal matrix containing the node degrees of \(G\) on the diagonal. The Laplacian \(\mathbf{L}\) is defined as \(\mathbf{L} = \Delta - A\) [1]. The coupling strength \(\alpha(t)\) is allowed to be time-varying but known, and to belong to a positive interval, upperbounded by \(\overline{\alpha} > 0\).

Suppose the corrupted measurement signal communicated over the network sent from the \(j\)th node is modelled as an additive fault:

\[
y_j(t) = w_j(t) + F_f j(t)
\]

for \(j = 1, \ldots, N\), where \(F \in \mathbb{R}^{m \times q}\) is a selector matrix and \(f_j(t)\) is a piecewise continuous signal representing a cyber attack [5], [7], [8]. Here the assumptions regarding the dimensions of the signals are \(n \geq p = m > q\). The case of \(m = q\) is treated as a special case since this scenario results in the loss of design freedom in the framework that will be proposed later.)

A. System in presence of corrupted information

When the network data becomes corrupted, then the control law (4) becomes

\[
u_i(t) = -\alpha(t) \sum_{j \in N_i} a_{ij} (w_i(t) - y_j(t)) + \alpha(t) \sum_{j \in N_i} a_{ij} F_f j(t)
\]

\[
= -\alpha(t) \sum_{j \in N_i} a_{ij} C (x_i(t) - x_j(t)) + \alpha(t) \sum_{j \in N_i} a_{ij} F_f j(t)
\]

for \(i = 1, \ldots, N\). Hence in the presence of communication of corrupted measurements (viewed as compromised measurements due to a cyber attack), using (6) in (1), the dynamics at the \(i\)th level become

\[
\dot{x}_i(t) = Ax_i(t) - \alpha(t) \sum_{j \in N_i} a_{ij} BC (x_i(t) - x_j(t)) + \alpha(t) \sum_{j \in N_i} a_{ij} BF_f j(t)
\]

\[
+ \alpha(t) \sum_{j \in N_i} a_{ij} F_f j(t)
\]

\[
w_i(t) = C x_i(t)
\]

Next define \(D := BF\) and

\[
\xi_i(t) := \alpha(t) \sum_{j \in N_i} a_{ij} f_j(t)
\]

It is assumed that the unknown signal \(|\xi_i(t)| \leq K(t)\) where the worst case upper bound \(K(t)\) is assumed to be known. Define

\[
\Gamma := BC
\]

then making use of (11), the dynamics of the \(i\)th node of the attacked network can be rewritten as

\[
\dot{x}_i(t) = Ax_i(t) + D \xi_i(t) - \alpha(t) \sum_{j = 1}^N L_{ij} \Gamma x_j(t)
\]

\[
w_i(t) = C x_i(t)
\]

B. Assumptions on the framework

Assume the system represented in (12) - (13) satisfies the following [6], [15], [17]:

- \((A, D, C)\) is a minimal realization with \(C\) and \(D\) full column and row rank respectively;
- \((A, D, C)\) is minimum phase;
- \(\text{rank}(CD) = q\).

Since \(C \in \mathbb{R}^{p \times n}, B \in \mathbb{R}^{n \times m}\) and \(F \in \mathbb{R}^{m \times q}\), and by definition \(D = BF\), the assumption \(\text{rank}(CD) = q\) together with the rank relationship \(\text{rank}(CD) = \text{rank}(CBF) \leq \min\{\text{rank}(CB), \text{rank}(F)\} = q\) implies \(\text{rank}(CB) \geq q\). Although the term \(B\) is not explicitly present in (12) its presence implicitly appears through the term \(\Gamma = BC\), emanating from the underlying structure of the problem.

Here the most generic case will be considered where \(m > \text{rank}(CB) > q\). Special cases when \(\text{rank}(CB) = \text{rank}(CB) = q\) and when \(\text{rank}(CB) = m\) will be discussed briefly at the end of the section.
With the imposed assumptions and structure on the triple \((A, D, C)\) at each node level, a nonsingular linear transformation \(x_i \rightarrow T_0 x_i\), for \(i = 1, \ldots, N\) exists and provides (with slight abuse of notation) the following canonical structure to the triple \((A, D, C)\) in (12):

\[
A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad D = \begin{bmatrix} 0 \\ D_2 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & C_2 \end{bmatrix} \tag{14}
\]

where \(A_{11} \in \mathbb{R}^{(n-p) \times (n-p)}\), \(D_2 \in \mathbb{R}^{p \times q}\) and \(C_2 \in \mathbb{R}^{p \times p}\) is orthogonal. Define \(A_{21}\) as the top \(p \times q\) rows of \(A_{21}\). By construction, the pair \((A_{11}, A_{21})\) is detectable and the unobservable modes of \((A_{11}, A_{21})\) are the invariant zeros of \((A, D, C)\) [15]. It is possible that the matrix \(D_2\) (conveniently written as a \(\mathbb{R}^{P \times Q}\) matrix to be consistent with the partition in (14)) can further be written as \(D_2 = \text{Col}(0 \quad D_0)\) where \(D_0 \in \mathbb{R}^{q \times q}\) is non-singular [15]. Furthermore, in this coordinate system,

\[
\Gamma = BC = \begin{bmatrix} 0 & \Gamma_{12} \\ 0 & \Gamma_{22} \end{bmatrix} \tag{15}
\]

where \(\Gamma_{12} \in \mathbb{R}^{(n-p) \times p}\) and \(\Gamma_{22} \in \mathbb{R}^{p \times p}\). Note this special structure follows from the specific form of the output distribution matrix (14). This coordinate system will be used for the observer synthesis.

### III. Sliding Mode Observer Design

The proposed observer dynamical system associated with the \(i\)th node, using only local node level measurements, is:

\[
\dot{x}_i(t) = A\tilde{x}_i(t) - \alpha(t) \sum_{j=1}^{N} \mathcal{L}_{ij} \Gamma \dot{x}_j(t) - G_ie_{y_i}(t) - G_n \nu_i(t)
\]

for \(i = 1, \ldots, N\), where \(e_{y_i}(t) = w_i(t) - C\tilde{x}_i(t)\). The discontinuous injection term, \(\nu_i(t) \in \mathbb{R}^{p}\), for \(i = 1, \ldots, N\) is given by:

\[
\nu_i = -r(t) \frac{e_{y_i}(t)}{\| e_{y_i}(t) \|}, \quad \text{if} \quad e_{y_i} \neq 0
\]

where \(r(t) : \mathbb{R}^+ \rightarrow \mathbb{R}^+\) is a scalar modulation function which will be specified explicitly in the sequel. In (16), \(G_i \in \mathbb{R}^{n \times p}\) and \(G_n \in \mathbb{R}^{p \times p}\) are the observer gain matrices to be designed. By choice of suitable gains, \(G_i\) and \(G_n\), the state estimation error is forced to an a-priori defined sliding surface in finite time, with certain level of performance. Consistent with the partition in (14), partition the estimated state vector as \(\hat{x}_i = \text{Col}(\hat{x}_{i1}, \hat{x}_{i2})\) such that \(\hat{x}_{i1} \in \mathbb{R}^{(n-p)}\) and \(\hat{x}_{i2} \in \mathbb{R}^{p}\). One choice of nonlinear injection gain \(G_n \in \mathbb{R}^{p \times p}\) in this co-ordinate framework is

\[
G_n = \begin{bmatrix} LP_2^{-1}C_i^T \\ P_2^{-1}C_i^T \end{bmatrix}
\]

where \(L = \begin{bmatrix} 0 & L^o \end{bmatrix} \in \mathbb{R}^{(n-p) \times (n-p)}\) (any) with \(L^o \in \mathbb{R}^{(n-p) \times (p-q)}\) and \(P_2 \in \mathbb{R}^{p \times p}\) is a symmetric positive definite (s.p.d) matrix [19]. A specific choice for \(P_2\) will be specified later. The objective here is to determine suitable gain matrices \(L^o\) and \(G_i\), which represent the design freedom, such that the dynamics of the state estimation error is asymptotically stable. Defining \(e_i(t) := x_i(t) - \hat{x}_i(t)\), the error dynamics are:

\[
\dot{e}_i(t) = Ae_i(t) + D\xi_i(t) - \alpha(t) \sum_{j=1}^{N} \mathcal{L}_{ij} \Gamma j e_j(t) + G_ie_{y_i}(t) + G_n \nu_i(t)
\]

for \(i = 1, \ldots, N\). Compatible with (14), the error states \(e_i\) is partitioned so that \(e_i = \text{Col}(e_{1i}, e_{2i})\) where \(e_{1i} \in \mathbb{R}^{n-p}\) and \(e_{2i} \in \mathbb{R}^p\). The idea is to force sliding on the surface [21]

\[
S = \{(e_1, \ldots, e_N) \mid e_{2i}(t) = 0, ..., e_{2N}(t) = 0\}
\]

in finite time \(T\).

#### A. Reduced order sliding motion

Suppose a sliding motion is attained on the surface \(S\) defined in (20) in finite time. During the reduced order sliding motion, \(e_{2i} = \dot{e}_{2i} = 0\). From (19) and exploiting the structure of \(\Gamma\) from (15), during sliding, it follows that \(\Gamma e_j = 0\) if \(e_{2j} = 0\) and therefore

\[
\begin{bmatrix} \dot{e}_{1i}(t) \\ 0 \end{bmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} e_{1i}(t) + \begin{bmatrix} 0 \\ D_2 \end{bmatrix} \xi_i(t) + \begin{bmatrix} LP_2^{-1}C_i^T \\ P_2^{-1}C_i^T \end{bmatrix} \nu_{eq,i}(t)
\]

for \(i = 1, \ldots, N\), where \(\nu_{eq,i}(t)\) is the so-called equivalent injection necessary to maintain sliding and represents the ‘average’ switched signal [20], [21].

The equivalent injection, \(\nu_{eq,i}(t)\) necessary to maintain sliding can be calculated from last \(p\) channels of the reduced order dynamics in (21): here

\[
P_2^{-1}C_i^T \nu_{eq,i}(t) = -A_{21}e_{1i}(t) - D_2 \xi_i(t)
\]

Subsequently, by substituting for \(P_2^{-1}C_i^T \nu_{eq,i}(t)\) from (22) in the first \(n-p\) channels of the reduced order dynamics in (21) yields

\[
\begin{bmatrix} \dot{e}_{1i}(t) \\ e_{2i}(t) \end{bmatrix} = \begin{bmatrix} A_{11} - L A_{21} \\ A_{21} \end{bmatrix} e_{1i}(t)
\]

for \(i = 1, \ldots, N\). Notice that since \(L = \begin{bmatrix} L^o & 0 \end{bmatrix}\) where \(L^o \in \mathbb{R}^{(n-p) \times (p-q)}\) a term corresponding to the uncertainty or disturbance \(\xi_i(t)\) is not present in (23), since \(LD_2 = 0\) because of the special structure of \(L\) and \(D_2\). Note that for a stable sliding motion to occur, the gain \(L\) must be chosen to ensure that \(A_{11}\) is stable.

#### B. Reachability

This section will prove that sliding indeed does occur. For a given matrix \(L\), define an invertible transformation \(e_i \mapsto T_L e_i\), for \(i = 1, \ldots, N\), where

\[
T_L := \begin{bmatrix} I_{n-p} & -L \\ 0 & I_p \end{bmatrix}
\]

and \(L := [L^o \quad 0] \in \mathbb{R}^{(n-p) \times p}\) with \(L^o \in \mathbb{R}^{(n-p) \times (p-q)}\) [15]. Consistent with (14), partition the error state vector in the transformed coordinates of (24) as

\[
\begin{bmatrix} \tilde{e}_{1i} \\ \tilde{e}_{2i} \end{bmatrix} = T_L e_i
\]
for \( i = 1,\ldots,N \) where \( \hat{e}_i := e_i - Le_{2i} \). In the new coordinates
\[
\hat{\Gamma} := T_L \Gamma T_L^{-1} = \begin{bmatrix}
0 & \Gamma_{12} - L \Gamma_{22} \\
\Gamma_{12} & \Gamma_{22}
\end{bmatrix}
\tag{26}
\]
Consistent with the partitioned form, define the gains associated with the linear injection as
\[
G_i = \begin{bmatrix}
G_{1i} \\
G_{2i}
\end{bmatrix}
\tag{27}
\]
for \( i = 1,\ldots,N \). In the new coordinates associated with (24), the error system from (19) can be written as
\[
\begin{align*}
\dot{\hat{e}}_i &= \hat{A}_1 \hat{e}_i + \hat{A}_2 e_{2i} + [L_{n-p} - L]G_i e_{y_i} \\
&\quad - \alpha(t) \sum_{j=1}^{N} L_{ij} (\Gamma_{12} - L \Gamma_{22}) e_{2j}
\end{align*}
\tag{28}
\]
\[
\dot{\hat{e}}_i = A_{21} \hat{e}_i + \hat{A}_{22} e_{2i} + D \xi_i(t) + G_2 e_{y_i}
+ P_2^{-1} C_2^T \nu_{eq,i} - \alpha(t) \sum_{j=1}^{N} L_{ij} \Gamma_{22} e_{2j}
\tag{29}
\]
for \( i = 1,\ldots,N \) where
\[
\begin{bmatrix}
\hat{A}_{11} & \hat{A}_{12} \\
A_{21} & \hat{A}_{22}
\end{bmatrix} = T_L \Gamma T_L^{-1}
\tag{30}
\]
and in particular \( \hat{A}_{11} := A_{11} - LA_{21} = (A_{11} - L^a A_{211}) \).
Define
\[
G_2 := -\hat{A}_{22} C_2^T + \Psi C_2^T
\tag{31}
\]
and
\[
G_1 := LG_2 - \hat{A}_{12} C_2^T
\tag{32}
\]
for \( i = 1,\ldots,N \). In (31), \( \Psi \in \mathbb{R}^{p \times p} \) is a designer selected stable matrix. Substituting for \( G_1 \) and \( G_2 \) in (28)-(29), the error dynamics in the transformed coordinates
\[
\dot{\hat{e}}_i = A_1 e_i - \alpha(t) \sum_{j=1}^{N} L_{ij} A_2 e_{2j} + D \xi_i(t) + G v_i
\tag{33}
\]
for \( i = 1,\ldots,N \), where
\[
G = \begin{bmatrix}
0 \\
-P_2^{-1} C_2^T
\end{bmatrix}
\tag{34}
\]
and the matrix \( D \) has the form from (14). The matrices
\[
A_1 := \begin{bmatrix}
A_{11} - LA_{21} & 0 \\
A_{21} & \Psi
\end{bmatrix}
\tag{35}
\]
\[
A_2 := \begin{bmatrix}
0 & \Gamma_{12} - L \Gamma_{22} \\
0 & \Gamma_{22}
\end{bmatrix}
\tag{36}
\]
At a network level the equations associated with (33) can be written as
\[
\dot{e}(t) = ((I_N \otimes A_1) - \alpha(t)(L \otimes A_2)) e(t) + (I_N \otimes D) \xi(t) + (I_N \otimes G) \nu
\tag{37}
\]
where \( e = \text{Col}(e_1, \ldots, e_N) \), the augmented injection signal \( \nu = \text{Col}(\nu_1, \ldots, \nu_N) \) whilst \( \xi = \text{Col}(\xi_1, \ldots, \xi_N) \). First the stability of the time varying linear part of (37), written as
\[
\dot{e}(t) = ((I_N \otimes A_1) - \alpha(t)(L \otimes A_2)) e(t)
\tag{38}
\]
will be investigated using a candidate Lyapunov function \( V(e) = e^T P_a e \) where the positive definite matrix \( P_a \) has the specific structure [22]
\[
P_a := (I_N \otimes P) \quad \text{where} \quad P = \text{Diag}(P_1, P_2)
\tag{39}
\]
where the \( P_1 \in \mathbb{R}^{(n-p) \times (n-p)} \) and \( P_2 \in \mathbb{R}^{p \times p} \) are s.p.d. First change coordinates to obtain a modal decomposition form for (38) as suggested in [23], using a state transformation \( e \mapsto Te = \eta \) where \( T := V^T \otimes I_n \) to have a transformed version of (38) given by
\[
\eta(t) = (I_N \otimes A_1) \eta(t) - \alpha(t)(\Lambda \otimes A_2) \eta(t)
\tag{40}
\]
Also, note that in the new coordinates
\[
V = \eta^T P_a T \eta = \eta^T (V^T \otimes I_n) (I_N \otimes P) (\Lambda \otimes I_n) \eta = \eta^T (I_N \otimes P) \eta
\tag{41}
\]
and so \( V \) retains a block-diagonal representation in the new coordinate system and \( V(\eta) = \sum_{i=1}^{N} \eta_i^T P_i \eta_i \) where \( \eta = \text{Col}(\eta_1, \ldots, \eta_N) \). Furthermore from (40), because \( \Lambda \) is diagonal
\[
\dot{\eta}_i(t) = A_1 \eta_i(t) - \alpha(t) \lambda_i A_2 \eta_i(t)
\tag{42}
\]
for \( i = 1,\ldots,N \) where \( \lambda_i \) is the \( i \)-th eigenvalue of \( \Lambda \). Write \( \rho_i(t) = \alpha(t) \lambda_i \) and so it follows that \( \rho_i(t) \in \mathbb{R} \) where the interval \( R = [0 \quad \tilde{\rho}] \), and the upper limit \( \tilde{\rho} = \alpha \lambda_{max}(\Lambda) \). Since \( \rho \in [0 \quad \tilde{\rho}] \), (42) can be thought of as the LPV system
\[
\dot{\eta}_i(t) = A(\rho(t)) \eta_i(t)
\tag{43}
\]
where \( A(\rho) = (A_1 - \rho A_2) \). The objective now is to select \( L \) and \( P \) such that
\[
PA(\rho) + A(\rho)^T P < 0 \quad \text{for all} \quad \rho \in R
\tag{44}
\]
A sufficient condition for this is that
\[
PA(0) + A(0)^T P < 0
\tag{45}
\]
It is easy to confirm that
\[
PA(\rho) + A(\rho)^T P = PA_1 + A_1^T P + \rho(PA_2 + A_2^T P)
\tag{46}
\]
where
\[
PA_1 + A_1^T P = \begin{bmatrix}
\Xi(P_1, M) & A_{T_1}^T P_2 \\
P_2 A_{21} & Q + Q^T
\end{bmatrix}
\tag{47}
\]
\[
PA_2 + A_2^T P = \begin{bmatrix}
0 & P_1 \Gamma_{12} - M \Gamma_{22}^T \\
(P_1 \Gamma_{12} - M \Gamma_{22}^T)^T & P_2 \Gamma_{22} + \Gamma^T P_2
\end{bmatrix}
\tag{48}
\]
with
\[
\Xi(P_1, M) = P_1 A_{11} - M A_{21} + A_{11}^T P_1 - A_{21}^T M^T
\tag{49}
\]
and \( M = P_1 L, \ Q = P_2 \Psi \). For a given \( \rho \), it is easy to see that the RHS of (46) is affine in the decision variable \( P_1, P_2, M \) and \( Q \) and so testing the conditions in (44)-(45) can be undertaken using standard LMI tools. Once a feasible solution is obtained for \( P_1, P_2, M \) and \( Q \), the required gains for the observer can be recovered as \( L = P_1^{-1} M \) and \( \Phi = P_2^{-1} Q \). Define the modulation function from (17) as
\[
r(t) = \|P_2\| \|D_2\| K(t) + 2 \eta_0 \|P_2\| + \eta_0
\tag{48}
\]
where $K(t)$ is the worst case bound on the faults from (10) and $\gamma_0$ is a positive scalar. Now it will be shown that $V = e^T(I_N \otimes P)e$ is a Lyapunov function for the nonlinear system in (33). Since $(I_N \otimes P)A_0(t) + A_0(t)^T(I_N \otimes P) < 0$ it follows that along the trajectories of (33)

$$\dot{V} \leq 2e^T(I_N \otimes P)(I_N \otimes D)\xi + 2e^T(I_N \otimes P)(I_N \otimes G)\nu$$

and therefore from the special structures of $P$, $D$, and the gain $G$,

$$\dot{V} \leq 2e^T(I_N \otimes P_2D_2)\xi + 2e^T(I_N \otimes C_2^T)\nu$$

where $e_2 = Col(e_2, \ldots, e_{2N})$. Consequently from (49)

$$\dot{V} \leq 2 \sum_{i=1}^{N} e_{2i}^TP_2D_2\xi_i + 2 \sum_{i=1}^{N} e_{2i}^TC_2^T\nu_i$$

$$\leq 2 \sum_{i=1}^{N} \|e_{2i}\|\|P_2\|\|D_2\|\|\xi_i\| - 2 \sum_{i=1}^{N} r(t)\|e_{yi}\|$$

$$= -2 \sum_{i=1}^{N} \|e_{yi}\|\|r(t)\|\|P_2\|\|D_2\|\|K(t)\| < 0 \quad \text{(50)}$$

since $\|e_{yi}\| = \|e_{2i}\|$ because $e_{yi} = C_2e_{2i}$ and $C_2$ is orthogonal, and $r(t)$ satisfies (48). Inequality (50) guarantees that $e(t) \to 0$ as $t \to \infty$ and consequently each $\|e_i(t)\| \to 0$ as $t \to \infty$. It will now be shown that in fact a sliding motion takes place at each node. Note that since, for $i = 1, \ldots, N$, $\|e_i(t)\| \to 0$ as $t \to \infty$ there exists a $t_0 > 0$ such that for all $t > t_0$ the following hold:

$$\|A_2\hat{e}_1(t)\| \leq \eta_0 \quad \text{(51)}$$

$$\alpha(t)\|\Gamma_22\|\sum_{j=1}^{N} \|e_{2j}\| \leq \eta_0 \quad \text{(52)}$$

Consider $V_i = \frac{1}{2}e_{2i}^TP_2e_{2i}$, then along trajectories of (29)

$$\dot{V}_i = e_{2i}^TP_2 (A_2\hat{e}_i + \Psi e_2 + D_2\xi_i(t) + P_2^{-1}C_2^T\nu_i$$

$$-\alpha(t)\sum_{j=1}^{N} \lambda_{ij}\Gamma_22e_{2j})$$

$$\leq \|e_{2i}\|\|P_2\|\|\lambda_2\|\|\hat{e}_1\| + \|D_2\|\|\xi_i(t)\|$$

$$-r(t)\|e_{yi}\| + \alpha(t)\|e_{yi}\|\|P_2\|\|\Gamma_22\|\sum_{j=1}^{N} \|e_{2j}\|$$

since $P_2\Psi + \Psi^TP_2 < 0$. Consequently for all $t > t_0$

$$\dot{V}_i \leq \|e_{yi}\|\|P_2\|\|\xi_i(t)\| + \|D_2\|\|\xi_i(t)\|$$

$$-r(t)\|e_{yi}\| + \|e_{yi}\|\|P_2\|\|\hat{e}_1\|$$

$$\leq -\eta_0\|e_{yi}\| = -\sqrt{2}\eta_0\sqrt{V_i}$$

The last inequality implies $V_i \equiv 0$ in finite time and a sliding motion takes place.

IV. FAULT RECONSTRUCTION AND ISOLATION

During the sliding motion, $e_2$, and its derivatives $e_{2i}$, are zero [20], [21]. From (29), during the sliding motion

$$0 = (I_N \otimes A_2)\hat{e}_1 + (I_N \otimes D_2)\xi_i(t) + (I_N \otimes P_2^{-1}C_2^T)\nu_{eq}(53)$$

where $\nu_{eq}$ is the so-called equivalent injection necessary to maintain a sliding motion [20], [21] in finite time. The equivalent injection can be approximated by either the use of a low pass filter [20] or by approximating the discontinuous unit vector with a smooth sigmoidal [21]. Thus $\nu_{eq,i}$ is an available signal and can be used to obtain an estimate of $\xi_i(t)$. Furthermore $\hat{e}_1 \to 0$, from (53) as $t \to \infty$ and therefore at a node level

$$D_2\xi_i(t) + P_2^{-1}C_2^T\nu_{eq} \to 0$$

and therefore at an individual node level

$$\nu_{eq,i}(t) \to -C_2P_2\xi_i(t)$$

since $C_2$ is orthogonal and thus the equivalent injection reconstructs the faults. Hence, an online approximation for $\xi_i(t)$ is given by

$$\hat{\xi}_i(t) = -(D_2^T D_2)^{-1}D_2^T (C_2P_2)^{-1} \nu_{eq,i}$$

which implies

$$\alpha(t)A\hat{f}(t) = -(I_N \otimes D_f)\nu_{eq}$$

where the augmented fault vector $\hat{f} = Col(\hat{f}_1, \ldots, \hat{f}_N)$ and $\nu_{eq} = Col(\nu_{eq,1}, \ldots, \nu_{eq,N})$. If $det(A) \neq 0$ then

$$\hat{f}(t) = -\frac{1}{\alpha(t)} A^{-1}(I_N \otimes D_f)\nu_{eq}$$

In this way the source of the fault can be identified. The limitation of the proposed approach is the requirement of invertibility of the adjacency matrix; which is not always the case. If $det(A) = 0$, isolation will not be possible. If $A$ has repeated identical columns, it will not be possible to isolate between faults in the associated nodes. However, some degree of isolability will still be possible. Furthermore, as in [6], at present a single node corruption case is considered. However, the proposed method can work for multiple detectable failures, (see the detectability arguments in [8]) but in general will not be possible with the fault isolation layer, i.e., the centralised coordination layer. Furthermore, for certain classes of networks detectability of such corruptions disappear due to symmetry properties of the interconnections.

V. NUMERICAL EXAMPLE

Consider a network of dynamical systems with a distributed proportional/derivative type control law as given in [6]. For this example the matrices in (1) - (3), are:

$$A = [0 \quad 1 \quad 0 \quad -0.01], B = [0 \quad 0.1], C = [1 \quad 1]^T$$

In the control law (4), the coupling strength $\alpha(t)$ is a sinusoidal signal with a bias to ensure $\alpha(t) > 0$. The observer design is carried out in the coordinate system associated with (14).

It is assumed that node 3 communicates an additive fault as described in (5) (with $F$ as unity) to its neighbourhood.
set consisting of nodes 1, 2, 4 and 8. A representation as given in (12)-(13) is obtained, and an observer network of the form given in (16) has been designed. Since in this example \(\text{rank}(CB) = m = 1\), the design freedom collapses as discussed in Remark 1, and \(\Psi\) is chosen as \(-2\). This ensures a solution to both (46) and (47) exists. The simulations have been carried out in Matlab 7.6.0 (R2008a) and the Simulink version 6.5. The solver ode15s has been used with a maximum and minimum time step size of 1e-3 and 1e-5 respectively. Figure 1(a) shows the convergence of the global state estimation error, \(e(t)\), i.e. the difference between the true states and the estimated state value. The output estimation error, \(e_y(t)\), for \(i = 1, \ldots, N\) is shown as an enlarged figure inside Figure 1(a), in which the occurrence of the sliding condition \((\dot{e}_2 = e_2 = 0)\) in finite time can be clearly seen. Here, as an intermediate step, the reconstruction of the \(\xi_i(t)\) signal is shown in the top subfigure of Figure 1(b). In the example, the \(\xi_i(t)\) signals corresponding to nodes 1, 2, 4 and 8 have nonzero values, and the corresponding \(\nu_{eq,i}\) signals have also nonzero values. In the lower subplot of Figure 1(b), the error signal \((\xi_i(t) - \nu_{eq,i}(t))\), for \(i = 1\) is shown. For the network \(\det(A) \neq 0\). Figure 1(c) shows a perfect reconstruction of the corruption that has been introduced at node 3.

VI. CONCLUSIONS

In this paper, a network of sliding mode observers has been proposed to estimate and detect corrupted data information that an agent communicates to its neighbours in the network. This is achieved by creating a separate monitoring network layer in which a distributed observer scheme exchanges local estimates among the other observers, in an identical manner to the original distributed system. Detection is achieved at this distributed observer level. Fault isolation is carried out centrally at a coordination layer, gathering all the equivalent injection signals from the individual observers and making use of knowledge of the underlying global interconnection topology. A numerical example taken from the existing literature has been used to demonstrate the efficacy of the proposed approach.

REFERENCES