Available online at www.sciencedirect.com

ScienceDirect

IFAC PapersOnLine 50-1 (2017) 2494–2499

Distributed consensus-based Kalman filtering considering subspace decomposition

A.R. del Nozal * L. Orihuela * P. Millán *

* Departamento de Ingeniería, Universidad Loyola Andalucía, Spain
{arodriguez,dorihuela,pmillan}@uloyola.es

Abstract: The aim of this paper is to provide a new observer structure able to deal with the distributed estimation of a discrete-time linear system from a network of agents. The main result is an innovative consensus-based structure that decompose the state in the observable and unobservable subspace of the agent using the observability staircase form. The paper proposes a design in which Kalman-like gains are synthetized to minimize the variance of the error on both subspaces. Finally some simulations are shown to compare the proposed estimator with centralized Kalman filter and other distributed schemes found in literture.

Keywords: Estimation and filtering. Distributed control and estimation. Sensor networks.

1. INTRODUCTION

The Kalman filter is, perhaps, the most well-known estimator for processes affected by statistical noises and disturbances. In recent years, its distributed implementation, the so-called Distributed Kalman Filter (DKF) (see Olfati-Saber (2009) and references therein), has attracted the attention of many researchers, not only those having a systems and control background, but also groups focused on signal processing (Ng et al. (2008)), computer vision (Chen (2012); Kamal et al. (2013)) or optics ((Gilles et al. (2013)) to name just a few.

One of the main reasons for this widespread expansion is that the distributed architectures offer interesting advantages with respect to old centralized schemes, such as scalability, flexibility, fault tolerance or robustness.

As mentioned before, scalability is one of the positive features that these estimators must present. Scalability refers to the adaptation of the infrastructure in order to include additional devices and to the computational and communication costs derived from this increasing set of agents. The preliminary estimators found in the literature lack of this feature designing the observers in a unique centralized step, see Millán et al. (2012); Ugrinovskii (2011).

Some authors, in pursuing reduced communication costs, propose estimators with limited communications among the agents, as the ones in Ribeiro et al. (2006); Msechu et al. (2008); Orihuela et al. (2013); Wang et al. (2017). Another interesting line of research concerns the diffusion DKF (see Hu et al. (2012); Zhang et al. (2015)), that reduces the communication steps required in consensus-based DKF (Carli et al. (2008)), while keeping the distributed design of the observer gains. The works in Ugrinovskii (2013); Yan et al. (2015) study the case with variable topology, which is interesting for networks with variable number of agents. The work Song et al. (2013) study the case with sensor noises cross-correlated.

This paper proposes a novel approach to the topic of distributed Kalman filtering, consisting in a decomposition of the state-space in two orthogonal subspaces (the *observable* and *unobservable* one). The proposed observer uses the local information from the plant to correct the locally observable subspace, whereas the locally unobservable subspace is estimated with the information provided by other agents using a consensus-based algorithm.

The proposed architecture has several positive features. The first one is that the design of the observer gains can be carried out in a distributed way. And this is done by exchanging a reduced amount of information between the agents, compared to the works of Olfati-Saber (2007); Cattivelli and Sayed (2010). Another interesting feature is that the design of the gains for the observable and unobservable subspaces are done in decoupled steps.

This paper constitutes a preliminary work where the authors intention is to illustrate the potential of this observer structure and the features introduced. It is also shown, through the simulations provided, that the proposed architecture also renders stabilizing estimators when coupling between observable subspaces of the agents is considered. However a mathematical proof is still required.

This paper is organized as follows. Section 2 introduces some notation and mathematical preliminaries needed for the rest of the paper. The distributed problem and the assumptions taken under consideration are presented in Section 3. Section 4 describes the proposed observer structure. In Section 5 a method to choose the estimation gains is given. Section 6 shows different simulation examples with a complete comparative with other observer structures. Finally, in Section 7 the main conclusions are drawn and discussed.

⋆ Research partially supported by grants DPI2013-44135-R funded by MCyT and TEC2016-80242-P by AEI/FEDER

10.1016/j.ifacol.2017.08.443

2. NOTATION AND PRELIMINARIES

2.1 Observability staircase form

Consider a discrete-time linear autonomous system:

$$x^+ = Ax, \tag{1}$$
$$y = Cx, \tag{2}$$

where $x \in \mathbb{R}^n$ is the state, $y \in \mathbb{R}^m$ is the measured output, and $A \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{m\times n}$ are known dynamic and output matrices. Then, it is possible to find a coordinate transformation matrix $T \in \mathbb{R}^{n\times n}$ such that the change of variable $\xi \triangleq Tx \in \mathbb{R}^n$ transforms the original state-space representation into the observability staircase form:

$$\xi^+ = TAT^T\xi = \begin{bmatrix} A^{\overline{o}} & \tilde{A} \\ 0 & A^o \end{bmatrix} \xi, \tag{3}$$
$$y = CT^T\xi = \begin{bmatrix} 0 & \tilde{C} \end{bmatrix} \xi, \tag{4}$$

where the transformed state ξ can be divided into: $\xi = \begin{bmatrix}\xi^{\overline{o}T} & \xi^{oT}\end{bmatrix}^T$, being $\xi^{\overline{o}} \in \mathbb{R}^{n\overline{o}}$ and $\xi^o \in \mathbb{R}^{no}$ are the unobservable and observable states of the system respectively. Note the complete decoupling of the observable part of the system from the unobservable one with $(A^o, \tilde{C})$ observable.

It is possible to write the coordinate transformation matrix T as $T = \begin{bmatrix}V^{\bar{o}} & V^o\end{bmatrix}^T$, where $V^{\bar{o}} \in \mathbb{R}^{n\times n\bar{o}}$ is composed by $n\bar{o}$ column vectors in $\mathbb{R}^n$ which form and orthonormal basis of the unobservable subspace of system (1)-(2). Correspondingly, $V^o \in \mathbb{R}^{n\times no}$ is an orthonormal basis of its orthogonal complement. Both basis are orthonormal and mutually orthogonal and altogether conform the whole space $\mathbb{R}^n$. Taking this under consideration, the following equations hold:

$$V^{\overline{o}T}V^{\overline{o}} = I_{n\overline{o}}, \quad V^{oT}V^o = I_{no},$$
$$V^{\overline{o}T}V^o = 0_{n\overline{o}\times no}, \quad V^{oT}V^{\overline{o}} = 0_{no\times n\overline{o}}.$$

Let $S^o = Gen\{V^o\}$ and $S^{\overline{o}} = Gen\{V^{\bar{o}}\}$ be the observable and unobservable subspaces associated to these bases.

The following two operations are at the core of the proposed observation structure given in Section 4. Let $u \in \mathbb{R}^n$ be any vector, let $V \in \mathbb{R}^{n\times m}$ be a matrix, composed by a subset of m linearly independent column vectors which form a base of a subspace of $\mathbb{R}^n$ and finally let $A \in \mathbb{R}^{m\times m}$ be any matrix. Then:

(1) Operator $\Pi_V(u) \doteq VV^Tu \in \mathbb{R}^n$ is defined as the projection of u onto $Gen\{V\}$.
(2) The following operator is also defined:
$$\Pi_V^A(u) = VAV^Tu$$
which represents a weighted projection of vector u onto the subspace $Gen\{V\}$. The weights are given by matrix A.

2.2 Graph theory

A graph is an ordered pair $G = (\mathcal{V}, \mathcal{E})$ comprising a set $\mathcal{V} = \{1, 2, \ldots, p\}$ of vertices and a set $\mathcal{E} \subset \mathcal{V}\times\mathcal{V}$ of edges. A *directed graph* is a graph in which edges have orientations, so if $(j,i) \in \mathcal{E}$, then agent i obtains information from agent j. A directed path from node i_1 to node i_k is a sequence of edges such as $(i_1, i_2), (i_2, i_3), \cdots, (i_{k-1}, i_k)$ in a directed graph.

The neighbourhood of i, $N_i \triangleq \{j : (j,i) \in \mathcal{E}\}$ is defined as the set of nodes with edges incoming to node i.

The *extended neighbourhood of* i *through* j, $N_i(j)$, is formed by those vertices with a direct path to node i that includes edge (j,i).

3. PROBLEM FORMULATION

Consider a graph G where the agents are located at the vertices intended to distributedly estimate the state of the following discrete-time linear system:

$$x^+ = Ax + w, \tag{5}$$
$$y_i = C_ix + n_i \quad \forall i \in \mathcal{V}, \tag{6}$$

where $y_i \in \mathbb{R}^{m_i}$ is a vector containing the measured output of agent i, $C_i \in \mathbb{R}^{m_i\times n}$ is the output matrix of agent i and $w \in \mathbb{R}^n$ and $n_i \in \mathbb{R}^{m_i}$ are mutually independent white Gaussian state and measurement noises, respectively, with covariance matrices $Q \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m\times m}$.

The observation structure proposed in the next section relies on system transformations in the observability staircase form of each agent, according to the localy measured outputs (defined by pairs $(A, C_i)\forall i \in \mathcal{V}$). Hence, for each agent the observable and unobservable subspaces are defined as $S_i^o = Gen\{V_i^o\}$ and $S_i^{\overline{o}} = Gen\{V_i^{\bar{o}}\}$, respectively.

Let $e_i \triangleq x - \widehat{x}_i \in \mathbb{R}^n$ be the estimation error of agent i and $P_{ii} = E\{e_ie_i^T\} \in \mathbb{R}^{n\times n}$ be the covariance matrix associated to the estimation error of that agent. Then, $tr(P_{ii})$ is the variance of the estimation error of agent i. Analogously it is defined $P_{ij} = E\{e_ie_j^T\}$ as the cross-covariance matrix.

The objective of each of the agents is to reconstruct the whole state of system (5), through a minimum variance estimator that makes use of local measurements and also estimates received from neighbouring agents.

Definition 1. The *extended observable subspace of* i *through* j, $\widehat{S}_i^o(j)$, is the union of the observable subspace of j and those observable subspaces of the agents belonging to the extended neighbourhood of i through j. This is:

$$\widehat{S}_i^o(j) \triangleq S_j^o \cup \left(\bigcup_{k\in N_i(j)} S_k^o\right). \tag{7}$$

Then $V_i^o(j) \in \mathbb{R}^{n\times no_i(j)}$ is the matrix comprised of the basis vectors of $\widehat{S}_i^o(j)$. The *extended unobservable subspace of* i *through* j, $\widehat{S}_i^{\overline{o}}(j)$, and matrix $V_i^{\overline{o}}(j) \in \mathbb{R}^{n\times n\overline{o}_i(j)}$ are analogously defined.

Assumption 2. Collective observability for each one of the agents is a necessary assumption. This means that, for each agent i, it holds:

$$\mathbb{R}^n \subseteq S_i^o \cup \left(\bigcup_{j\in N_i} \widehat{S}_i^o(j)\right),$$

which implies that any agent i must be able to observe the complete state using the information provided by its neighbourhood.

It is worth mentioning that local observability is not assumed, that is, no node is able to estimate the full plant state based only on its direct measurements of the plant.

Assumption 3. Any agent i knows $\widehat{S}_i^o(j)$ for every $j \in N_i$.

This assumption will be necessary to define the observer structure.

4. PROPOSED OBSERVATION STRUCTURE

Consider the following observer structure:

$$\begin{aligned}\widehat{x}_i^+ &= A\widehat{x}_i + V_i^o L_i(y_i - \widehat{y}_i) + \\ &+ V_i^{\overline{o}} V_i^{\overline{o}T} \sum_{j\in N_i} V_i^o(j) N_{ij} V_i^o(j)^T (\widehat{x}_j - \widehat{x}_i)\end{aligned} \quad (8)$$

where $\widehat{x}_i \in \mathbb{R}^n$ and $\widehat{y}_i \in \mathbb{R}^{m_i}$ are the estimation of the plant state x and the output y_i performed by agent i and $L_i \in \mathbb{R}^{no_i \times m_i}$ and $N_{ij} \in \mathbb{R}^{no_i(j)\times no_i(j)}$ are, respectively, a local correction-term and consensus-based correction gains to be designed.

It is worth pointing out that the second term is a Kalman-like correction term in which the estimation error, weighted with the output matrix and the local Kalman gain,L_i, is projected onto the observable states of the agent:

$$V_i^o L_i(y_i - \widehat{y}_i) = \Pi_{V_i^o}^{L_i\tilde{C}_i} e_i + V_i^o L_i n_i.$$

On the other hand, the consensus term weights the difference between the estimation made by the neighbors j and agent i with matrix N_{ij}, and the resulting vectors are finally projected onto the unobservable part of agent i:

$$\Pi_{v_i^{\overline{o}}} \left(\sum_{j\in N_i} \Pi_{V_i^o(j)}^{N_{ij}} (\widehat{x}_j - \widehat{x}_i) \right).$$

Thus, the information provided by the observable states of the neighbourhood of i is used to correct its unobservable states.

5. OBSERVER METHOD

The aim of this section is to provide a design method for the observer defined in (8) in order to minimize $tr(P_{ii})$.

Proposition 4. The evolution of the covariance matrix P_{ii} and the cross-covariance matrix P_{ij} are given by:

$$\begin{aligned}P_{ii}^+ &= E\left\{e_i^+ e_i^{+T}\right\} \quad (9)\\ &= \widehat{A}_i P_{ii} \widehat{A}_i^T + \widehat{A}_i \sum_{p\in N_i} P_{ip}\widehat{N}_{ip}^T + \sum_{p\in N_i} \widehat{N}_{ip} P_{ip} \widehat{A}_i^T \\ &+ \sum_{p\in N_i}\sum_{q\in N_i} \widehat{N}_{ip} P_{pq} \widehat{N}_{iq}^T + Q + V_i^o L_i R_i L_i^T V_i^{oT},\end{aligned}$$

$$\begin{aligned}P_{ij}^+ &= E\left\{e_i^+ e_j^{+T}\right\} \quad (10)\\ &= \widehat{A}_i P_{ij} \widehat{A}_j^T + \widehat{A}_i \sum_{q\in N_j} P_{iq}\widehat{N}_{jq}^T + \sum_{p\in N_i} \widehat{N}_{ip} P_{jp} \widehat{A}_j^T \\ &+ \sum_{p\in N_i}\sum_{q\in N_j} \widehat{N}_{ip} P_{pq} \widehat{N}_{jq}^T + Q,\end{aligned}$$

where matrices $\widehat{A}_i$ and $\widehat{N}_{ip}$ are defined as:

$$\widehat{A}_i = A - V_i^o L_i C_i - \sum_{p\in N_i} \widehat{N}_{ip},$$

$$\widehat{N}_{ip} = V_i^{\overline{o}} V_i^{\overline{o}T} V_i^o(p) N_{ip} V_i^o(p)^T.$$

Proof. Considering (5) and (8) the dynamics of the estimation error is given by:

$$\begin{aligned}e_i^+ &= x^+ - \widehat{x}_i^+ \\ &= \left(A - V_i^o L_i C_i - V_i^{\overline{o}} V_i^{\overline{o}T} \sum_{p\in N_i} V_i^o(p) N_{ip} V_i^o(p)^T \right) e_i \\ &+ \sum_{p\in N_i} \left(V_i^{\overline{o}} V_i^{\overline{o}T} V_i^o(p) N_{ip} V_i^o(p)^T \right) e_p + (w - V_i^o L_i n_i) \\ &= \widehat{A}_i e_i + \sum_{p\in N_i} \widehat{N}_{ip} + w - V_i^o L_i n_i.\end{aligned}$$

By multiplying e_i and its transpose it is easy to obtain expression (9). Proceeding analogously with e_i and e_j it is possible to obtain expression (10). Note that the expectation of the product of two uncorrelated variables is zero. Thus, the expectation of the product of the noise term and any estimation error vanishes from the structure. □

In order to keep track of the cross-covariance matrices, the agents need to receive the Kalman and consensus gains of the other agents. This is a serious drawback in distributed systems, since the amount of information grows exponentially with the number of agents. This problem is tackled in Subsection 5.2.

Theorem 5. The set of matrices (L_i, N_{ij}) that minimize $tr(P_{ii}(k+1))$ are given by the following equations:

$$L_i = V_i^{oT} A P_{ii} C_i^T \left[C_i P_{ii} C_i^T + R_i\right]^{-1}, \quad (11)$$

$$\begin{aligned}V_i^o(j)^T V_i^{\overline{o}} V_i^{\overline{o}T} A\left(P_{ij} - P_{ii}\right) V_i^o(j) &= V_i^o(j)^T V_i^{\overline{o}} V_i^{\overline{o}T} \times \\ \times \sum_{p\in N_i} V_i^o(p) N_{ip} V_i^o(p)^T M_{ijp} V_i^o(j), &\quad \forall j \in N_i, \quad (12)\end{aligned}$$

where $M_{ijp} = (P_{ip} - P_{jp} + P_{ij} - P_{ii})$.

Proof. With the purpose of minimizing the trace of the covariance matrix, equation (9) is partially derived with respect to L_i and N_{ij}, yielding:

$$\begin{aligned}\frac{\partial tr(P_{ii}^+)}{\partial L_i} &= -2V_i^{oT}\left[A P_{ii} C_i^T - \right. \\ &\left. - V_i^o L_i \left(C_i P_{ii} C_i^T + R_i\right)\right] = 0, \quad (13)\end{aligned}$$

$$\begin{aligned}\frac{\partial tr(P_{ii}^+)}{\partial N_{ij}} &= -2V_i^o(j)^T V_i^{\overline{o}} V_i^{\overline{o}T} A\left(P_{ii} - P_{ij}\right) V_i^o(j) - \\ &- 2V_i^o(j)^T V_i^{\overline{o}} V_i^{\overline{o}T} \sum_{p\in N_i} V_i^o(p) N_{ip} V_i^o(p)^T M_{ijp} V_i^o(j) = 0.\end{aligned} \quad (14)$$

Finally, isolating $L_i(k)$ in (13) and operating in (14) it is possible to obtain expressions (11) and (12). □

This theorem presents an explicit equation for L_i and an expression that, if fulfilled, the observer structure presented in (8) minimizes the variance of the estimation error in the next step.

5.1 Decoupled observable subspaces

This subsection proves that, under some mild assumptions, the consensus gains can be obtained by solving a system of linear equations

Assumption 6. There is no intersection between the observable subspaces of any pair of agents, this is:

$$S_i^o \cap S_j^o = \emptyset \quad \forall i \neq j.$$

This assumption does not incur in additional constrains with respect to Assumption 2. By neglecting the information contained in the coupled observable subspaces, Assumption 2 implies that Assumption 6 is fulfilled. In fact, it can be written as a particular case of Assumption 2:

$$\mathbb{R}^n = S_i^o \cup \left(\bigcup_{j \in N_i} \widehat{S}_i^o(j) \right)$$

Proposition 7. Considering Assumption 6, then the next equality holds:

$$V_i^o(j)^T V_i^{\overline{o}} V_i^{\overline{o}^T} V_i^o(j) = I_{no_i^a(j) \times no_i^a(j)}. \tag{15}$$

The proof of this proposition is reported in the appendix.

Next result, whose proof can be also found in the appendix, states that the value of each matrix N_{ij} for every pair of agents (i, j) can be always obtained.

Proposition 8. Considering that Assumption 6 holds, the system of equations (14) is always consistent, namely it has at least one solution for N_{ij}.

The case in which there is intersection between observable subspaces of agents, $S_i^o \cap S_j^o \neq \emptyset, \quad i \neq j$, will be tackled in future research. However some simulations are shown in which a better response is obtained. In Example 3 equation (14) is solved with an iterative method but its consistency is not proved.

5.2 Distributed computational observers

In this section we propose a simplification of Proposition 4 introducing an additional approximation.

Approximation. In order to reduce the exchange of information favoring the distributed topology, it will be considered that the value of the cross-covariance matrices are zero for any pair of agents.

This approximation, although is not mathematically proven to be exact, renders similar results to the original case, as it will be shown in the simulations.

Proposition 9. Considering the previous approximation, the expression of the covariance matrix exposed in (9) can be rewritten as:

$$P_{ii}^+ = \widehat{A}_i P_{ii} \widehat{A}_i^T + \sum_{p \in N_i} \widehat{N}_{ip} P_{pp} \widehat{N}_{ip}^T + Q + V_i^o L_i R_i L_i^T V_i^{oT}.$$

This proposition implies that gain matrices L_i and N_{ij} are calculated locally. To calculate the set of matrices $N_{ij}, \quad \forall j \in N_i$ agent i needs the covariance matrix of each of its neighbours P_{jj}, pre and post multiplied by $v_i^o(p)$. Hence, the total amount of data required from neighboring agents is lower than needed in Olfati-Saber (2007) and Cattivelli and Sayed (2010).

6. SIMULATIONS

In this section, three simulation examples are presented in order to show the effectiveness of the proposed observer. The algorithm presented will be compared with two other observer structures: the Centralized Kalman Filter (CKF) and the algorithm 3 introduced in Olfati-Saber (2007). In all the cases the covariance matrices considered for the system and measurements noises are $Q = I_n \otimes 0.02, \quad R = I_m \otimes 0.02$.

Example 1: Consider the following system:

$$x^+ = \begin{bmatrix} 1.005 & 0 & 0 & 0 \\ 0 & 0.9954 & -0.08757 & 0 \\ 0 & 0.1248 & 0.9945 & 0 \\ 0 & 0 & 0 & 0.9775 \end{bmatrix} x + w,$$

which is being observed by three agents in such a way that $y_1 = x_1$, $y_2 = x_2$ and $y_3 = x_4$. The linear topology $1 \leftrightarrow 2 \leftrightarrow 3$ is considered.

It is clear from the chosen system that each agent observes different modes of the states fulfilling assumption 6. The basis vectors of the observable subspace of agents can be easily obtained as:

$$V_1^o = [1\ 0\ 0\ 0]^T, \quad V_2^o = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}^T, \quad V_3^o = [0\ 0\ 0\ 1]^T.$$

Figure 1 shows the evolution of the state modes and agent 3 estimates. Note that, for this agent, a local design is made to estimate state x_4 meanwhile the rest of states are estimated with the information provided by agent 2.

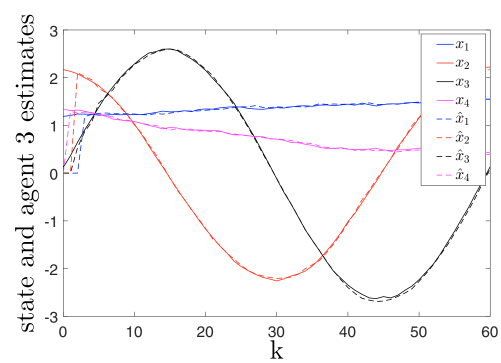

Fig. 1. State of the plant in solid lines and agent 3 estimates in dashed lines considering Proposition 4.

Regarding the other observer structures considered, Figure 2 shows the evolution of the trace of the covariance matrix for each one of the algorithms considered. The algorithm introduced by Olfati-Saber fails at estimating the error due to the fact that the dynamic consensus algorithm proposed in Spanos and Murray (2005) does not converge for unstable poles. This simulation shows that approximation of zero cross-covariances does not incur in a big worsening of the performance with respect to the optimal case.

Example 2: Consider now the same network topology than Example 1 and consider the same plant changing the unstable mode of the system for the stable one $x_1^+ = 0.99x_1 + w_1$.

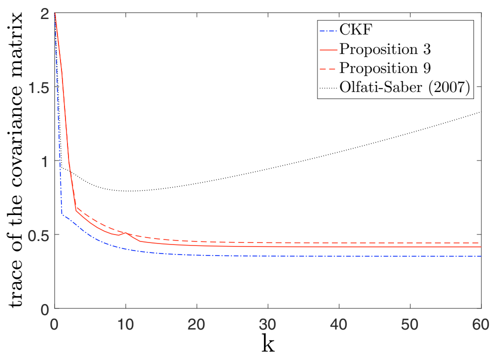

Fig. 2. Evolution of the average trace of the covariance matrix of all the agents for each one of the algorithms considered in the Example 1.

The estimator in Olfati-Saber (2007) is now able to estimate the whole state. Nevertheless, the value of the trace of the covariance matrix in steady state is higher than the algorithm presented which get closer to the performance of the CKF.

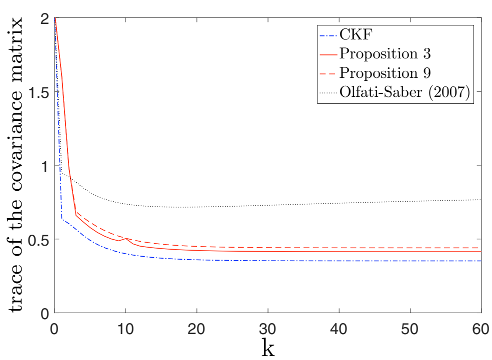

Fig. 3. Evolution of the average trace of the covariance matrix of all the agents for each one of the algorithms considered in the Example 2.

Example 3: Consider the system and the topology defined in Example 1 with the next output matrices for the agents:

$$C_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}^T, \quad C_2 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0.9 & 0 \\ 0 & 1.05 \end{bmatrix}^T, \quad C_3 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1.1 & 0 \\ 0 & 0.8 \end{bmatrix}^T.$$

In spite that Assumption 6 is not applicable in this case, the estimators are still able to observe the complete state. Furthermore, agents 2 and 3 are able to observe the same modes of the system and share their estimates with agent 1 reducing in that way the variance of the estimation error of the set. Figure 4 shows the evolution of the trace of the covariance matrix of agent 1 for the output matrices defined in Example 1 and Example 3 considering Proposition 9 in the design of (L_i, N_{ij}). The same initial conditions are considered.

Due to the fact that states x_2, x_3 and x_4 are measured by two agents, the variance of the estimates of each one of these modes decrease considerably respect to the case exposed in Examples 1. This is deducible from Figure 4.

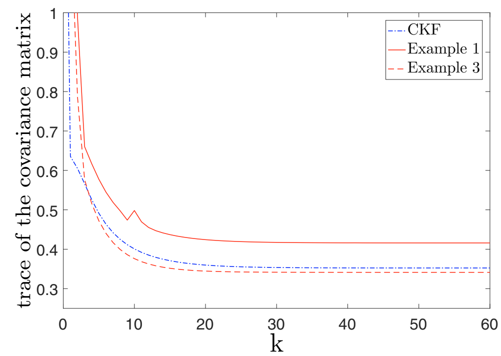

Fig. 4. Evolution of the trace of the covariance matrix of agent 1 for the agents configuration defined in Example 1 and Example 3.

7. CONCLUSIONS

The observer structure presented in this paper introduces a new way to analyze the distributed estimation problem for a network of agents. Decomposing the system using the observability staircase form it is possible to design a set of gains that affect to the observable and unobservable subspace of the system independently. Furthermore, it is possible to design the estimator minimizing the variance of the estimation error of every agent. It is worth mentioning that the design can be done in a distributed way. This work establishes a preliminary result and a few topics related with the application of subspace decomposition will be tackled in future research, such as the study of consistency of the resolution of the system of equations presented in (14) in a general case or quantify the impact of the zero cross-covariance approximation in the observer.

APPENDIX

Proof of Proposition 7: With some abuse of notation, $V_i^o(j)$ and $V_i^{\bar{o}}$ will be denoted as V^o and $V^{\bar{o}}$ respectively. Each of these bases are composed by no and $n\bar{o}$ orthonormal vectors:

$$V^o = [v_1^o \; v_2^o \; \cdots \; v_{no}^o], \quad V^{\bar{o}} = \left[v_1^{\bar{o}} \; v_2^{\bar{o}} \; \cdots \; v_{n\bar{o}}^{\bar{o}}\right].$$

Let's define matrix H as $H = V^{\bar{o}^T} V^o$. Then, equation (15) can be rewritten as $H^T H = I_{no \times no}$.

Let's decompose H in the product of the vectors of bases V^o and $V^{\bar{o}}$:

$$H = \begin{bmatrix} v_1^{\bar{o}T} \\ v_2^{\bar{o}T} \\ \vdots \\ v_{n\bar{o}}^{\bar{o}\;T} \end{bmatrix} \begin{bmatrix} v_1^{oT} \\ v_2^{oT} \\ \cdots \\ v_{no}^{o\;T} \end{bmatrix}^T = \begin{bmatrix} v_1^{\bar{o}T} v_1^o & v_1^{\bar{o}T} v_2^o & \cdots & v_1^{\bar{o}T} v_{no}^o \\ v_2^{\bar{o}T} v_1^o & v_2^{\bar{o}T} v_2^o & \cdots & v_2^{\bar{o}T} v_{no}^o \\ \vdots & \vdots & \ddots & \vdots \\ v_{n\bar{o}}^{\bar{o}\;T} v_1^o & v_{n\bar{o}}^{\bar{o}\;T} v_2^o & \cdots & v_{n\bar{o}}^{\bar{o}\;T} v_{no}^o \end{bmatrix}.$$

Under Assumption 6 every vector of V^o can be expressed as a linear combination of the vectors of $V^{\bar{o}}$. Thus, the product of that columns of $V^{\bar{o}}$ that are not a linear combination of the columns vector of V^o with these vectors is zero. Then, it is possible to reorder matrix H as:

$$H = \begin{bmatrix} \overline{H} \\ 0 \end{bmatrix}, \quad H^T H = \overline{H}^T \overline{H} = I_{no \times no}, \qquad (.1)$$

and the linear combination is expressed as:

$$v_1^{\overline{o}} = \sum_{p=1}^{no} \alpha_{1p} v_p^o, \quad , v_2^{\overline{o}} = \sum_{p=1}^{no} \alpha_{2p} v_p^o, \qquad (.2)$$

where $\alpha_{qp} \in \mathbb{R}$ are the weights of the linear combination.

Taking under consideration that column vectors of every base considered are orthonormal, then it holds:

$$v_1^{\overline{o}T} v_2^{\overline{o}} = \left(\sum_{p=1}^{no} \alpha_{1p} v_p^{oT}\right)\left(\sum_{p=1}^{no} \alpha_{2p} v_p^o\right) = \sum_{p=1}^{no} \alpha_{1p}\alpha_{2p} = 0.$$

Proving (.1) is equivalent to prove that $\overline{H}$ is orthonormal, this is, its row vectors are orthogonal between them and then the dot product of whatever pair of rows considered must be zero:

$$v_i^{\overline{o}T} \left[v_1^o \ v_2^o \ \cdots \ v_n^o\right] \left[v_1^o \ v_2^o \ \cdots \ v_n^o\right]^T v_j^{\overline{o}} = 0, \quad i \neq j.$$

Finally, applying the linear combination defined in (.2) to the expression introduced above, the proposition is proved:

$$\sum_{p=1}^{no} \alpha_{1p} v_p^{oT} V^o V^{oT} \sum_{p=1}^{no} \alpha_{2p} v_p^o = \sum_{p=1}^{n} \alpha_{1p}\alpha_{2p} = 0.$$

□

Proof of Proposition 8: By transposing equation (14), a linear system of equations with the structure $A_s N_{ij} = B_s$ is obtained where:

$$A_s = V_i^o(j)^T \left(2P_{ij} - P_{ii} - P_{jj}\right) V_i^o(j),$$
$$B_s = V_i^o(j)^T \left(P_{ij} - P_{ii}\right) A^T V_i^{\overline{o}} V_i^{\overline{o}T} V_i^o(j).$$

If Assumption 6 is satisfied then $v_i^{\overline{o}} v_i^{\overline{o}T} v_i^o(j) = v_i^o(j)$ due to the fact that $S_j^o \subseteq S_i^{\overline{o}}$. Then, matrices A_s and B_s can be rewritten as:

$$\begin{aligned} A_s &= 2V_i^o(j)^T P_{ij} V_i^o(j) - V_i^o(j)^T P_{ii} V_i^o(j) \\ &\quad - V_i^o(j)^T P_{jj} V_i^o(j), \\ B_s &= V_i^o(j)^T P_{ij} A^T V_i^o(j) - V_i^o(j)^T P_{ii} A^T V_i^o(j). \end{aligned}$$

It is well-known that a linear system of equations with the structure $A_s x = b_s$ is consistent if and only if b_s is a linear combination of column vectors of A_s. Thus, it can be seen that the columns of matrix B_s are linear combination of columns vectors of A_s multiplied by weight matrix A^T. □

REFERENCES

Carli, R., Chiuso, A., Schenato, L., and Zampieri, S. (2008). Distributed Kalman filtering based on consensus strategies. *IEEE Journal on Selected Areas in Communications*, 26(4), 622–633.

Cattivelli, F.S. and Sayed, A.H. (2010). Diffusion lms strategies for distributed estimation. *IEEE Transactions on Signal Processing*, 58(3), 1035–1048.

Chen, S.Y. (2012). Kalman filter for robot vision: a survey. *IEEE Transactions on Industrial Electronics*, 59(11), 4409–4420.

Gilles, L., Massioni, P., Kulcsár, C., Raynaud, H.F., and Ellerbroek, B. (2013). Distributed Kalman filtering compared to Fourier domain preconditioned conjugate gradient for laser guide star tomography on extremely large telescopes. *Journal of the Optical Society of America A*, 30(5), 898–909.

Hu, J., Xie, L., and Zhang, C. (2012). Diffusion Kalman filtering based on covariance intersection. *IEEE Transactions on Signal Processing*, 60(2), 891–902.

Kamal, A.T., Farrell, J.A., and Roy-Chowdhury, A.K. (2013). Information weighted consensus filters and their application in distributed camera networks. *IEEE Transactions on Automatic Control*, 58(12), 3112–3125. doi:10.1109/TAC.2013.2277621.

Millán, P., Orihuela, L., Vivas, C., and Rubio, F.R. (2012). Distributed consensus-based estimation considering network induced delays and dropouts. *Automatica*, 48(10), 2726–2729.

Msechu, E.J., Roumeliotis, S.I., Ribeiro, A., and Giannakis, G.B. (2008). Decentralized quantized Kalman filtering with scalable communication cost. *IEEE Transactions on Signal Processing*, 56(8), 3727–3741.

Ng, B.L., Evans, J.S., Hanly, S.V., and Aktas, D. (2008). Distributed dowlink beamforming with cooperative base stations. *IEEE Transactions on Information Theory*, 54(12), 5491–5499.

Olfati-Saber, R. (2007). Distributed Kalman filtering for sensor networks. In *46th IEEE Conference on Decision and Control*, 5492–5498. New Orleans, LA, USA.

Olfati-Saber, R. (2009). Kalman-consensus filter: Optimality, stability, and performance. In *48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, 7036–7042. Shanghai, P.R. China.

Orihuela, L., Millán, P., Vivas, C., and Rubio, F.R. (2013). Reduced-order H_2/H_∞ distributed observer for sensor networks. *International Journal of Control*, 86(10), 1870–1879.

Ribeiro, A., Giannakis, G., and Roumeliotis, S. (2006). SOI-KF: Distributed Kalman filtering with low-cost communications using the sign of innovations. *IEEE Transactions on Signal Processing*, 54(12), 4782–4795.

Song, H., Yu, L., and Zhang, D. (2013). Distributed set-valued estimation in sensor networks with limited communication data rate. *Journal of the Franklin Institute*, 350(5), 1264–1283.

Spanos, D.P. and Murray, R.M. (2005). Distributed sensor fusion using dynamic consensus. In *16th IFAC World Congress*, 133–139. Prague, Czech Republic.

Ugrinovskii, V. (2011). Distributed robust filtering with H_∞ consensus of estimates. *Automatica*, 47(1), 1–13.

Ugrinovskii, V. (2013). Distributed robust estimation over randomly switching networks using H_∞ consensus. *Automatica*, 49(1), 160–168.

Wang, G., Li, N., and Zhang, Y. (2017). Diffusion distributed Kalman filter over sensor networks without exchanging raw measurements. *Signal Processing*, 132, 1 – 7.

Yan, L., Zhang, X., and Lv, J. (2015). Consensus-based distributed estimators in sensor networks with sampled information and intermittent communication links. In *Chinese Control and Decision Conference*, 6057–6062. Chongqing, China.

Zhang, Y., Wang, C., Li, N., and Chambers, J. (2015). Diffusion Kalman filter based on local estimate exchanges. In *IEEE International Conference on Digital Signal Processing (DSP)*, 828–832. Singapore.