A moving horizon scheme for distributed state estimation

Marcello Farina, Giancarlo Ferrari-Trecate and Riccardo Scattolini

Abstract—This paper presents a novel distributed estimation algorithm based on the concept of moving horizon estimation. Under weak observability conditions we prove convergence of the state estimates computed by any sensor to the correct state even when constraints on noise are taken into account in the estimation process. Simulation examples are provided in order to show the main features of the proposed method.

I. INTRODUCTION

A sensor network consists of a set of electronic devices, with sensing and computational capabilities, which coordinate their activity through a communication network. They can be employed in wide range of applications, such as monitoring, exploration, surveillance or to track targets over specific regions. The diffusion of sensor networks is partly due to the recent developments in wireless communications and to the availability of low cost devices. On the other hand, many theoretical and technological challenges have still to be tackled in order to fully exploit their potentialities. Among the open problems, the use of sensor networks for distributed state estimation is of paramount importance. The problem can be described as follows. Assume that any sensor of the network measures some variables, computes a local estimate of the overall state of the system under monitoring, and transmits to its neighbors both the measured values and the computed state estimation. Then, the main challenge is to provide a methodology which guarantees that all the sensor asymptotically reach a common reliable estimate of the state variables, i.e. the local estimates reach a consensus. This goal must be achieved even if the measurements performed by any sensor are not sufficient to guarantee observability of the process state (namely, local observability), provided that all the sensors, if put together, guarantee such property (namely, collective observability). The transmission of measurements and of estimates among the sensors must lead to the twofold advantage of enhancing the property of observability of the sensors and of reducing the uncertainty of state estimates computed by each node.

Consensus algorithms for distributed state estimation based on Kalman filters have been recently described in [1], [3], [7], [8], [9], [10], [13]. In particular, in [8], [10], [13], consensus on the measurements is used to reduce their uncertainty and Kalman filters are applied by each agent. In [9], three algorithms for distributed filtering are proposed. The first algorithm is similar to the one described in [8], save for the fact that sensors exploit only partial measurements of the state vector. The second approach relies on communicating the state estimates among neighboring agents (consensus on the state estimates). The third algorithm, named iterative Kalman consensus filter, is based on the discrete-time version of a continuous-time Kalman filter plus a consensus step on the state estimates, which is proved to be stable. However, stability has not been proved for the discrete-time version of the algorithm and optimality of the estimates has not been addressed. Recently, convergence in mean of the local state estimates obtained with the algorithm presented in [8] has been proved in [7], provided that the observed process is stable.

In [1] consensus on the estimates is used together with Kalman filters. The weights of the sensors’ estimates in the consensus step and the Kalman gain are optimized (in two subsequent steps) to minimize the estimation error covariance: however, in so doing optimality is not guaranteed. A two-step procedure is also used in [3], where the considered observed signal is a random walk. In the proposed algorithm filtering and consensus are performed subsequently, and the estimation error is minimized with respect to both the observer gain and the consensus weights. This guarantees optimality of the solution. Recently, an interesting solution to the problem of distributed estimation of a parameter vector, with noisy linear measurements, has been proposed in [2]. The algorithm accounts for dynamically changing interconnections among sensors, unreliable communication links, and faults. Asymptotic convergence of the estimates to the true values has been proved, under suitable hypothesis of “dynamical” graph connectivity.

In this paper we propose a distributed algorithm based on the concept of Moving Horizon Estimation (MHE), [11], [12]. This approach has many advantages; first of all, the observer is optimal in a sense, since a suitable minimization problem must be solved on-line at each time instant. Furthermore, we prove that, under weak observability conditions, convergence of the state estimate is guaranteed in a deterministic framework. Finally, constraints on the noise are taken into account, as it is common in receding horizon approaches in control and estimation [6].

The paper is structured as follows. In Section II we introduce the dynamical system to be observed and the structure of the sensor network. We define a number of observability properties of this network and we describe the distributed state estimation algorithm. In Section III we investigate the convergence properties of the algorithm. In Section IV we present a simulation example, while Section V reports some concluding remarks and hints for future developments. For space limitations, all the results are reported without proofs.
which can be found in [5].

II. PROBLEM FORMULATION AND SOLUTION

The observed process is described by the linear discrete-time dynamics

\[ x_{t+1} = Ax_t + w_t, \]

where \( x_t \in \mathbb{R}^n \) is the state vector and \( w_t \in \mathbb{W} \subseteq \mathbb{R}^r \) (\( 0 \in \mathbb{W} \)) represents a disturbance with variance \( Q = \mathbb{W} \in \mathbb{R}^{r \times r} \). We assume that \( y_t \in \mathbb{C} \), where the set \( \mathbb{C} \) is convex. The initial condition \( x_0 \in \mathbb{R}^n \) is a Gaussian random variable with mean \( \mu \) and covariance matrix \( \Sigma_0 \).

In other words, we assume sensors \( \mathcal{E} \) models that sensor \( j \) communicates only once within a sampling interval. In the following, we assume the matrix \( \bar{C} \in \mathbb{R}^{r \times p_r} \) and the matrix \( C_i \) is non null. The communication network among sensors is described by the directed graph \( \mathcal{G} = (\mathcal{V}, \mathcal{E}) \), where the nodes in \( \mathcal{V} \) represent the sensors and the edge \( (j, i) \) in the set \( \mathcal{E} \subseteq \mathcal{V} \times \mathcal{V} \) models that sensor \( j \) can transmit information to sensor \( i \). We assume \( (i, j) \in \mathcal{E}, \forall i \in \mathcal{V} \). Moreover, we denote by \( \mathcal{V}_i \) the set of neighbors to node \( i \), i.e., \( \mathcal{V}_i = \{ j : (j, i) \in \mathcal{E} \} \). We associate to the graph the stochastic matrix \( K \in \mathbb{R}^{M \times M} \) with entries

\[
\begin{cases}
  k_{i,j} = 1 & \text{if } (j, i) \in \mathcal{E} \\
  k_{i,j} = 0 & \text{otherwise} \quad \forall i = 1, ..., M
\end{cases}
\]

In the following, we assume the matrix \( K \) is given, i.e., it will not be considered as a design parameter for state estimation.

A. Local, regional and collective models

We assume that, at a generic time instant \( t \), a given sensor \( i \) can collect the measurement produced by itself and by its neighboring sensors. In other words, we assume sensors communicate only once within a sampling interval. We can now distinguish three types of quantities: local, regional, and collective. With reference to sensor \( i \), a quantity is referred to as local when it is related to the node \( i \) solely, while it is called regional if it is related to the nodes in \( \mathcal{V}_i \). Finally, we say that a quantity is collective, if it is related to the whole network. For the sake of clarity, we use different notations for local, regional and collective variables. Namely, given a variable \( z, z' \) represents its local version, \( z \) is its regional counterpart, and \( z \) the collective version. For instance, we refer to \( y_i \) in (2) as local measurement. On the other hand, if \( \mathcal{V}_i = \{ j_1, ..., j_n \}, \) the regional measurement of node \( i \) is

\[ \tilde{y}_i = \bar{C}^i x_t + v_i, \]

where \( \tilde{y}_i = \left[ (\hat{y}_i^1)^T \ldots (\hat{y}_i^m)^T \right]^T \) and \( \bar{C}^i = \left[ (C_i^1)^T \ldots (C_i^m)^T \right]^T \). The dimension of vectors \( \tilde{y}_i \) and \( v_i \), and the number of rows of matrix \( \bar{C}^i \) is \( \bar{p}_i = \sum_{k=1}^{\mathcal{V}_i} p_{k.i} \).

Furthermore, we denote by \( \bar{R} \in \mathbb{R}^{r \times p_r} \), the covariance matrix related to the regional noise \( v_i \) on sensor \( i \), i.e.,

\[ \bar{R}_i = \text{diag}(R_{j_1}^i, \ldots, R_{j_n}^i). \]

According to the adopted terminology, three different observability notions can be introduced in this framework, namely local, regional and collective observability.

Definition 1: The system is locally observable by sensor \( i \) if the pair \((A, C)\) is observable. The system is regionally observable by sensor \( i \) if the pair \((A, \bar{C}^i)\) is observable. The system is collectively observable if the pair \((A, \bar{C})\) is observable, where \( \bar{C} = [C_1^T \ldots C_m^T]^T \).

Notice that, for a given sensor \( i \), local observability implies regional observability, and regional observability of any sensor implies collective observability, while all opposite implications are not true.

Given a single sensor model (1)-(2), the \( i \)-th sensor regional observability matrix \( \bar{R}_i \) is

\[ \bar{R}_i = [(C_i^1)^T \ldots (C_i^m)^T]^T \]

Let \( \bar{P}_O \) be the orthogonal projection matrix on \( \ker(\bar{R}_i) \), that is the regionally unobservable subspace. Similarly, let \( \bar{P}_O \) be the orthogonal projection on the regional observability subspace \( \ker(\bar{R}_i) \). Next, we recall how \( \bar{P}_O \) and \( \bar{P}_NO \) can be computed. Let \( r_i = \text{rank}(\bar{R}_i) \) and denote with \( \xi_{i1}, \ldots, \xi_{in} \) an orthonormal basis of \( \ker(\bar{P}_O) \). Let also \( \xi_{i1}, \ldots, \xi_{in} \) be an orthonormal basis of \( \ker(\bar{P}_O) \) and define the non-singular matrix

\[ T = [\xi_{i1}, \ldots, \xi_{in}] \]

Defining the matrices \( S_O \) and \( \bar{S}_O \) as

\[ S_O = \begin{bmatrix} I_{r_i} & 0_{(n-r_i) \times n} \end{bmatrix}, \quad \bar{S}_O = \begin{bmatrix} 0_{r_i \times (n-r_i)} & I_{n-r_i} \end{bmatrix} \]

we have \( \bar{P}_O = T_s O S_O T (T_s O S_O)^{-1} \) and \( \bar{P}_NO = T_s O S_O T (T_s O S_O)^{-1} \). Furthermore, defining \( T = \text{diag}(T_1, \ldots, T_m), \quad S_O = \text{diag}(S_O, \ldots, S_O), \) and \( S_NO = \text{diag}(S_NO, \ldots, S_NO) \), the collective projection matrices are \( P_O = S_O S_O T^{-1} \) and \( P_NO = S_0 S_0 T^{-1} \).

Note that \( S_NO \) is empty when the system is regionally observable by sensor \( i \). In this case we assume that \( \bar{P}_NO = 0_{nxn} \).

B. The distributed estimation algorithm

Our aim is to design, for a generic sensor \( i \in \mathcal{V} \), an algorithm for computing a reliable estimate of the system’s state based on regional measurements \( \bar{y}_i \) and further pieces of information provided by sensors \( j \in \mathcal{V}_i \). The proposed solution relies on the use of MHE, see [6], [11], [12], in view of its capability to handle noise constraints. More specifically, we propose a Distributed MHE (DMHE) scheme where each sensor solves a MHE problem. For a given estimation horizon \( N \geq 1 \), each node \( i \in \mathcal{V} \) at time \( t \) solves the constrained minimization problem MHE-i defined as

\[ \Theta_{t+1}^i = \min_{\hat{y}_i^{t+1}, \hat{y}_i^{t+1}, \hat{y}_i^{t+1}, \Gamma_{t+1}} \| J^i(t-N,t,N -N,\hat{y}_i^t, \hat{y}_i^t, \Gamma_{t-N}) \]

under the constraints

\[ \hat{y}_i^{t+1} = \hat{A}_i \hat{y}_i^t + w_i^t \]

\[ \hat{y}_i^{t+1} = \bar{C} \hat{y}_i^t + \bar{v}_i \]

\[ \bar{w}_k^t \in \mathbb{W} \]
\[ J(t) = J(t-N,t)\frac{\delta^2_j(t)}{i} + J(t-N,N-1)\frac{\delta^2_j(t)}{i}, \]

In (8) and hereafter, the notation \[ ||S||_F \] stands for \[ S^T S \], where \( S \) is a positive-semidefinite matrix.

We denote by \( \delta_{x_i}^j(i) \) and \( \delta_{\tilde{x}_i}^j(i) \) the optimizers to (6) and with \( \delta_{x_i}^j(i) \) and \( \delta_{\tilde{x}_i}^j(i) \) the local state sequence stemming from \( x_i^j(i) \) and \( \tilde{x}_i(i) \). Furthermore, \( \delta_{x_i}^j(i) \) denotes the weighted average of state estimates produced by sensors \( j \in \mathcal{N}' \), i.e.

\[ \delta_{x_i}^j(i) = \sum_{i=1}^{M} k_{ij} \delta_{x_i}^j(i-1) \]

In (8), the function \( \Gamma_{t-N}(\delta_{x_i}^j(i),\delta_{\tilde{x}_i}^j(i)) \) is the so-called initial penalty, defined as follows

\[ \Gamma_{t-N}(\delta_{x_i}^j(i),\delta_{\tilde{x}_i}^j(i)) = \Gamma_{NO}(\delta_{x_i}^j(i),\delta_{\tilde{x}_i}^j(i)) + \Gamma_{OJ}(\delta_{x_i}^j(i),\delta_{\tilde{x}_i}^j(i)) \]

where

\[ \Gamma_{NO}(\delta_{x_i}^j(i),\delta_{\tilde{x}_i}^j(i)) = \frac{1}{2} \left\| \delta_{x_i}^j(i)-\delta_{\tilde{x}_i}^j(i) \right\|_F^2 \]

and the constant term \( \Theta_{c-1}^j(i) \) is defined in (6) and it is known at time \( t \). For this reason, it could be neglected when solving the optimization problem. However, since it plays a major role in establishing the main convergence properties of the proposed DMHE, it is here maintained for clarity of presentation.

Note that \( \delta_{x_i}^j(i) \) is the estimate of \( x_{i-N} \) computed by sensor \( j \) at time \( t \) and \( t-1 \) and therefore, in view of the definition of \( k_{ij} \) in (3), \( \Gamma_i(\cdot) \) depends only upon regional quantities. Since also the cost (8) and the constraints (7) depend only upon regional variables, the overall estimation scheme is decentralized.

Finally, notice that \( \Gamma_i(\cdot) \) embodies a consensus term, in the sense that it penalizes deviations of \( \delta_{x_i}^j(i-1) \) from the local average of the state estimates produced by the neighbors to sensor \( i \). Consensus, besides increasing accuracy of the local estimates, is fundamental to guarantee convergence of the state estimates to the state of the system even if regional observability does not hold. In fact, it allows sensor \( i \) to reconstruct components of the system state that cannot be estimated by the \( i \)-th regional model.

The positive definite symmetric matrix \( \Pi_{i-N/i-1}^i \) appearing in (10) plays the role of a covariance matrix and is a design parameter whose choice will be discussed in details in the next section.

---

C. Computation of the matrices \( \Pi_{i-N/i-1}^i \) and estimation procedure

Let us define

\[ \Pi_{i/t} = \text{diag}\left(\Pi_{i/t}^1, \ldots, \Pi_{i/t}^M\right) \]

We require the matrix \( \Pi_{i-N/i-1} \) to satisfy the following Linear Matrix Inequality (LMI)

\[ K^T \left( P_{i-O}^i \Pi_{i-N/i-1}^i \Pi_{i-N/i-1}^i \right) K \leq \Pi_{i-N/i-1}^{-1} \]

where \( K = K \otimes I_0 \in \mathbb{R}^{MN \times nM} \), the symbol \( \otimes \) denotes the Kronecker product, \( I_0 \) is the \( n \times n \) identity matrix and \( \hat{\Pi}_{i-N/i-1} = \text{diag}\left(\hat{\Pi}_{i-N/i-1}^1, \ldots, \hat{\Pi}_{i-N/i-1}^M\right) \). The matrix \( \hat{\Pi}_{i-N/i-1} \) is, for given one iteration of the difference Riccati equation associated to the Kalman filter for the system

\[ \begin{cases} x_{i-N} = A_{i-N} -1 + w_{i-N-1} \\ \hat{x}_{i-N} = \hat{\Sigma}_N x_{i-N} + \hat{\Sigma}_N \end{cases} \]

where matrix \( \hat{\Sigma}_N \) is defined in (5) and

\[ \hat{\Sigma}_N = \begin{bmatrix} 0 & 0 & \ldots & 0 \\ \hat{C} \hat{C} & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \hat{C} \hat{A}^{N-2} & \hat{C} \hat{A}^{N-3} & \ldots & \hat{C} \end{bmatrix} \in \mathbb{R}^{N \times n(N-1)} \]

for all nodes store the matrix \( \Pi_{i} \) and the estimate \( \tilde{x}_{0/0} = \mu \) of \( x_0 \), where \( \mu \) is given.

if \( 1 \leq t \leq N \), the estimation horizon \( N \) is reduced to \( \hat{N} = t \) and node \( i \in \mathcal{N}' \) performs the following steps

- compute \( \hat{\Pi}_{i-N/i-1} = \hat{\Pi}_{i/0}^i \) from \( \Pi_0 \) according to (12), for all \( i \in \mathcal{N}' \),
- solve the problem \( \text{MHE-}i \), with initial penalty

\[ \pi_{i-N}^i = \frac{1}{2} \left\| P_{i-O}^i (\hat{x}_{0/0} - \hat{\Sigma}_N (\hat{x}_{0/0} - \hat{\Sigma}_N)^T) \right\|_F^2 + \frac{1}{2} \left\| P_{i-O}^i \hat{x}_{0/0} \right\|_F^2 \]
These limitations are severe since, for the matrix \(LMI\) hence hampering the application of DMHE to large networks. The next proposition provides a way to circumvent this problem.

**Proposition 1:** The matrices \(\Pi_{i-N/N-i}^j\) which satisfy, \(\forall i \in \mathcal{Y}'\)

\[
\left[ (P_{i}^T) \Pi_{i-N/N-i}^j \Pi_{i-N/N-i}^j \right] \geq 2 \sum_{j=1}^{M} k_{ij} \Pi_{i-N/N-i}^j
\]  

(21)
also satisfy the LMI (12). Therefore, the minimization problem (20) can be replaced by the following decentralized one, performed by each sensor \(i \in \mathcal{Y}'\):

\[
\min \left( \text{trace}(\Pi_{i-N/N-i}^j) \right), \text{ subject to (21)}
\]

Notice, that in the solution provided by Proposition 1, each node computes \(\Pi_{i-N/N-i}^j\) solely on the basis of the information provided by its neighbors, that amounts to the matrices \(\Pi_{i-N/N-i}^j, j \in \mathcal{Y}'\).

### III. CONVERGENCE PROPERTIES OF DMHE

When the network is composed by a single sensor one has \(K = 1\) and DMHE reduces to the MHE scheme, for which convergence and stability have been established in [12]. The main purpose of this Section is to extend the convergence results of [12] to the proposed DMHE scheme.

**Definition 2:** Let \(\Sigma\) be system (1) with \(w = 0\) and denote by \(x_\Sigma(t, x_0)\) the state reached by \(\Sigma\) at time \(t\) starting from initial condition \(x_0\). Assume that the trajectory \(x_\Sigma(t, x_0)\) is feasible, \(i.e., x_\Sigma(t, x_0) \in \mathcal{X}\) for all \(t\). DMHE is convergent if

\[
\|x_\Sigma(t, x_0) - x_i(t, x_0)\| \to 0 \text{ for all } i \in \mathcal{Y}'
\]

Note that, as in [12], convergence is defined assuming that the model generating the data is noiseless, but the possible presence of noise is taken into account in the state estimation algorithm. Now, if we define the collective vectors \(\hat{s}_{i/N-N-i} = [\hat{q}_{i1}^T, \ldots, \hat{q}_{iM_i}^T]^T \in \mathbb{R}^{MN}\) and \(x_\Sigma(t, x_0) = I_n \otimes x_\Sigma(t, x_0)\), the following intermediate result can be stated.

**Lemma 1:** If \(\Pi_{i-N/N-i}^j\) are feasible, \(i.e., x_\Sigma(t, x_0) \in \mathcal{X}\) for all \(t\). DMHE is convergent if

\[
\|x_\Sigma(t, x_0) - x_i(t, x_0)\| \to 0 \text{ for all } i \in \mathcal{Y}'
\]

The LMI (12) deserves a few comments. First, condition (12) is required to guarantee convergence of the DMHE scheme. Second, the choice of \(\Pi_{i-N/N-i}^j\) verifying (12) is not unique. Intuitively, matrices \(\Pi_{i-N/N-i}^j\) model the uncertainty one has about the term \(\tilde{z}_{i-N} - \tilde{z}_{i-N-1}\) and therefore one would make the left hand side of the inequality (12) “as close as possible” to the right hand side. A way for achieving this is to solve the LMI problem

\[
\text{min } (\text{trace}(\Pi_{i-N/N-i}^j)) \text{ subj. to (12)}
\]  

(20)
where \(\Pi_{i-N/N-i}^j\) has the structure given in (11). Notice that (20) can be solved by each sensor since, similarly to the formula for updating covariances in Kalman filtering, the computation of \(\Pi_{i-N/N-i}^j\) does not depend upon the collected measurements. However, problem (20) has a centralized flavor since each sensor needs to know the matrices \(P_{i}^T\) and \(Q_{i}^T\) of all sensors, and the matrix \(K\) encoding the graph topology. These limitations are severe, since, for instance, the LMI (12) has size \(n \times M\) which implies that the computational burden required at each sensor for solving (20) scales with the number of sensors, hence hampering the application of DMHE to large networks. The next proposition provides a way to circumvent this problem.

**Theorem 1:** Under the assumptions of Lemma 1, DMHE is convergent if the matrix \(\Phi\) is Schur.

The fundamental assumption that matrix \(\Phi\) is Schur is motivated by Lemma 1 and does not require that system (1) is asymptotically stable. Moreover, Theorem 1 does not hinge on observability properties. For instance, it embraces the case where a component of the state of system (1) is not observed by any sensor but the estimation error decays to zero because the state component has the same property. Nevertheless, understanding how observability conditions affect the error dynamics is a topic of great interest. As a motivation, consider the problem of designing a linear state estimator when the network is composed by a single sensor. It is well known that if system (1) is detectable the error dynamics will inherit the unobservable eigenvalues of \(A\). However, if system (1) is observable one can design estimators such that the eigenvalues of the error dynamics do not depend upon the eigenvalues of \(A\).

In our framework, this rises the problem of studying conditions for guarantees that the matrix \(\Phi\) does not inherit any (non zero) eigenvalues of \(A\). More formally, let \(\lambda_{j}^{i}\) and \(\nu_{j}^{i}\) be the eigenvalues and the eigenvectors of \(A\), respectively, with \(i = 1, \ldots, n\). Note that, in view of its definition, the eigenvalues of \(A\) are \(\lambda_{j}^{i}\) \((i = 1, \ldots, n)\), each one with multiplicity \(M\). Moreover, denote by \(e_{j}\) \((j = 1, \ldots, M)\) the canonical basis vectors of \(\mathbb{R}^{M}\), so that the eigenspace related to \(\lambda_{j}^{i}\) is \(\text{span}(e_{1} \otimes V_{A}, \ldots, e_{M} \otimes V_{A})\). We consider the following property.

**Property 1:** If \(\lambda_{j}^{i}\) is a non-zero eigenvalue of \(A\), for all \(x \in \text{span}(e_{1} \otimes V_{A}, \ldots, e_{M} \otimes V_{A})\), \(\lambda_{j}^{i}\) and \(x\) are not an eigenvalue/eigenvector pair for \(\Phi\).

Before giving the main results, we introduce the definition of isolated subgraph. If the graph \(G\) is not strongly connected \((i.e., \text{it is reducible}), one can partition \(G\) into \(l\) nonempty irreducible subgraphs \(G_{j}^{i} = (\lambda_{j}^{i}, e_{j}^{i}), j = 1, \ldots, l, \) see e.g. [4]. If \(p \in A_{j}^{*}\) and \(q \in T_{p}\) imply that \(q \in A_{j}^{*}\) we say
that \( G^*_i \) is isolated. Remark that, if \( \mathcal{G} \) is strongly connected, it is also isolated.

**Theorem 2:** Consider a partition of \( \mathcal{G} \) into the irreducible subgraphs \( \mathcal{G}^*_i, \ i = 1, \ldots, l \). If for all the isolated strongly connected subgraphs \( \mathcal{G}^*_i \), it holds

\[
\bigcap_{j \in \mathcal{N}_i} \ker(\partial_j^G) = 0 \tag{24}
\]

then Property 1 is verified. □

The example in Fig. 1 illustrates the meaning of condition (24).

![Figure 1](image)

**Fig. 1.** The graph is decomposed into three connected subgraphs: \( \mathcal{G}_1, \mathcal{G}_2 \) and \( \mathcal{G}_3 \). Notice that the node 2 (of \( \mathcal{G}_2 \)) is a neighbor of the node 3 of \( \mathcal{G}_3 \). Therefore, graph \( \mathcal{G}_1 \) is not isolated. Analogously, the graph \( \mathcal{G}_2 \) is not isolated, in that the node 1 (of \( \mathcal{G}_1 \)) is a neighbor of the node 2. According to the definition, the subgraph \( \mathcal{G}^*_1 \) is isolated. Condition (24) requires collective observability (see Section II-A) for the nodes of the subgraph \( \mathcal{G}^*_1 \) solely.

In the case of strongly connected graphs we have the following result.

**Corollary 1:** If \( \mathcal{G} \) is strongly connected and the system is collectively observable, then Property 1 is verified. □

As a particular case, assume that all sensors enjoy regional observability and are arranged in a strongly connected graph \( \mathcal{G} \). This yields \( P_{\text{NO}} = \Phi = 0_{nM \times nM} \) and convergence of DMHE follows from Theorem 1. Moreover, since the system is also collectively observable, Corollary 1 guarantees that Property 1 holds.

![Figure 2](image)

**Fig. 2.** Communication network and associated matrix \( K \) used in the example.

**IV. EXAMPLE**

We consider the fourth order system

\[
x_{t+1} = \begin{bmatrix} 0.9962 & 0.1949 & 0 & 0 \\ -0.1949 & 0.3819 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -1.21 & 1.98 \end{bmatrix} x_t + w_t \tag{25}
\]

where \( x_t = [x_{1,t}, x_{2,t}, x_{3,t}, x_{4,t}]^T \). Notice that the eigenvalues of the system’s matrix \( A \) are 0.9264, 0.4517, 0.990.4795 and, since \( |0.990.4795| > 1 \), the system is unstable.

Let \( e_t \in \mathbb{R}^4 \), be white noise with covariance \( Q_e = \text{diag}(0.0012, 0.038, 0.0012, 0.038) \). In the following we consider two cases

A. \( w_t = e_t, \ Q = Q_e \) and \( \mathbb{W} = \mathbb{R}^4 \) (unconstrained input noise)

B. \( w_t = |e_t|, \ Q = Q_e \) and \( \mathbb{W} = \mathbb{R}^4_{\geq 0} \) (constrained input noise)

In both cases, we set \( \mu = [0 \ 0 \ 0 \ 0]^T \), \( \Pi_0 = 100I_4 \) and \( N = 2 \) in the DMHE algorithm.

The state of (25) is measured by \( M = 4 \) sensors with sensing model

\[
y_i^t = [1 \ 0 \ 0 \ 0] x_t + v_i^t \quad \text{if} \ i = 1, 2
\]

\[
y_i^t = [0 \ 0 \ 1 \ 0] x_t + v_i^t \quad \text{if} \ i = 3, 4
\]

where \( \text{Cov}(v_i^t) = R_i, \ i = 1, 2 \ldots, 4 \). Sensors are connected according to the graph in Fig. 2, where the matrix \( K \) is also given. It is apparent that the information available, at each instant, to node 1 consists of the measurements of \( x_{1,t} \) and \( x_{3,t} \) (transmitted by sensor 4). Analogously, the information available to node 3 consists of \( x_{1,t} \) (transmitted by sensor 2) and \( x_{3,t} \). It is easy to check that the system is regionally observable by sensors 1 and 3. On the other hand, at each time instant sensor 2 can only use two different measurements of \( x_{1,t} \) (produced by sensors 1 and 2). Similarly, sensor 4 can only use two different measures of \( x_{3,t} \) (produced by sensors 3 and 4). Therefore, the system is not regionally observable by sensors 2 and 4. In fact, \( P_{\text{RLO}} = \text{diag}(0, 0, 1, 1), \ P_{\text{RLO}} = \text{diag}(1, 1, 0, 0) \). The eigenvalues of the matrix \( \Phi \) defined in (23) are 0, 0.4632, 0.2258 and 0.49500.2397i. Since \( \Phi \) is Schur, convergence of DMHE is guaranteed by Theorem 1. Moreover, since the graph is strongly connected and collective observability holds, Corollary 1 guarantees that also Property 1 holds.

In Fig. 3 the estimation errors produced by all sensors in the scenario A are shown. It is worth noticing that the estimates produced by sensors 2 and 4, relative to states \( x_{3,t} \) and \( x_{1,t} \) and \( x_{3,t} \) and \( x_{1,t} \), respectively, display big errors for \( t < 6 \). In fact, these states cannot be observed by these sensors using regional measurements. Nonetheless, the estimation errors of all sensors asymptotically converge to the same values, thanks to the consensus action embodied in the proposed algorithm.

Fig. 4 depicts the evolution of the eigenvalues of matrices \( \Pi_{t+t-N+1} \) over time. Note that these matrices are the same in the cases A and B. Indeed, the update procedure described in Section II-C does not depend on the estimates and can be run off-line.

The estimation errors for case B are depicted in Fig. 5. Analogously to case A, convergence of DMHE can be noticed.

![Figure 5](image)

**V. CONCLUSIONS**

Many generalizations of the DMHE scheme described in the paper can be considered in order to enhance its potentialities.
A first one is the development of DMHE schemes enjoying convergence even in presence of state constraints. In addition we will study how to: (i) exploit the degrees of freedom in the choice of the matrix $K$ in order to improve the speed of convergence of the state estimates provided by sensors; (ii) extend the DMHE scheme to time-varying graphs; (iii) use properly multiple transmissions between sensors within a sampling interval when the network bandwidth is sufficiently large.

**REFERENCES**


