On Distributed Pinning Observers for a Network of Dynamical Systems

Christopher Edwards and Prathyush P. Menon

***Abstract*—In this paper, a distributed observer structure is proposed to estimate the states of a large scale network of semi-linear systems interconnected by a positive, time varying coupling strength. The distributed observer comprises distinct sub-observers which require only local node level information and exchange their local state estimates with their 'neighbouring' observers. The key idea here is to use a minimum number, or at least relatively few, measurements from the network being monitored to reduce the sensor requirements. The problem is formulated as a two stage LMI optimization problem.**

***Index Terms*—Observers, Pinning, Complex Networks**

I. INTRODUCTION

The objective of this paper is to develop an observer structure to estimate all the states in a network of interconnected dynamical systems. Such systems of systems arise in applications such as environmental monitoring, formation control, synthetic biological networks and arrays of nano synchronizing devices. Observers which can address the complexity arising in large scale networks are therefore of value and are timely. When employing model based estimation, designers can follow paradigms based on centralized, decentralized, semi-decentralized and distributed architectures. Centralized monitoring schemes require all the measurements to be collected at one 'central' location– thus adopting a star topology with the other agents in the network. However decentralized, semi-decentralized, and distributed monitoring (and control) schemes are often preferred over centralized schemes due to their lower complexity and improved computational aspects.

The research described in this paper can be viewed as part of the ongoing interest in distributed estimation based on extending traditional Kalman filters, Luenberger observers and unknown input observers etc to large scale systems: see for example [3] where the authors employ unknown input observers for fault detection in networked control systems. Here the system is modelled as a scale free network where the nodes are coupled linearly and diffusively [1], [2]. The problem considered in this paper is essentially the dual of the 'pinning control' problem in [2], [7], [21] and the objective here is to observe the entire state of the network from measurements at only a subset of the nodes. This 'pinning observer' problem has been recently posed in [9] and relates to the earlier observer work in [20]. The work described in this paper extends the results in [11] by including the presence of disturbnces/uncertainty at node level.

Also there has been significant progress in terms of developing consensus based ideas in conjunction with filtering algorithms to create distributed filtering algorithms: for example in [4] each Kalman filter estimates the states of the system locally, and then all the filter node estimates, reach consensus by reducing the measure of disagreement [5] between the neighbouring filter nodes. Recent work in this area (for discrete time systems and considering uncertainty) has appeared in [8]. For a specific class of uncertain systems subject to disturbances, estimator networks with a specific structure are designed in [6] wherein the estimator gains and their coupling strengths are computed via LMI conditions, to ensure a certain level of $\mathcal{H}_\infty$ performance. The problem formulation considered in this paper *is not consensus based* although the approach does involve the exchange of information between neighbouring nodes in a graph theoretic sense. Consequently strong parallels exist, and similarities in the underlying analysis can be found.

In this paper, the focus is to develop a network of 'local' observers to asymptotically estimate all the states in the network. The key objective is to use measurements from only a certain subset of the nodes when creating the observer output error injection signals. By imposing the requirement of using a minimum number of (or at least relatively few) injection signals, this implies only a subset of the nodes need to be 'instrumented' with sensors. This is useful in certain engineering applications where sensing is 'costly'. For example in multi-core processors, the build up of temperature has an adverse affect on power and reliability [10]. Monitoring the temperature distribution is important (and a precursor to developing control systems to dissipate the unwanted heat). Incorporating temperature sensors encroaches on the available silicon area on the chips, and therefore comes at a high 'cost' in terms of space utilization. Consequently in such systems there is a clear trade-off between the requirement for establishing an accurate estimate of the temperature distribution while maintaining a minimal footprint in terms of the 'real-estate' costs of deploying physical sensors [10]. Of course this paradigm is not applicable to all large scale problem formulations: for example in certain engineering systems such as wireless networks the most significant 'cost' is associated with communication rather than sensing per se. In such situations the results in this paper are less significant and approaches focussing on 'topology control' such as [22] can be pursued. In this paper a single a-priori off-line design is undertaken rather than online self-organization [23]. Here the network is not spatially dependent and hence the design is a 'one off'. Once the optimization problem is solved it is not revisited. Consequently there are no online computational issues and no requirement for distributing the computations.

In this paper $\mathcal{C}ol(.)$ denotes a column vector and $\mathcal{D}iag(.)$ denotes a diagonal matrix. The expressions det(.) and rank(.) denote the determinant and rank of a matrix respectively and $\mathcal{N}(\cdot)$ and $\mathcal{R}(\cdot)$ represent the null space and range space of a matrix. The symbol $\otimes$ denotes the Kronecker product.

II. SYSTEM DESCRIPTION

Consider a network of interconnected dynamical systems, represented from the perspective of a graph. Each node represents a dynamical system, and an edge denotes an interaction between two nodes. These interconnections are assumed to be bidirectional and hence the network is considered as a static undirected graph. As in [1], [2] the dynamics of the network

Christopher Edwards and Prathyush P Menon are affiliated to the College of Engineering, Mathematics and Physical Sciences, University of Exeter, EX4 4QF, UK. (email: C.Edwards@exeter.ac.uk, P.M.Prathyush@exeter.ac.uk) The work was supported by a UKIERI research grant.

considered in this paper is given by

$$\dot{x}_i(t) = Ax_i(t) + G\phi(Hx_i(t)) - \alpha(t)\sum_{j=1}^{N}\mathcal{L}_{ij}\Gamma x_j(t) \quad (1)$$

$$y_i(t) = Cx_i(t) \quad (2)$$

for $i = 1, \ldots, N$ where $\mathcal{L}$ represents the Laplacian of the underlying undirected topology. In (1)-(2), $x_i(t) \in \mathbb{R}^n$ and $y_i(t) \in \mathbb{R}^p$ represent the states and measured outputs at the i^{th} node, where $p \leq n$. The matrices $A \in \mathbb{R}^{n\times n}$, and $C \in \mathbb{R}^{p\times n}$ represent a linear system and C is assumed to be full row rank. The matrices $G \in \mathbb{R}^{n\times r}$ and $H \in \mathbb{R}^{r\times n}$ and the nonlinear function $\phi(\cdot) : \mathbb{R}^r \mapsto \mathbb{R}^r$, is assumed to be Lipschitz with respect to x with Lipschitz gain $l_\phi > 0$. The positive scalar $\alpha(t) \in \mathbb{R}^+$ is a time varying coupling strength, which is assumed to satisfy $\inf_t \alpha(t) = \alpha_0 > 0$ almost everywhere. The matrix $\Gamma \in \mathbb{R}^{n\times n}$ describes the state interconnection structure and is comprised of elements which are zero or one [7]. In this paper it is assumed the interconnection structure is fully connected and so rank$(\mathcal{L}) = N-1$. Without loss of generality, the states can be chosen such that the output distribution matrix in (2) has the following structure:

$$C = \begin{bmatrix} 0 & C_2 \end{bmatrix} \quad (3)$$

where $C_2 \in \mathbb{R}^{p\times p}$ and $\det(C_2) \neq 0$.

***Assumption** 1:* The matrix $\Gamma \in \mathbb{R}^{n\times n}$, which represents the node level coupling among the states, in a coordinate system in which C has the structure in (3) has the form

$$\Gamma = \begin{bmatrix} 0 & 0 \\ 0 & I_q \end{bmatrix} \quad \text{where} \quad q \leq p \quad (4)$$

(Assumption 1 implies that coupling is via the node outputs and this impacts on the states of 'neighbouring' nodes.)

***Assumption** 2:* The matrix Γ can be factorized as $\Gamma = DFC$ where $F \in \mathbb{R}^{q\times p}$ and $D \in \mathbb{R}^{n\times q}$ has full column rank, with the geometric property that

$$\mathcal{R}(D) \cap \mathcal{N}(FC) = \{\emptyset\} \quad (5)$$

***Assumption** 3:* The triple (A, D, FC), thought of as representing a linear system, is minimum phase.

***Remark** 1:* The system in (1) can arise from an engineering situation in which N identical subsystems of the form

$$\dot{x}_i(t) = Ax_i(t)+Bu_i(t), \quad y_i = Cx_i(t), \quad \text{for all } i = 1, \ldots, N.$$

have control inputs $u_i(t)$ of a distributed/consensus type:

$$u_i(t) = -\alpha(t)\sum_{j=1}^{N} a_{ij}\mathcal{F}\left(y_i(t) - y_j(t)\right) \quad (6)$$

where $\mathcal{F} \in \mathbb{R}^{m\times p}$ is a feedback gain matrix, and $a_{ij} = 1$ or 0, i.e., the ij^{th} element of the adjacency matrix $\mathcal{A}(\mathcal{G})$. In this situation $\Gamma = B\mathcal{F}C$ where B is the input distribution matrix associated with each node of the network and with a geometric constraint $\mathcal{R}(B) \cap \mathcal{R}(\mathcal{F}C) = \{\emptyset\}$.

***Lemma** 1:* For the system in (1)-(2) satisfying Assumptions 1-3, there exists a coordinate system in which

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad D = \begin{bmatrix} 0 \\ D_2 \end{bmatrix} \quad FC = \begin{bmatrix} 0 & D_2^{-1} \end{bmatrix} \quad (7)$$

where $A_{11} \in \mathbb{R}^{(n-q)\times(n-q)}$ is Hurwitz and $D_2 \in \mathbb{R}^{q\times q}$ is a non-singular matrix.

Proof: The geometric property in Assumption 2 implies $\det(FCD) \neq 0$. To prove this, suppose for a contradiction that $\det(FCD) = 0$. If this is the case there exists a vector $\xi \neq 0$ such that $(FCD)\xi = 0$. This implies that the vector $\tilde{\xi} = D\xi$ belongs both to $\mathcal{R}(D)$ and $\mathcal{N}(FC)$ and so therefore from the constraint in (5), $\tilde{\xi} = 0$. Since D has full column rank, $\tilde{\xi} = D\xi = 0$ implies $\xi = 0$ which is a contradiction. Therefore $\det(FCD) \neq 0$ as claimed. As a consequence of Assumption 1, the matrix D also has a special structure. To expose this write D generically as

$$D = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix}$$

where $D_2 \in \mathbb{R}^{q\times q}$. It follows from (4) that

$$\underbrace{(DFC)}_{\Gamma} D = \Gamma D = \begin{bmatrix} 0 & 0 \\ 0 & I_q \end{bmatrix}\begin{bmatrix} D_1 \\ D_2 \end{bmatrix} = \begin{bmatrix} 0 \\ D_2 \end{bmatrix}$$

Using the associative property of matrices, it follows that

$$D(FCD) = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} FCD = \begin{bmatrix} 0 \\ D_2 \end{bmatrix} \quad (8)$$

Considering the first $n-q$ rows of (8) yields $D_1(FCD) = 0$, which implies $D_1 = 0$ since $\det(FCD) \neq 0$. Since $D \in \mathbb{R}^{n\times q}$ is full column rank and $D_1 = 0$, $\det(D_2) \neq 0$. The last q rows of (8) then implies $FCD = I_q$. Furthermore the triple (A, D, FC), which is minimum phase from Assumption 3, has the property rank$(FCD) = q$. Hence, using the results from [15], there exists a state-space realization in which (A, D, FC) has the canonical structure in (7). Furthermore, as argued in [15], the eigenvalues of A_{11} are the invariant zeros of (A, D, FC) and therefore by assumption A_{11} is Hurwitz. ■

For a network of the form described in (1) - (2), with each subsystem (A, D, FC) satisfying Assumptions 1 and 3, the objective is to estimate all the states using an interconnection of observers (observer network) employing the fewest number of observer gains/output error injection signals. The distributed observer network is assumed to have an identical interconnection topology among the observer subsystems. A schematic of such an arrangement is depicted in Figure 1. The observer subsystems communicate their estimated states to other neighbouring observer subsystem nodes, and use measurements from only a certain subset of the nodes. Imposing the requirement of using a minimum number of injection signals, implies only sensors belong to a subset of the nodes need to 'actively' take measurements. This minimizes the consumption of power and engenders resilience because of the inherent reconfiguration capability this creates.

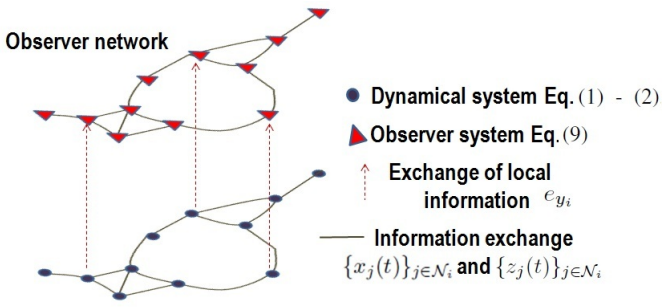

Fig. 1. Schematic of distributed network

III. Observer structure

Let the distributed observer be given by

$$\dot{z}_i(t) = Az_i(t) + G\phi(Hz_i) - \alpha(t)\sum_{j=1}^{N}\mathcal{L}_{ij}\Gamma z_j(t) - L_i(t)e_{y_i} \tag{9}$$

for $i = 1, \ldots, N$ where the state estimate of the i^{th} node is $z_i \in \mathbb{R}^n$ and $e_{y_i} := C(x_i - z_i)$ is the local output state estimation error at the i^{th} node. The gain $L_i(t) \in \mathbb{R}^{n\times p}$ at the i^{th} node is to be determined. Furthermore the objective is to achieve a sparse solution in which as many of the $L_i(t) \equiv 0$ as possible, and so measurements are not required at the associated node. Define the error in the state estimate of the i^{th} node as $e_i := x_i - z_i$ then

$$\begin{aligned}\dot{e}_i(t) &= (A + L_i(t)C)e_i(t) + G\phi(Hx_i) - G\phi(Hz_i)\\ &\quad -\alpha(t)\sum_{j=1}^{N}\mathcal{L}_{ij}\Gamma e_j(t)\ ,\end{aligned} \tag{10}$$

Suppose the observer gain at an individual node has the form

$$L_i(t) = -\alpha(t)\ell_i DF\ , \tag{11}$$

for all $i = 1, \ldots, N$ where the ℓ_i are positive scalar gains to be determined and the matrices D, F from Assumption 2 are considered as fixed and given. A series of lemmas will first be presented before developing the main results.

A. Sufficiency conditions at a single node level

Consider a symmetric positive definite (s.p.d) matrix $P \in \mathbb{R}^{n\times n}$ with a block diagonal form as considered in [14]

$$P = \begin{bmatrix} P_1 & 0 \\ 0 & P_2 \end{bmatrix}\ , \tag{12}$$

where the matrix sub-block $P_2 \in \mathbb{R}^{q\times q}$. Also partition the matrices associated with the nonlinear term in (1) conformably with the canonical form in (7) as

$$G = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix}, \quad H = \begin{bmatrix} H_1 & H_2 \end{bmatrix}\ , \tag{13}$$

where $G_2 \in \mathbb{R}^{q\times r}$ and $H_2 \in \mathbb{R}^{r\times q}$.

Lemma 2: Let γ be a positive scalar which satisfies the inequality $\|H_1(sI - A_{11})^{-1}G_1\|_\infty < \gamma$, then in the coordinates of Lemma 1, there exists a scalar ℓ_0 and a s.p.d matrix $P \in \mathbb{R}^{n\times n}$ of the form given in (12) such that

$$P(A-\ell_0 DFC)+(A-\ell_0 DFC)^{\mathrm{T}}P+\frac{1}{\gamma}\big(PGG^{\mathrm{T}}P+H^TH\big) < 0\ . \tag{14}$$

Proof: First notice that, in the partitioned form consistent with (7) in Lemma 1,

$$\begin{aligned}&P(A - \ell_0 DFC) + (A - \ell_0 DFC)^T P =\\ &\quad\begin{bmatrix} P_1A_{11} + A_{11}^{\mathrm{T}}P_1 & P_1A_{12} + A_{21}^{\mathrm{T}}P_2 \\ P_2A_{21} + A_{12}^{\mathrm{T}}P_1 & P_2A_{22} + A_{22}^{\mathrm{T}}P_2 - 2\ell_0 P_2 \end{bmatrix}.\end{aligned} \tag{15}$$

By applying a Schur complement, (14) is equivalent to

$$\begin{bmatrix} P(A - \ell_0 DFC) + (A - \ell_0 DFC)^T P & PG & H^{\mathrm{T}} \\ G^{\mathrm{T}}P & -\gamma I_r & 0 \\ H & 0 & -\gamma I_r \end{bmatrix} < 0\ . \tag{16}$$

Writing (16) in terms of its block sub-components, by making use of the block partitions in (12), (13) and (15), and re-arranging the order of the resultant columns 1,2,3 and 4 into the order 1,3,4 and 2, the inequality in (16) becomes

$$\left[\begin{array}{ccc|c} P_1A_{11} + A_{11}^{\mathrm{T}}P_1 & P_1G_1 & H_1^{\mathrm{T}} & P_1A_{12} + A_{21}^{\mathrm{T}}P_2 \\ * & -\gamma I_r & 0 & G_2^{\mathrm{T}}P_2 \\ * & * & -\gamma I_r & H_2 \\ \hline * & * & * & P_2A_{22} + A_{22}^{\mathrm{T}}P_2 - 2\ell_0 P_2 \end{array}\right] < 0\ . \tag{17}$$

A necessary condition for (17) to hold is that the top left square sub-block matrix

$$\Psi(P_1) = \begin{bmatrix} P_1A_{11} + A_{11}^{\mathrm{T}}P_1 & P_1G_1 & H_1^{\mathrm{T}} \\ * & -\gamma I_r & 0 \\ * & * & -\gamma I_r \end{bmatrix} < 0\ . \tag{18}$$

From the Bounded real lemma, (18) is equivalent to the condition $\|H_1(sI - A_{11})^{-1}G_1\|_\infty < \gamma$. Consequently from the hypothesis of Lemma 2, there exists a s.p.d matrix P_1 such that $\Psi(P_1) < 0$. Now let $P_2 := I_q$, and define

$$X^{\mathrm{T}} = \begin{bmatrix} A_{12}^{\mathrm{T}}P_1 + A_{21} & G_2 & H_2^{\mathrm{T}} \end{bmatrix}\ .$$

Then from the Schur complement, the matrix inequality (17) holds if and only if

$$\ell_0 I > -X^{\mathrm{T}}\Psi(P_1)X - (A_{22} + A_{22}^{\mathrm{T}}). \tag{19}$$

This condition can always be achieved for large enough $\ell_0 > 0$ and so the lemma is proved. ■

Remark 2: Lemma 2 demonstrated the existence of a gain $\ell_0 > 0$ for which the matrix inequality in (14) is satisfied. It is more useful to find the minimum value of ℓ_0 for which (14) is satisfied. This is equivalent to the problem of finding the smallest $\gamma > 0$ such that (17) holds for some $P_1, P_2 > 0$. For a given $\gamma > 0$, finding a solution to (17) with respect to the decision variables P_1 and $P_2 > 0$ is an LMI feasibility problem and can be tested easily. The minimum value of γ for which a solution can be found reduces to a 'line search' over $\begin{bmatrix} 0 & \ell_0 \end{bmatrix}$, which can be solved using a bisection algorithm.

Lemma 3: Suppose the conditions of Lemma 2 hold, and let ℓ_0^* be the smallest positive value for which the inequality (14) (or equivalently (17)) holds. Then if $\ell(t) : \mathbb{R}^+ \mapsto \mathbb{R}^+$ is a time varying gain satisfying $\ell(t) > \ell_0^*$

$$\begin{aligned}\Theta(t) &:= P(A - \ell(t)DFC) + (A - \ell(t)DFC)^{\mathrm{T}}P\\ &\quad +\tfrac{1}{\gamma}\big(PGG^{\mathrm{T}}P + H^TH\big) < 0\ ,\end{aligned} \tag{20}$$

for all $t \in \mathbb{R}^+$.

Proof: From Assumption 2, $DFC = \Gamma$, and therefore

$$\begin{aligned}\Theta(t) \ &\equiv\ P(A - \ell_0^*\Gamma)+(A - \ell_0^*\Gamma)^{\mathrm{T}}P+\frac{1}{\gamma}\big(PGG^{\mathrm{T}}P + H^{\mathrm{T}}H\big)\\ &\quad -2(\ell(t) - \ell_0^*)(P\Gamma + \Gamma^{\mathrm{T}}P)\ .\end{aligned} \tag{21}$$

From the block partition structures of P and Γ it follows by direct computation that $P\Gamma + \Gamma^{\mathrm{T}}P = \mathcal{D}iag(0, P_2) \geq 0$ and hence, $-2(\ell(t) - \ell_0^*)(P\Gamma + \Gamma^{\mathrm{T}}P) < 0$ if $\ell(t) > \ell_0^*$. Therefore from (21) for all $\ell(t) > \ell_0^*$

$$\Theta(t) \leq P(A - \ell_0^*\Gamma) + (A - \ell_0^*\Gamma)^{\mathrm{T}}P + \frac{1}{\gamma}\big(PGG^{\mathrm{T}}P + H^TH\big)\ . \tag{22}$$

and consequently $\Theta(t) < 0$ for all $t \in \mathbb{R}^+$ if the conditions in Lemma 2 holds. ■

B. Sparse distribution of observer gains at network level

For the observer system in (9) with gains defined in (11) let

$$\mathfrak{L} = \mathcal{D}iag(\ell_1, \ldots, \ell_N) \tag{23}$$

where the $\ell_i \geq 0$ for $i = 1 \ldots N$. Define the network state estimation error $e := \mathcal{C}ol(e_1, \ldots, e_N)$. Then since by construction $DFC = \Gamma$, it follows from (10) that

$$\begin{aligned} \dot{e}(t) &= ((I_N \otimes A) - \alpha(t)(\mathfrak{L} \otimes DFC) - \alpha(t)(\mathcal{L} \otimes \Gamma))\, e(t) \\ &\quad +(I_N \otimes G)\Phi(\cdot) \\ &= ((I_N \otimes A) - \alpha(t)((\mathcal{L} + \mathfrak{L}) \otimes \Gamma))\, e(t) \\ &\quad +(I_N \otimes G)\Phi(\cdot) \end{aligned} \tag{24}$$

where $\Phi(.) = \mathcal{C}ol(\Phi_1(.), \ldots, \Phi_N(.))$ and where the component $\Phi_i(.) = \phi(Hx_i) - \phi(Hz_i)$. The dynamics of the overall network can further conveniently be written as

$$\dot{e}(t) = \big((I_N \otimes A) - \alpha(t)(\tilde{\mathcal{L}} \otimes \Gamma)\big)e(t) + (I_N \otimes G)\Phi(\cdot) \tag{25}$$

where $\tilde{\mathcal{L}} = \mathcal{L} + \mathfrak{L}$. Note that by construction, $\tilde{\mathcal{L}}$ is dependent on the observer gains $\ell_i \geq 0$, for $i = 1 \ldots N$ to be designed.

Proposition 1: Suppose the conditions of Lemma 2 are satisfied, for $\gamma = \frac{1}{l_\phi}$, where l_ϕ is Lipschitz gain for $\phi(.)$, and $\mathfrak{L}$ is chosen so that

$$\mathcal{L} + \mathfrak{L} > \frac{\ell_0^*}{\alpha_0} I_N \ , \tag{26}$$

then $\alpha(t)\lambda_i > \ell_0^*$, where λ_i is any eigenvalue of $\tilde{\mathcal{L}} = \mathcal{L} + \mathfrak{L}$, and hence the error system in (25) is stable.

Proof: Note that

$$\mathcal{L} + \mathfrak{L} > \frac{\ell_0}{\alpha_0} I_N \quad \Rightarrow \lambda_i > \frac{\ell_0}{\alpha_0}, \ \text{ for } i = 1...N \ ,$$

where λ_i is any eigenvalue of $\mathcal{L} + \mathfrak{L}$. (Note all eigenvalues of $\mathcal{L} + \mathfrak{L}$ are real.) Consequently if (26) holds then

$$\alpha(t)\lambda_i > \alpha(t)\frac{\ell_0^*}{\alpha_0} \Rightarrow \alpha(t)\lambda_i > \ell_0^* \ , \tag{27}$$

since by assumption $\inf_t \alpha(t) = \alpha_0 > 0$, for all $t \in \mathbb{R}^+$. By construction $\tilde{\mathcal{L}}$ is symmetric since both $\mathcal{L}$ and $\mathfrak{L}$ are symmetric and so by spectral decomposition

$$\tilde{\mathcal{L}} = V \Lambda V^{\mathrm{T}} \ , \tag{28}$$

where the orthogonal matrix $V \in \mathbb{R}^{N \times N}$ is formed from the eigenvectors of $\tilde{\mathcal{L}}$, and $\Lambda \in \mathbb{R}^{N \times N}$ is a diagonal matrix $\Lambda = \mathcal{D}iag(\lambda_1, \ldots \lambda_N)$ formed from the eigenvalues. Define a co-ordinate transformation $T : e \mapsto \eta := Tx$, where

$$T = (V^{\mathrm{T}} \otimes I_n) \ , \tag{29}$$

and V is the orthogonal matrix from (28). The matrix T is an orthogonal transformation since V is orthogonal. Applying the coordinate transformation given in (29) to (25), after algebraic manipulation

$$\dot{\eta} = ((I_N \otimes A) - \alpha(t)(\Lambda \otimes \Gamma))\, \eta + (V^{\mathrm{T}} \otimes G)\Phi(\cdot) \ . \tag{30}$$

Now consider $\mathcal{V} = \eta^{\mathrm{T}}(I_N \otimes P)\eta = \sum_{i=1}^N \eta_i^{\mathrm{T}} P \eta_i$ as a candidate Lyapunov function for (30) where P has the block diagonal structure in (12). Since Λ is diagonal it follows that

$$\begin{aligned} \dot{\mathcal{V}} &= \sum_{i=1}^N \eta_i^{\mathrm{T}}\big(P(A - \alpha(t)\lambda_i\Gamma) + (A - \alpha(t)\lambda_i\Gamma)^{\mathrm{T}}P\big)\eta_i \\ &\quad +2\eta^{\mathrm{T}}(V^{\mathrm{T}} \otimes PG)\Phi(\cdot) \ . \end{aligned} \tag{31}$$

Since $(V^{\mathrm{T}} \otimes PG) = (I_N \otimes PG)(V^{\mathrm{T}} \otimes I_n)$ and $\alpha(t)\lambda_i > \ell_0^*$ for all $i = 1 \ldots N$, it follows from Lemma 3 that P can be chosen so that

$$\begin{aligned} \dot{\mathcal{V}} &< \sum_{i=1}^N -\frac{1}{\gamma}\eta_i^{\mathrm{T}}\big(PGG^{\mathrm{T}}P + H^{\mathrm{T}}H\big)\eta_i \\ &\quad +2\eta^{\mathrm{T}}(I_N \otimes PG)(V^{\mathrm{T}} \otimes I_n)\Phi(\cdot) \ . \end{aligned} \tag{32}$$

Note that $\|(V^{\mathrm{T}} \otimes I_n)\Phi(\cdot)\|^2 = \|\Phi(\cdot)\|^2$ since $(V^{\mathrm{T}} \otimes I_n)$ is orthogonal, and

$$\|\Phi(\cdot)\|^2 = \sum_{i=1}^N \|\Phi_i(\cdot)\|^2 \leq l_\phi^2 \sum_{i=1}^N \|He_i\|^2 \ , \tag{33}$$

based on the Lipschitz properties of $\Phi_i(\cdot)$ inherited from the Lipschitz properties of the function $\phi(\cdot)$ in (1). Furthermore

$$l_\phi^2 \sum_{i=1}^N \|He_i\|^2 = l_\phi^2 \|(I_N \otimes H)e\|^2 = l_\phi^2 \|(V^{\mathrm{T}} \otimes I_n)(I_N \otimes H)e\|^2, \tag{34}$$

because $(V^{\mathrm{T}} \otimes I_n)$ is orthogonal. From the property of the Kronecker product operation and the definition of the coordinate transformation $T : e \mapsto \eta := Tx$

$$\begin{aligned} l_\phi^2 \|(V^{\mathrm{T}} \otimes I_n)(I_N \otimes H)e\|^2 &= l_\phi^2 \|(I_N \otimes H)(V^{\mathrm{T}} \otimes I_n)e\|^2 \\ &= l_\phi^2 \|(I_N \otimes H)\eta\|^2 \ . \end{aligned} \tag{35}$$

By hypothesis $l_\phi = \frac{1}{\gamma}$, and hence

$$l_\phi^2 \|(I_N \otimes H)\eta\|^2 = l_\phi^2 \sum_{i=1}^N \|H\eta_i\|^2 = \frac{1}{\gamma^2} \sum_{i=1}^N \|H\eta_i\|^2 \ .$$

Therefore from (33)-(35) for all x_i, z_i and for $i = 1, \ldots, N$

$$\|\Phi(\cdot)\|^2 \leq \frac{1}{\gamma^2} \sum_{i=1}^N \|H\eta_i\|^2 \ . \tag{36}$$

From Young's inequality [17]

$$\begin{aligned} 2\eta^{\mathrm{T}}(I_N \otimes PG)(V^{\mathrm{T}} \otimes I_n)\Phi(\cdot) &\leq \frac{1}{\gamma}\sum_{i=1}^N \eta_i^{\mathrm{T}} PGG^{\mathrm{T}} P\eta_i + \gamma\|\Phi(\cdot)\|^2, \\ &\leq \frac{1}{\gamma}\sum_{i=1}^N \eta_i^{\mathrm{T}} PGG^{\mathrm{T}} P\eta_i + \|H\eta_i\|^2, \end{aligned} \tag{37}$$

and therefore from (32) and (37), $\dot{\mathcal{V}} < 0$ for all $\eta_i \neq 0$ for $i = 1, \ldots, N$ and asymptotic stability is guaranteed. ■

C. LMI formulation

The problem of designing the observer gains employing output error injection feedback at a minimum number of nodes will now be formulated as a two stage optimization problem.

In the first step, the convex optimization problem is solved to obtain a value for ℓ_0^* according to Remark 2. Subsequently, in the second stage of the optimization problem, the diagonal matrix of observer gains $\mathfrak{L}$ from (23) is computed. A natural cost function is trace($\mathfrak{L}$). Since $\mathfrak{L}$ is diagonal with non-negative diagonal elements, this quantity represents $\sum_{i=1}^N |\ell_i|$ and so this penalizes the total norm of the injection gains ℓ_i. However, here the requirement is to have as many as possible of the diagonal entries of the matrix $\mathfrak{L}$ equal to zero, which in turn, from (11), makes many of the $\ell_i = 0$. Making many $\ell_i = 0$ essentially imposes a rank constraint on the matrix $\mathfrak{L}$. Optimization problems involving rank constraints are usually

non-convex, and in general known to be computationally intractable (NP-hard) [16], [18]. To obviate this difficulty a known heuristic involving the $\log\det(\cdot)$ function, that solves an approximation to this problem, will be employed [19]. Specifically:

Minimize $\log\det(\mathfrak{L})$+Trace($\mathfrak{L}$) w.r.t $\ell_1, \ldots \ell_N$, subject to:

$$\mathcal{L}+\mathfrak{L} \;>\; \frac{\ell_0}{\alpha_0} I_N \;, \tag{38}$$

$$\mathfrak{L} \;\geq\; 0 \;. \tag{39}$$

Because of the specific diagonal structure of $\mathfrak{L}$, the function $\log\det(\mathfrak{L}) = \sum_{i=1}^{N}\log(\ell_i)$. This problem is still not convex and so an iterative solution is employed [18]. The function $\sum_{i=1}^{N}\log(\ell_i)$ is then linearized so that at the kth iteration, for $\mathfrak{L}^k = \mathcal{D}iag(\ell_1^k, \ell_2^k, \ldots \ell_N^k)$, from a Taylor's series expansion,

$$\log\det(\mathfrak{L}) \approx \log\det(\mathfrak{L}^k) + \sum_{i=1}^{N}\frac{1}{\ell_i^k}(\ell_i - \ell_i^k) \;. \tag{40}$$

This optimization problem is solved as given in Algorithm 1.

ALGORITHM 1

1) Initialize $M^k \in \mathbb{R}^{1\times N}$, where $M^k = (m_1^k, \ldots, m_N^k)$ and $m_i^k \in \mathbb{R}^+$ as $m_i^1 = 1$ for $i = 1 \ldots N$,
2) Define $k = 1$,$\epsilon > 0$,$J[0] = 0$, and the improvement in cost function value $\Delta J[k] = 0$, accuracy (δ) and maximum iterations ($Iter_{max}$);
3) **While** $(\Delta J[k] \leq 0)$ & $(k < Iter_{max})$ **do**
{
   - Increment the iteration counter, $k = k + 1$
   - Minimize w.r.t. $(\ell_1, \ldots, \ell_N)$

$$J[k] := \sum_{i=1}^{N}\frac{\ell_i}{m_i^k} + \text{Trace}(\mathfrak{L})$$

   - Subject to:

$$\mathcal{L}+\mathfrak{L} > \frac{\ell_0}{\alpha_0} I_N$$

   and

$$\mathfrak{L} > \epsilon I_N$$

   - Update the improvement in cost function value:

$$\Delta J[k] := J[k] - J[k-1]$$

   - Update M^k element-wise as $m_i^{k+1} = \ell_i$

}
4) **End**

In this problem there is a performance trade-off in terms of the magnitude of the gains. As the solutions become more sparse, the nonzero elements tend to increase in magnitude. To justify this comment, notice that if inequality (38) holds, multiplying on the left by a row vector of ones and on the right by a column vector of ones, implies

$$\textstyle\sum_{i=1}^{N} l_i > N l_0/\alpha_0 \tag{41}$$

must hold. (In (41) $\mathcal{L}$ vanishes because of its row-sum-equal-to-zero property.) From (41) the effect of sparseness and the increasing magnitude of the nonzero terms is apparent.

***Remark** 3:* Algorithm 1 is guaranteed to find a locally optimal solution since $\mathfrak{L} = \frac{\ell_0}{\alpha_0} I_N$ is a feasible solution. However the algorithm is not guaranteed to find the global optimal solution since the problem as posed is not convex.

***Remark** 4:* The formulation in this paper is quite flexible and can be extended to include further constraints (in some cases at the cost of no longer being able to guarantee a feasible solution). Variations could include:

- Ensuring that each $l_i \in [\ 0 \quad \hat{l}\ ]$, i.e. belongs to a specified finite interval of the real line.
- Guaranteeing certain sensors are never used: setting the corresponding $l_i = 0$ enforces this constraint.
- Preferentially ensuring certain nodes are used by removing the associated index from the $\log(\det(\mathfrak{L}))$ heuristic in (40) and instead employing the expression

$$\textstyle\sum_{i\in\bar{\mathcal{S}}} \frac{1}{\ell_i^k}(\ell_i - \ell_i^k) \tag{42}$$

  where $\bar{\mathcal{S}} = \{1, 2, \ldots, N\}/\mathcal{S}$ and $\mathcal{S}$ is the set of indices representing the nodes required to be used. In this way the nodes represented $\mathcal{S}$ are not 'encouraged' to be zero and will be used in preference to those in $\bar{\mathcal{S}}$.

All these modifications retain convexity, but a feasible solution cannot always be guaranteed.

IV. Simulation Example

Here a network of interconnected flexible link robot systems is considered [12], [13]. Each system is modelled as:

$$\begin{bmatrix} \dot{\theta}_m^i \\ \dot{\omega}_m^i \\ \dot{\theta}_l^i \\ \dot{\omega}_l^i \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\frac{K_l}{J_m} & -\frac{B}{J_m} & \frac{K_l}{J_m} & 0 \\ 0 & 0 & 0 & 1 \\ \frac{K_l}{J_l} & 0 & a_{43} & 0 \end{bmatrix} \begin{bmatrix} \theta_m^i \\ \omega_m^i \\ \theta_l^i \\ \omega_l^i \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ \frac{1}{J_m} & 0 \\ 0 & 0 \\ \frac{-1}{J_1} & \frac{mgh}{J_1} \end{bmatrix} \begin{bmatrix} \kappa_2\left(\theta_m^i\right)^3 \\ \omega_m^i - \sin(\omega_m^i) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} u^i \tag{43}$$

where $a_{43} = -\frac{K_l}{J_l} - \frac{mgh}{J_l}$ and the states x_i are given by the angular position and velocity of the motor shaft (θ_m^i and ω_m^i) and the angular position and velocity of the link (θ_l^i and ω_l^i). For a detailed explanation of the physical parameters and their interpretation see [13]. Here a consensus controller

$$u_i = \sum_{j\in\mathcal{J}_i} FC(x_i - x_j) \quad \text{for } i = 1 \ldots N \tag{44}$$

has been introduced where the output distribution matrix is

$$C = [\ I_3 \quad 0\ ] \tag{45}$$

In the example which follows the consensus gains are

$$F = [\ 6 \quad 1 \quad -10\ ] \tag{46}$$

With parameters taken from [12], [13] the system matrix

$$A = \begin{bmatrix} 0 & 1.0000 & 0 & 0 \\ -48.6486 & -12.4324 & 48.6486 & 0 \\ 0 & 0 & 0 & 1.0000 \\ 19.7802 & 0 & -53.7379 & 0 \end{bmatrix}$$

and

$$D^{\mathrm{T}} = [\ 0 \quad 1 \quad 0 \quad 0\ ]$$

It can be easily verified $\mathcal{R}(D) \cap \mathcal{N}(FC) = \{0\}$ and that the transmission zeros of (A, FC, D) are

$$\{-1.6086 \pm 6.4960i, -2.7827\}$$

and so the theory developed earlier in the paper is applicable. In the canonical form from (7)

$$A = \left[\begin{array}{ccc|c} -6.0273 & -0.2923 & 10.0000 & -6.0565 \\ 0.5621 & 0.0273 & 1.0000 & 0.5648 \\ -4.2831 & -10.9068 & 0 & 0 \\ \hline 0.1025 & -16.7010 & -10.0000 & -6.4324 \end{array}\right]$$

and

$$D^{\mathrm{T}} = (FC) = [\ 0 \quad 0 \quad 0 \quad 1\ ] \tag{47}$$

The method described earlier in the paper ensures stability, but not necessarily any performance. To introduce performance, the system matrix A is replaced by $A + dI_n$ where the scalar $d > 0$. Performing the design on the plant $(A + dI_n, D, FC)$ will ensure all the poles of the error system (for frozen time) will be to the left of $-d$ in the Complex plane.

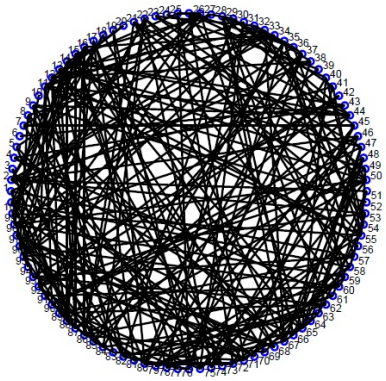

Fig. 2. Connection topology

A random network with 100 nodes has been generated. The interconnection structure is shown in Figure 2. When $d = 0.45$ and $\gamma = 50$ it can be verified that the LMI in (17) is feasible when $\ell_0^* = 2.405$. This value has been obtained using a bisection approach based on the original interval $[\ 0 \quad \ell_0\ ]$ where $\ell_0 = 7.286 \times 10^6$ has been calculated from (19). Using $\ell_0^* = 2.405$ and $\alpha_0 = 5$ a sparse solution has been obtained using Algorithm 1 in which 88 out of 100 of the diagonal elements of $\mathfrak{L}$ are zero. The nodes which need to be instrumented are given below.

TABLE I
NODE IDENTIFIER AND GAIN VALUES

Node	Gain l_i	Node	Gain l_i
2	24.5303	15	18.4830
4	2.5007	22	18.3550
11	34.8853	24	0.6965
12	14.0205	37	11.3347
13	20.9083	57	3.5074
14	35.2417	67	27.1469

If a centralized approach, or a conventional full node approach is utilized, output measurements from all nodes of the network, i.e., 100 measurements, would be necessary. In the proposed approach, only 12 measurements are required. Figure 3 shows the state estimation errors in 10 representative nodes from the network as a whole, together with the overall mean squared estimation error for the complete network.

V. CONCLUSIONS

In this paper a distributed observer has been proposed to estimate the states in a large scale network of linear time invariant systems interconnected by a positive, time varying coupling strength. The distributed observer comprises distinct sub-observers driven by only local node level information, but they exchange their local state estimates with their neighbouring observers. The key idea here is to use a minimum number of, or at least relatively few, measurements from the network being monitored, to reduce the sensor requirements. The problem has been formulated as a two stage optimization problem. Currently the proposed scheme assumes the global time varying coupling strength is measurable for all the time – although this could be overcome in the future with the use of an appropriate adaptive scheme.

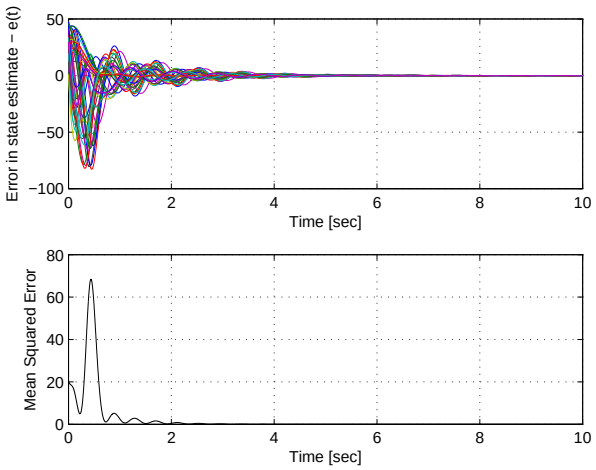

Fig. 3. Simulation Results

REFERENCES

[1] C. W. Wu, "*Synchronisation of complex networks of nonlinear dynamical systems*", World Scientific Publishing, 2007.

[2] X.F. Wang and G. Chen, "Pinning control of scale-free network", Physica A, vol. 310, pp. 521-531, 2002.

[3] I. Shames, A. M. H. Teixeira, H. Sandberg, K. H. Johansson," Distributed fault detection for interconnected second-order systems". *Automatica*, 47, pp. 2757-2764, 2011.

[4] R. Olfati-Saber, "Distributed Kalman filtering for sensor networks," *In Proceedings of* 46^{th} *IEEE Conference of Decision and Control*, 2007.

[5] R. Olfati-Saber and R. M. Murray, "Consensus problems in networks of agents with switching topology and time delays," *IEEE Transactions on Automatic Control*, vol. 49, pp. 1520-1533, 2004.

[6] V. Ugrinovskii, "Distributed robust filtering with H_∞ consensus of estimates," *Automatica*, vol. 47, pp. 1–13, 2011.

[7] W. Yu, G. Chen, and J. Lu, "On pinning synchronization of complex dynamical networks," *Automatica*, 45, pp. 429-435, 2009.

[8] B. Acikmese, M. Mandic, and J.L. Speyer, "Decentralized observers with consensus filters for distributed discrete-time linear systems", *Automatica*, 50, pp 1037-1052, 2014.

[9] W. Yu, G. Wen, J. Lu and X. Yu, "Pinning observability on complex networks", *IET Control Theory & Applications*, 8, pp. 2136–2144, 2014.

[10] S. Sarama and N. Dutt, "Minimal Sparse Observability of Complex Networks: Application to MPSoC Sensor Placement and Run-Time Themral Estimaion and Tracking" *Design, Automation and Test in Europe Conference and Exhibition (DATE)*, Dresden, 2014.

[11] C. Edwards and P.P. Menon, "Sparse distributed observers for a network of dynamical systems", *Proceedings of 52nd IEEE Conference on Decision and Control*, 2013.

[12] S. Raghavan and J. K. Hedrick. "Observer design for a class of nonlinear systems," *International Journal of Control*, 59:515–528, 1994.

[13] X. Fan and M. Arcack. "Observer design for systems with multivariable monotone nonlinearities," *System & Control Letters*, 50:319–330, 2003.

[14] G. Gu. "Stabilizability conditions of multivariable uncertain systems via output feedback control", *IEEE Transactions on Automatic Control*, 35, pp. 925–927, 1990.

[15] C Edwards, XG Yan and SK Spurgeon. On the solvability of the constrained Lyapunov problem, IEEE Transactions Automatic Control, 52, pp. 1982-1987, 2007.

[16] S. Boyd and L. Vandenberghe. "*Convex Optimization,*" Cambridge University Press, 2004.

[17] H. Khalil, "*Nonlinear Systems,*", Pearson, 3rd Edition, 2008.

[18] M. Fazel, H. Hindi and S. Boyd. "Rank minimization and applications in system theory," In *Proceedings of American Control Conference*, 2004.

[19] M. Fazel, H. Hindi and S.P. Boyd, "Log-det heuristic for matrix rank minimization with applications to Hankel and Euclidean distance matrices", American Control Conference, 2003.

[20] D. Zelazo and M. Mesbahi. "On the observability properties of homogeneous and heterogeneous networked dynamic systems." Proceedings of the IEEE Conference on Decision and Control, 2008.

[21] Y. Liu, J.-J. Slotine and A Barabsi, "Controllability of complex networks", *Nature*, 473, pp 167-173, 2011.

[22] P. Santi, "Topology control in wireless ad hoc and sensor networks." *ACM computing surveys (CSUR) 37*, pp 164-194, 2005.

[23] L. Buttyn and J.P. Hubaux, "Stimulating cooperation in self-organizing mobile ad hoc networks". *Mobile Networks and Applications*, 8, pp 579-592, 2003.