Distributed estimation design for LTI systems: a linear quadratic approach

Article *in* International Journal of Systems Science · October 2019

DOI: 10.1080/00207721.2019.1674407

Álvaro Rodríguez del Nozal
Universidad de Sevilla

Pablo Millan
Universidad Loyola Andalucía

Distributed Estimation Design for LTI systems: A Linear Quadratic Approach

A. Rodríguez del Nozal[a], L. Orihuela[b], and P. Millán[b]

[a]Department of Electrical Engineering, Universidad de Sevilla, Seville, Spain
[b]Departamento de Ingeniería, Universidad Loyola Andalucía, Sevilla, Spain

ARTICLE HISTORY
Compiled September 26, 2019

ABSTRACT
This paper deals with the problem of distributedly estimate the state of a plant through a network of interconnected agents. Each of these agents must perform a real-time monitoring of the plant state, counting on the measurements of local plant outputs and on the exchange of information with neighboring agents. The paper introduces a distributed LQ-based design that is applied to a distributed observer structure based on a multi-hop subspace decomposition. Stability and optimality conditions are derived and tested in simulation. Finally, the design method presented allows the user, through the tune of two scalar parameters, to modify the observer gains according to their experience about the plant.

KEYWORDS
Multi-agent systems; State Estimation; Distributed estimation; LTI-systems; Linera Quadratic

1. Introduction

Nowadays, the cheapening of electronic devices and the raising use of information technologies have triggered the implantation in process control of distributed architectures in place of the increasingly outdated centralized strategies. These architectures present a great amount of advantages with respect to the aforementioned centralized schemes such as scalability, flexibility, fault tolerance or the suitability for learning from large datasets, just to mention a few of them (Sayed, 2014). However, distributed topologies pose new challenges for the control community, such as communication delays, packet losses and the generation of distributed control/estimation algorithms without having knowledge of all the information handled by each device.

This paper deals with the problem of distributedly estimate the state of a plant using a set of interconnected agents. Each of these agents must perform a real-time monitoring of the plant state, counting on the measurements of local plant outputs and on the exchange of information with the rest of the network. Therefore, each agent is able to observe a portion of the state using only direct measurements, and, then, might ignore the contributions of other agents to ensure stability. On the other hand, the locally unobservable modes of the plant dynamics can be observed using information gathered from the communication with other

Research partially supported by grant TEC2016-80242-P by AEI/FEDER
Email: arnozal@us.es, dorihuela@uloyola.es, pmillan@uloyola.es

agents in the network. A recent literature survey about this topic can be found in (Li et al., 2015).

The literature in distributed estimation is broad and important contributions to this topic can be found. We will focus our literature review to the case of the distributed estimation of a plant by a network of agents that can design the observer in a distributed fashion. That is, a centralized design of the estimation strategy is not required. Maybe, one of the main contributions in this frame is the Distributed Kalman Filter (DKF). The strategy, first introduced in (Olfati-Saber, 2007) allows the estimator to minimize the uncertainties in the estimation using consensus among the agents. The design introduced minimizes the expected value of the estimation error. For an overview of technical details associated with consensus-based estimation strategies, reader is referred to (Garin & Schenato, 2010). This kind of distributed estimation has attracted the attention of many researchers that have presented succesive modifications of the DKF. To mention some results, let's focus on (Das & Moura, 2015), where an experimental evaluation of the sensitivity of the performance of the distributed estimation to the model parameters and noise statistics is studied. Reference (Kar & Moura, 2011) proposes a gossip Kalman filter, this is, a distributed estimator in which consensus and estimation appears at the same time scale. Another approach can be found in (Rodríguez del Nozal, Orihuela, & Millán, 2017), where the authors introduce a state decomposition in order to minimize the information exchanged during the estimation phase among the agents in the network. Nevertheless, the information needed in the design phase is huge which complicates the distributed design. In (Khan, Kar, Jadbabaie, & Moura, 2010) the concept of Network Tracking Capacity is introduced in order to characterize a class of dynamical system that a network and given observations can track with bounded error. All these approaches have the advantage of minimizing the uncertainties in the estimation at the expense of a high exchange of information through the network.

Another approach to the same problem is the Luenberger-based distributed observer, introduced in (Park & Martins, 2017) and (Rego, Aguiar, Pascoal, & Jones, 2017). In the former, necessary and sufficient conditions for the stability of the observer structure presented are determined. Nevertheless, the design method is proposed for the unperturbed plant, not considering the impact of disturbances and noises. The latter presents a distributed observer whose design can be done in a distributed way. However, the performance of the distributed observer relays on some parameters whose choice remains unclear. In the same line, two novel observer structures are presented in (Mitra & Sundaram, 2018) and (Kim, Shim, & Cho, 2016), where a matrix transformation is used in order to decouple the observable subspaces of each agent. Although the idea introduced is interesting, both structures present problems to design the observer in a distributed fashion under mild assumptions.

Fuzzy logic has been also applied to the distributed estimation problem. In (Qiu, Sun, Wang, & Gao, 2019) a fuzzy state observer is designed to estimate unmeasured states. Another interesting work that includes the basis for its application in estimation is (Sun, Mou, Qiu, Wang, & Gao, 2018) where the problem of adaptive fuzzy control is investigated for a class of non-triangular structural stochastic switched nonlinear systems with full state constraints.

Another observer design method consists in applying linear quadratic techniques to estimation problems. This strategy has been deeply studied and applied in control theory but rarely used to design observers. In this framework a few works can be pointed out. For instance, in (Zhang, Lewis, & Das, 2011) a cooperative tracking problem is proposed and a

LQR based optimal design approach is applied to design the control and estimation gains. In (Kishor, Singh, & Raghuvanshi, 2006) an LQ-based method is used to design a hydro turbine speed control estimator. Another approach can be found in (Orihuela, Gómez-Estern, & Rubio, 2014), where the authors relay on an observer design in which a quadratic function in terms of the estimation error is minimized. All this approaches have in common the use of LQ-based techniques to design centralized estimators. Nevertheless, authors are not aware of any paper dealing with the application of linear quadratic-based techniques to design distributed observers.

In distributed systems, all the devices involved in the problem work cooperatively or non-cooperatively to achieve an end. In this framework, there are many works that study distributed computation methods for multi-agent systems with the main aim of reducing the complexity of the problem and achieving shorter resolution rates than in the centralized scheme. For example, in (Nedic & Ozdaglar, 2009) a distributed computation model for optimizing a sum of convex objective functions corresponding to multiple agents is studied. In (Droge, Kawashima, & Egerstedt, 2014) the relationship between dual decomposition and the consensus-based method for distributed optimization is studied and compared. Another interesting work in this field is (Shi & Yang, 2018), where an augmented Lagrange algorithm for distributed optimization is proposed.

This is where the novelty of this study lies with respect to available literature, in the design of distributed observers using a LQ framework. In this paper a distributed approach is applied to estimate the state of the plant from a network of collaborative agents. The main contributions are listed below:

- The design of the observer gains, namely local Luenberger gain and consensus matrices, is tackled by minimizing a quadratic cost function. This minimization problem is proven to be solvable in a distributed way.
- By using linear programming, the optimality and stability of the observer is proven in a distributed framework.
- It is presented an empirical way to choose the weights of the cost function. Instead of designing the full matrices a task that is sometimes difficult for the practitioner, the method only requires to tune two scalar parameters. Simulations show that the method gets nice results when the practitioner has certain knowledge about the reliability of the model and the accuracy of the measurement.

The paper is organized as follows. In Section 2, the problem is formally stated, together with some definitions and assumptions. Section 3 presents the observer structure and some properties useful in the sequel sections. Section 4 constitutes the main result of the paper where a LQ-based method to design the observer is introduced. This section also presents stability and optimality results. In Section 5 some simulation examples that show the robustness of the observer are presented. Finally, conclusions are drawn in Section 6.

Notation 1.1. A graph is a pair $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ comprising a set $\mathscr{V} = \{1, 2, \ldots, p\}$ of *vertices* or *agents*, and a set $\mathscr{E} \subset \mathscr{V} \times \mathscr{V}$ of *edges* or *links*. A *directed graph* is a graph in which edges have orientations, so that if $(j,i) \in \mathscr{E}$, 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), …, (i_{k-1}, i_k) in a directed graph. The *neighborhood* of i, $\mathscr{N}_i \triangleq \{j : (j,i) \in \mathscr{E}\}$ is defined as the set of nodes with edges incoming to node i. Given $\rho \in \mathbb{Z} > 0$, the *ρ-hop reachable set* of i, $\mathscr{N}_{i,\rho}$, is defined as the set of nodes with a direct path to i involving ρ edges. Note that the 1-hop reachable set of i corresponds to the neighborhood of i. The operator $\mathrm{col}(\cdot,\cdot)$ stacks subsequent matrices into a column vector, e.g. for $A \in \mathbb{R}^{m_1 \times n}$ and $B \in \mathbb{R}^{m_2 \times n}$, $\text{col}(A,B) = [A^\top \ B^\top]^\top \in \mathbb{R}^{(m_1+m_2)\times n}$. $Im(A)$ denotes the image of matrix A, i.e., the subspace generated by the columns of matrix A. $\sigma(A)$ denotes the set of eigenvalues of matrix A. Let $||x||_\infty = \max\{|x_1|, \ldots, |x_p|\}$ be the infinity norm of vector $x = [x_1, \ldots, x_p]$. Let I_n denotes the identity matrix of dimension n. The evolution of a dynamical system $x^+ = Ax$ is Globally Uniformly Bounded (GUUB) with ultimate bound b if it exists positive constants b and c, independent of $t_0 \geq 0$, such that for every $0 < a < c$ arbitrarily large, there exist a $T = T(a,b) > 0$, independent of t_0, such that $||x(t_0)|| \leq a \Rightarrow ||x(t)|| \leq b$, for all $t \geq t_0 + T$.

2. Problem formulation

Consider a set of agents $\mathscr{V} = \{1, 2, ..., p\}$ connected through a communication network characterized by a directed graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$ where the vertices of the graph, $\mathscr{V}$, represents the agents, and the edges of the graph, $\mathscr{E}$, indicates the connection among them (see for instance Figure 1). The main aim of each agent $i \in \mathscr{V}$ is to distributedly estimate the state $x \in \mathbb{R}^n$ of the following LTI system:

$$\begin{aligned} x^+ &= Ax + w, \\ y_i &= C_i x + n_i, \quad \forall i \in \mathscr{V}, \end{aligned} \tag{1}$$

where $y_i \in \mathbb{R}^{m_i}$ is the output locally measured by each agent i at time k, $C_i \in \mathbb{R}^{m_i \times n}$ is the output matrix of agent i, and $w \in \mathbb{R}^n$ and $n_i \in \mathbb{R}^{m_i}$ are state and measurement noises at time k, respectively.

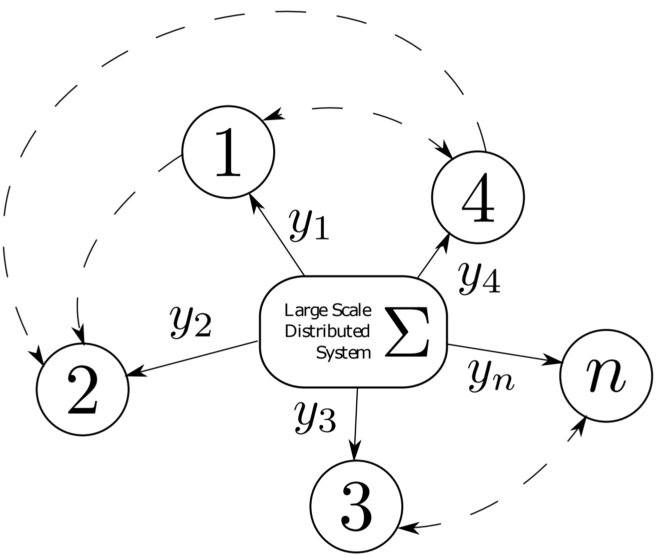

Figure 1. Distributed estimation problem scheme, where a set of agents take measurements of system Σ and exchange information among them through a communication network (in dashed lines).

The observation structure proposed in the next section relies on system transformations to the observability staircase form (see for instance Theorem 16.2 in (Hespanha, 2009)). Prior to introducing this structure, the following definitions are needed.

Definition 2.1. The ρ-hop output matrix of agent i, $C_{i,\rho}$, is a matrix that stacks the $(\rho-1)$-hop output matrix of agent i and the $(\rho-1)$-hop output matrices of its neighborhood, $\mathscr{N}_i$. That is:

$$C_{i,\rho} := \begin{bmatrix} C_{i,\rho-1} \\ \mathrm{col}(C_{j,\rho-1})_{j\in\mathscr{N}_i} \end{bmatrix}, \quad \forall\rho \geq 1,$$

where $C_{i,0} := C_i$.

Intuitively speaking, the ρ-hop output matrix of agent i, $C_{i,\rho}$, is composed by its output matrix C_i and the output matrices of all the agents j with a direct path to i involving ρ or less edges.

Definition 2.2. System (1) is locally detectable from agent i if pair (C_i,A) is detectable. System (1) is collectively detectable if for each agent $i \in \mathscr{V}$ there exists a finite number of hops $\ell_i \in Z > 0$ such that pair (C_{i,ℓ_i},A) is detectable.

Assumption 2.3. *We assume that system* (1) *is collectively detectable.*

Note that Assumption 2.3 is in general less restrictive than other approaches find in the literature, as it does not enforce connectivity of the network.

There always exists a coordinate transformation matrix $\begin{bmatrix}\bar{V}_{i,\rho} \; V_{i,\rho}\end{bmatrix} \in \mathbb{R}^{n\times n}$ according to pair $(C_{i,\rho},A)$ such that the change of variable $\xi_{i,\rho} \triangleq [\bar{V}_{i,\rho} \; V_{i,\rho}]^\top x \in \mathbb{R}^n$ transforms the original state-space representation into the observability staircase form. Note that $\bar{V}_{i,\rho} \in \mathbb{R}^{n\times n^{\bar{o}}_{i,\rho}}$ is composed by $n^{\bar{o}}_{i,\rho}$ column vectors in $\mathbb{R}^n$ that form an orthogonal basis of the unobservable subspace of pair $(C_{i,\rho},A)$. Correspondingly, $V_{i,\rho} \in \mathbb{R}^{n\times n^{o}_{i,\rho}}$ is an orthogonal basis of its orthogonal complement.

Definition 2.4. The ρ-hop unobservable subspace from agent i, denoted $\bar{\mathscr{O}}_{i,\rho}$, is composed of all system modes that cannot be observed from the output locally measured by agent i and those measured by all the agents belonging to the s-hop reachable set of i, $\forall s \in \{0,\dots,\rho\}$. Equivalently, the ρ-hop unobservable subspace from agent i is the unobservable subspace related to pair $(C_{i,\rho},A)$ using the above coordinate transformation:

$$\bar{\mathscr{O}}_{i,\rho} := Im(\bar{V}_{i,\rho}).$$

The orthogonal complement of $\bar{\mathscr{O}}_{i,\rho}$, with some abuse of notation, is denoted ρ-hop observable subspace from agent i, $\mathscr{O}_{i,\rho} := Im(V_{i,\rho})$. We denote $n^o_{i,\rho} = dim(\mathscr{O}_{i,\rho})$.

According to Definition 2.4, it is clear that:

$$\mathscr{O}_{i,\rho-1} \subseteq \mathscr{O}_{i,\rho}, \quad \forall i \in \mathscr{V}, \quad \rho \geq 0. \tag{2}$$

where we consider $\mathscr{O}_{i,-1} = \emptyset$. Then, the vectors of the "innovation" basis that generates $\mathscr{O}_{i,\rho} \cap (\mathscr{O}_{i,\rho-1})^\perp$ can be stacked into a matrix $W_{i,\rho} \in \mathbb{R}^{n\times n_{i,\rho}}$, where $n_{i,\rho} = n^o_{i,\rho} - n^o_{i,\rho-1}$, in such a way that:

$$Im(W_{i,\rho}) := \mathscr{O}_{i,\rho} \cap (\mathscr{O}_{i,\rho-1})^\perp, \quad \rho \geq 0, \tag{3}$$

Let us, to be selected later, define $\ell_i \in \mathbb{Z}_{>0}$ as an arbitrary number of hops. From these definitions it is clear that for all $\rho \in \{0,\dots,\ell_i\}$ and all $i \in \mathscr{V}$, it holds that

$$Im(V_{i,\rho}) = Im\left(\begin{bmatrix} W_{i,\rho} & V_{i,\rho-1} \end{bmatrix}\right), \tag{4}$$
$$Im(\bar{V}_{i,\rho-1}) = Im\left(\begin{bmatrix} W_{i,\rho} & \bar{V}_{i,\rho} \end{bmatrix}\right), \tag{5}$$

with $\bar{V}_{i,-1} := I_n$.

It is worth pointing out that $Im(W_{i,\rho})$ corresponds to the innovation introduced by the ρ-hop reachable set $\mathscr{N}_{i,\rho}$ of agent i, that is, the observable modes for agent i at hop ρ that are not observable at hop $\rho - 1$. Accordingly, the transformation matrix T_i, defined as $T_i = [\bar{V}_{i,\ell_i}\ V_{i,\ell_i}]$, can be divided using the innovations at each hop:

$$T_i := \big[\underbrace{\bar{V}_{i,\ell_i} \quad W_{i,\ell_i} \quad \cdots \quad W_{i,\rho+1}}_{\bar{V}_{i,\rho}} \quad \underbrace{W_{i,\rho} \quad \cdots \quad W_{i,0}}_{V_{i,\rho}}\big] \in \mathbb{R}^{n\times n}, \tag{6}$$

for all $\rho \in \{0,\dots,\ell_i\}$, where it is easy to identify the observable and unobservable subspaces of the system by agent i at hop ρ. Please, note that $\bar{V}_{i,\ell_i}$ represents the basis of the unobservable but detectable subspace of agent i.

The following lemma, previously presented in (Rodríguez del Nozal, Millán, Orihuela, Seuret, & Zaccarian, 2019), introduces some important properties that are central for the subsequent derivations.

Lemma 2.5. *For any agent $i \in \mathscr{V}$, the next properties hold, $\forall\rho,\ \rho' \in \{1,\dots,\ell_i\}$ such that $\rho \neq \rho'$:*

(i) $W_{i,\rho}^\top W_{i,\rho'} = 0$,
(ii) $Im(W_{j,\rho-1}) \subseteq Im(V_{i,\rho})$, $\forall j \in \mathscr{N}_i$,
(iii) $Im(W_{i,\rho}) \subseteq \bigoplus_{j\in\mathscr{N}_i} Im(W_{j,\rho-1})$,
(iv) $Im(A\bar{V}_{i,\rho}) \subseteq Im(\bar{V}_{i,\rho})$,

Once introduced the above lemma, the following proposition can be established.

Proposition 2.6. *For each agent i, the orthogonal similarity transformation given by T_i in* (6) *transforms the system matrix A into a block upper-triangular matrix in the form (Rodríguez del Nozal et al., 2019):*

$$T_i^\top A T_i = \begin{bmatrix} \bar{V}_{i,\ell_i}^\top A\bar{V}_{i,\ell_i} & \bar{V}_{i,\ell_i}^\top AW_{i,\ell_i} & \dots & \bar{V}_{i,\ell_i}^\top AW_{i,1} & \bar{V}_{i,\ell_i}^\top AW_{i,0} \\ 0 & W_{i,\ell_i}^\top AW_{i,\ell_i} & \dots & W_{i,\ell_i}^\top AW_{i,1} & W_{i,\ell_i}^\top AW_{i,0} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & W_{i,1}^\top AW_{i,1} & W_{i,1}^\top AW_{i,0} \\ 0 & 0 & \dots & 0 & W_{i,0}^\top AW_{i,0} \end{bmatrix} \tag{7}$$

3. Observer Structure

The structure of the proposed distributed observer is as follows:

$$\widehat{x}_i^+ = \underbrace{A\widehat{x}_i}_{(a)} + \underbrace{W_{i,0}L_i(y_i - \widehat{y}_i)}_{(b)} + \underbrace{\sum_{\rho=1}^{\ell_i}\sum_{j\in\mathcal{N}_i} W_{i,\rho}N_{i,j,\rho}W_{j,\rho-1}^\top(\widehat{x}_j - \widehat{x}_i)}_{(c)} \tag{8}$$

where $\hat{x}_i$ is the estimation of system state x by agent i and L_i and $N_{i,j,\rho}$ are, respectively, a local observer gain and consensus gains to be designed. It is worth pointing out the role of the three terms in (8):

(a) The first term, $A\hat{x}_i$, is the classical model-based open-loop prediction.
(b) The second term, containing $L_i(y_i - \widehat{y}_i)$, is a local Luenberger-like output injection term, intended to correct the previous prediction with the difference between the local measures and its predicted outputs $\widehat{y}_i := C_i\hat{x}_i$. This term is pre-multiplied by $W_{i,0}$, which implies that the elements in the correction vector $L_i(y_i - \hat{y}_i)$ are actually used as weights to perform linear combinations of the column vectors forming $W_{i,0}$, and therefore only affect the observable subspace of agent i. This makes full sense, as the locally available output y_i only contains information about this subspace.
(c) This last term aims at adjusting the estimates $\hat{x}_i$ with the information received by the neighboring agents. Thus, the differences between the estimates of i and j are multiplied by matrix $W_{j,\rho-1}^\top$. The result is multiplied by gain matrix $N_{i,j,\rho}$ and is used as weights to perform linear combinations of $W_{i,\rho}$.

For each agent $i \in \mathcal{V}$, let us define the estimation error as $e_i := x - \hat{x}_i$, and similarly, the transformed estimation error as $\varepsilon_i := \xi_i - \hat{\xi}_i = T_i^\top e_i$, which can be decomposed in the transformed estimation error of agent i at each hop ρ:

$$\varepsilon_i = \begin{bmatrix} \bar{\varepsilon}_{i,\ell_i} \\ \varepsilon_{i,\ell_i} \\ \vdots \\ \varepsilon_{i,1} \\ \varepsilon_{i,0} \end{bmatrix} = \begin{bmatrix} \bar{V}_{i,\ell_i}^\top \\ W_{i,\ell_i}^\top \\ \vdots \\ W_{i,1}^\top \\ W_{i,0}^\top \end{bmatrix} e_i = T_i^\top e_i, \tag{9}$$

and thus, due to the fact that T_i is an orthogonal matrix (and therefore $T_i^\top T_i = I_n$), the expression of the estimation error in $\varepsilon_{i,\rho}$ coordinates yields:

$$e_i = \bar{V}_{i,\ell_i}\bar{\varepsilon}_{i,\ell_i} + \sum_{r=0}^{\ell_i} W_{i,r}\varepsilon_{i,r}. \tag{10}$$

The following lemma, whose proof can be found in (Rodríguez del Nozal et al., 2019), will be used later on.

Lemma 3.1. *The next equation holds for any $i \in \mathcal{V}$, any $j \in \mathcal{N}_i$, and any $\rho \in \{0,\dots,\ell_i\}$*

$$W_{j,\rho-1}^\top(\hat{x}_j - \hat{x}_i) = W_{j,\rho-1}^\top\left(\sum_{r=0}^{\rho} W_{i,r}\varepsilon_{i,r} - W_{j,\rho-1}\varepsilon_{j,\rho-1}\right). \tag{11}$$

Proposition 3.2. *Consider the network of agents described by the graph $\mathscr{G} = (\mathscr{V}, \mathscr{E})$, where every agent i implements the observer structure* (8) *to estimate the state of the system* (1)*. Then, the transformed estimation error dynamics at every hop ρ is given by the following equations:*

$$\varepsilon_{i,0}^{+} = (W_{i,0}^{\top} A W_{i,0} - L_i C_i W_{i,0}) \varepsilon_{i,0} + W_{i,0}^{\top} w - L_i n_i, \tag{12}$$

$$\varepsilon_{i,\rho}^{+} = \sum_{r=0}^{\rho} \left(W_{i,\rho}^{\top} A - \sum_{j \in \mathscr{N}_i} N_{i,j,\rho} W_{j,\rho-1}^{\top} \right) W_{i,r} \varepsilon_{i,r} + W_{i,\rho}^{\top} w + \sum_{j \in \mathscr{N}_i} N_{i,j,\rho} \varepsilon_{j,\rho-1}, \tag{13}$$

for all $\rho = \{1, \ldots, \ell_i\}$.

Proof. Let us write first the evolution of the estimation error dynamics for system (1) under the observation structure in (8):

$$e_i^{+} = x^{+} - \hat{x}_i^{+} = Ae_i + w - W_{i,0} L_i C_i e_i - W_{i,0} L_i n_i - \sum_{\rho=1}^{\ell_i} \sum_{j \in \mathscr{N}_i} W_{i,\rho} N_{i,j,\rho} W_{j,\rho-1}^{\top} (\hat{x}_j - \hat{x}_i).$$

Using (9), we can write the transformed estimation error dynamics for agent i at hop 0:

$$\varepsilon_{i,0}^{+} = W_{i,0}^{\top} e_i = W_{i,0}^{\top} A e_i - L_i C_i e_i + W_{i,0}^{\top} w - L_i n_i,$$

where $W_{i,0}^{\top} W_{i,0} = I_{n_{i,0}}$ and Lemma 2.5 (i) has been used. Next, thanks to equation (10), we know the expression of e_i in $\varepsilon_{i,\rho}$ coordinates and using Lemma 2.5 (iv) it implies that $W_{i,0}^{\top} A \left(\bar{V}_{i,\ell_i} \bar{\varepsilon}_{i,\ell_i} + \sum_{r=1}^{\ell_i} W_{i,r} \varepsilon_{i,r} \right) = 0$ so we can rewrite the equation above as:

$$\varepsilon_{i,0}^{+} = W_{i,0}^{\top} A W_{i,0} \varepsilon_{i,0} - L_i C_i W_{i,0} \varepsilon_{i,0} + W_{i,0}^{\top} w - L_i n_i,$$

which is the desired equation exposed in (12).

The second part of the proof consists in obtaining expression (13). Let us write the transformed estimation error dynamics for agent i at hop ρ, with $\rho \geq 0$. Using the orthogonality in Lemma (2.5) (i):

$$\varepsilon_{i,\rho}^{+} = W_{i,\rho}^{\top} e_i^{+} = W_{i,\rho}^{\top} A e_i + W_{i,\rho}^{\top} w - \sum_{j \in \mathscr{N}_i} N_{i,j,\rho} W_{j,\rho-1}^{\top} (\hat{x}_j - \hat{x}_i),$$

where we have applied $W_{i,\rho}^{\top} W_{i,\rho} = I_{n_{i,\rho}}$. Next, anagolously as with $\rho = 0$, thanks to equation (10), we can substitute e_i in $\varepsilon_{i,\rho}$ coordinates and using Lemma (2.5) (iv) we know that $W_{i,\rho}^{\top} A \left(\bar{V}_{i,\ell_i} \bar{\varepsilon}_{i,\ell_i} + \sum_{r=\rho+1}^{\ell_i} W_{i,r} \varepsilon_{i,r} \right) = 0$. Hence, the above equation yields:

$$\varepsilon_{i,\rho}^{+} = W_{i,\rho}^{\top} A \sum_{r=0}^{\rho} W_{i,r} + W_{i,\rho}^{\top} w - \sum_{j \in \mathscr{N}_i} N_{i,j,\rho} W_{j,\rho-1}^{\top} (\hat{x}_j - \hat{x}_i).$$

Finally, applying Lemma 3.1, equation (13) is obtained, completing the proof. □

4. LQ based observer design

This section presents an LQ-based design for the observers in (8). It is first shown that the proposed design guarantees optimality and asymptotic convergence in the absence of plant and measurement noises. Then, it is demonstrated that Assumption 2.3 suffices to ensure the feasibility of the proposed design. After that, it is shown that the dynamics of the estimation error is Globally Uniformly Ultimately Bounded (GUUB) in the presence of noises. Finally, a tuning method is proposed to choose the weights of the cost functions associated to the LQ-design.

4.1. Design of the distributed observer

Let us consider the following local quadratic cost function:

$$J_i(k) = \sum_{\rho=0}^{\ell_i} \sum_{t=k}^{\infty} \Big(\varepsilon_{i,\rho}(t)^\top U_{i,\rho} \varepsilon_{i,\rho}(t) + u_{i,\rho}(t)^\top S_{i,\rho} u_{i,\rho}(t) \Big), \tag{14}$$

where

$$u_{i,0}(t) = -L_i C_i W_{i,0} \varepsilon_{i,0}(t), \tag{15}$$

$$u_{i,\rho}(t) = -\sum_{j \in N_i} N_{i,j,\rho} W_{j,\rho-1}{}^\top W_{i,\rho} \varepsilon_{i,\rho}(t), \quad \rho = \{1, \ldots, \ell_i\}, \tag{16}$$

and $U_{i,\rho} \in \mathbb{R}^{n_{i,\rho} \times n_{i,\rho}}$ and $S_{i,\rho} \in \mathbb{R}^{n_{i,\rho} \times n_{i,\rho}}$ are diagonal positive definite weighting matrices.

The term $\varepsilon_{i,\rho}(t)^\top U_{i,\rho} \varepsilon_{i,\rho}(t)$ in (14) is the stabilization cost of the estimation error, computed for every hop ρ. By analogy with the classical cost functions in LQ control problems, this term is inspired by the term $x^\top Qx$ that weights the deviation of the system state/estimation error from the reference. The purpose of term $u_{i,\rho}(t)^\top S_{i,\rho} u_{i,\rho}(t)$ is to weight, on the one hand, the information feedback at hop 0, which involves only the local measured plant output and the corresponding gain L_i, and on the other hand, the information feedback at further hops, involving neighbors estimates and consensus matrices $N_{i,j,\rho}$. Using the same analogy, it weights the influence of the feedback signal as the term $u^\top Ru$, typically used in LQ control.

Remark 4.1. The relation between weighting matrices $U_{i,\rho}$ and $S_{i,\rho}$ has a direct effect in the values of the observer gains L_i and $N_{i,j,\rho}$, and, consequently, in the performance of the observer. Higher values in $U_{i,\rho}$ implies less confidence in the system model, obtaining as a result more aggressive observer gains. Conversely, higher values in $S_{i,\rho}$ implies less confidence in the agent's measurements providing a design in which the system model have more influence in the state estimation than the measurements taken.

Property 4.2. *For every agent $i \in \mathscr{V}$, the estimation gains L_i and $N_{i,j,\rho}$ are designed in such a way that for all $\rho \in \{0, \ldots, \ell_i\}$:*

$$(S_{i,0} + P_{i,0})^{-1} P_{i,0} W_{i,0}^\top A W_{i,0} \varepsilon_{i,0} \varepsilon_{i,0}^\top W_{i,0}^\top C_i = L_i C_i W_{i,0} \varepsilon_{i,0} \varepsilon_{i,0}^\top W_{i,0}^\top C_i^\top, \tag{17}$$

$$\big(S_{i,\rho} + P_{i,\rho}\big)^{-1} P_{i,\rho} W_{i,\rho}^\top A W_{i,\rho} = \sum_{j \in N_i} N_{i,j,\rho} W_{j,\rho-1}{}^\top W_{i,\rho}, \ \forall \rho \in \{1, \ldots, \ell_i\} \tag{18}$$

where $P_{i,\rho} \in \mathbb{R}^{n_{i,\rho} \times n_{i,\rho}}$ are positive definite matrices solution of the following Discrete-time Algebraic Riccati Equation (DARE):

$$W_{i,\rho}^\top A^\top W_{i,\rho} P_{i,\rho} W_{i,\rho}^\top A W_{i,\rho} - P_{i,\rho} + U_{i,\rho} = W_{i,\rho}^\top A^\top W_{i,\rho} P_{i,\rho} \left(S_{i,\rho} + P_{i,\rho}\right)^{-1} P_{i,\rho} W_{i,\rho}^\top A W_{i,\rho}, \quad (19)$$

for all $\rho \in \{0, \dots, \ell_i\}$.

Next, the computational complexity of the design proposed in analyzed. In order to do that, let us define a *flop* as the amount of work associated with a floating-point add and multiplication. There exists several methods to solve DAREs. We will consider the one introduced in (Lin-Zhang & Wen-Wei, 1993, Algorithm 4.1B) that quantifies in $72n_d$ flops the computational cost of solving a DARE in which the searched variable has dimension $n_d \times n_d$.

As it can be seen, the observer design introduced in Property 4.2 implies the resolution of $\ell_i + 1$ discrete-time algebraic Riccati equations for each agent $i \in \mathcal{V}$. Note that the dimension of the searched variable in each equation is given by the size of the "innovation" at each hop ρ, that is, $n_{i,\rho}$. Next, based on equation (6) it is easy to see that $\sum_{\rho=0}^{\ell_i} n_{i,\rho} = n$. Thus, the computational complexity of the observer design requires less than $72n$ flops.

Based on the above property, we can now state the main result of the paper.

Theorem 4.3. *Consider system* (1) *in the absence of plant and measurements noises, and the observation structure defined in* (8)*. Then,*

(1) If all the estimation errors $\varepsilon_{i,r}$ converge to zero for $0 \le r < \rho$, $\forall i \in \mathcal{V}$, the gain matrices L_i and $N_{i,j,\rho}$ that minimize the cost function (14) *at hop ρ are given by Property 4.2.*
(2) If the estimations gains L_i and $N_{i,j,\rho}$ for every $\rho \in \{0, \dots, \ell_i\}$ are designed satisfying Property 4.2, then the estimates of all the agents tend asymptotically to the actual plant state.

Proof. First, it will be shown that, provided that all the estimation errors $\varepsilon_{i,r}$ converge to zero for $0 \le r < \rho$, $\forall i \in \mathcal{V}$, the optimal design of the estimation gains are given by (17)-(19). After that, the asymptotic stability will be proven by induction.

For the first part, let us write the dynamics of $\varepsilon_{i,\rho}$ according to (13) in the absence of noises and with $\varepsilon_{i,r} \equiv 0$, $\forall r : 0 \le r < \rho$.

$$\varepsilon_{i,\rho}^+ = W_{i,\rho}^\top A W_{i,\rho} \varepsilon_{i,\rho} - \sum_{j \in \mathcal{N}_i} N_{i,j,\rho} W_{j,\rho-1}^\top W_{i,\rho} \varepsilon_{i,\rho} = W_{i,\rho}^\top A W_{i,\rho} \varepsilon_{i,\rho} + u_{i,\rho}, \quad (20)$$

where (16) has been used.

Now, let us write the cost function in (14) as $J_i(k) = \sum_{\rho=0}^{\ell_i} J_{i,\rho}(k)$, with $J_{i,\rho}(k) = \sum_{t=k}^{\infty} \Big(\varepsilon_{i,\rho}(t)^\top U_{i,\rho} \varepsilon_{i,\rho}(t) + u_{i,\rho}(t)^\top S_{i,\rho} u_{i,\rho}(t)\Big)$.

Given the quadratic dependence of $\varepsilon_{i,\rho}(k)$, it is clear that the values $u_{i,\rho}(t)$ that minimize $J_{i,\rho}(k)$ are linear functions of $\varepsilon_{i,\rho}(k)$, and therefore it is possible to write the optimal costs of each agent as:

$$J_{i,\rho}{}^{*}(k) = \varepsilon_{i,\rho}(k)^{\top} P_{i,\rho} \varepsilon_{i,\rho}(k), \tag{21}$$

for some $P_{i,\rho} \in \mathbb{R}^{n_{i,\rho} \times n_{i,\rho}} > 0$. Furthermore:

$$J_{i,\rho}(k) = \kappa_{i,\rho}(\varepsilon_{i,\rho}, u_{i,\rho}, k) + \sum_{t=k+1}^{\infty} \kappa_{i,\rho}(\varepsilon_{i,\rho}, u_{i,\rho}, t) = \kappa_{i,\rho}(\varepsilon_{i,\rho}, u_{i,\rho}, k) + J_{i,\rho}(k+1), \tag{22}$$

where $\kappa_{i,\rho}(\varepsilon_{i,\rho}, u_{i,\rho}, k) := \varepsilon_{i,\rho}(k)^{\top} U_{i,\rho} \varepsilon_{i,\rho}(k) + u_{i,\rho}(k)^{\top} S_{i,\rho} u_{i,\rho}(k)$.

To compute the optimal $u_{i,\rho}^{*}(k)$ for which the minimum cost $J_{i,\rho}^{*}(k)$ is attained, let us rewrite the equation above using (21), which results in:

$$\begin{aligned} J_{i,\rho}(k) &= \varepsilon_{i,\rho}(k)^{\top} U_{i,\rho} \varepsilon_{i,\rho}(k) + u_{i,\rho}(k)^{\top} S_{i,\rho} u_{i,\rho}(k) + \varepsilon_{i,\rho}(k+1)^{\top} P_{i,\rho} \varepsilon_{i,\rho}(k+1) \\ &= \varepsilon_{i,\rho}(k)^{\top} U_{i,\rho} \varepsilon_{i,\rho}(k) + u_{i,\rho}(k)^{\top} S_{i,\rho} u_{i,\rho}(k) \\ &+ (\varepsilon_{i,\rho}(k)^{\top} W_{i,\rho}^{\top} A^{\top} W_{i,\rho} + u_{i,\rho}(k)^{\top}) P_{i,\rho} (W_{i,\rho}^{\top} A W_{i,\rho} \varepsilon_{i,\rho}(k) + u_{i,\rho}(k)). \end{aligned} \tag{23}$$

Thus $u_{i,\rho}^{*}(k) = \arg \min\limits_{u_{i,\rho}(k)} J_{i,\rho}(k) = u_{i,\rho}(k) : \dfrac{\partial J_{i,\rho}(k)}{\partial u_{i,\rho}(k)} = 0$, which yields to:

$$u_{i,\rho}^{*}(k) = -\left(S_{i,\rho} + P_{i,\rho}\right)^{-1} P_{i,\rho} W_{i,\rho}^{\top} A W_{i,\rho} \varepsilon_{i,\rho}(k). \tag{24}$$

Substituting (24) in (23), it is straightforward to obtain (19), from which $P_{i,\rho}$ can be computed. Then, by comparing (24) and (16), it is clear that the observer gains must be designed according to equation (18).

Now let us move to the second claim of the theorem. First of all, we assume that all the estimation errors $\varepsilon_{i,r}$ converge to zero for $0 \leq r < \rho$, $\forall i \in \mathscr{V}$, and this lead us to prove the stability of $\varepsilon_{i,\rho}$. Later, it will be proven the convergence of $\varepsilon_{i,0}$.

To show the stabilization of the error $\varepsilon_{i,\rho}$, consider that from (22) it is directly obtained that $J_{i,\rho}(k+1) - J_{i,\rho}(k) = -\kappa_{i,\rho}(\varepsilon_{i,\rho}, u_{i,\rho}, k)$. Thus, taking as a Lyapunov function $V_{i,\rho}(k) = J_{i,\rho}^{*}(k) = \varepsilon_{i,\rho}(k)^{\top} P_{i,\rho} \varepsilon_{i,\rho}(k)$, it holds that $\Delta V_{i,\rho}(k) = V_{i,\rho}(k+1) - V_{i,\rho}(k) = -\kappa_{i,\rho}(\varepsilon_{i,\rho}, u_{i,\rho}, k) < 0$, which ensures the asymptotic convergence of $\varepsilon_{i,\rho}$ to the origin in absence of noise.

Finally, it suffices to show that in the absence of noises the proposed design guarantees the stabilization of $\varepsilon_{i,\rho}$ for $\rho = 0$. From (12), it yields that $\varepsilon_{i,0}^{+} = (W_{i,0}^{\top} A W_{i,0} - L_i C_i W_{i,0}) \varepsilon_{i,0}$. Repeating the same procedure above, but this time computing the optimal L_i that minimizes $J_{i,0}(k)$ the following expression is obtained:

$$(S_{i,0} + P_{i,0})^{-1} P_{i,0} W_{i,0}^{\top} A W_{i,0} \varepsilon_{i,0} \varepsilon_{i,0}^{\top} W_{i,0}^{\top} C_i = L_i C_i W_{i,0} \varepsilon_{i,0} \varepsilon_{i,0}^{\top} W_{i,0}^{\top} C_i^{\top},$$

where $P_{i,0}$ can be also computed from (19).

Then, comparing previous equation and (15), it is clear that the gains L_i must be obtained according to (17), and the Lyapunov function $V_{i,0}(k) = J_{i,0}^{*}(k) = \varepsilon_{i,0}(k)^{\top} P_{i,0} \varepsilon_{i,0}(k)$ ensures the asymptotic convergence of $\varepsilon_{i,0}$. $\square$

Note that matrices $U_{i,\rho}$ weight the knowledge on the dynamics of the system whereas matrices $S_{i,\rho}$ weight the accuracy of the information provided by the local measurements of the system (when $\rho = 0$) and the information provided by the neighborhood (when $\rho > 0$). Thus, if the measurements are highly affected by noise, it will be reflected in $u_{i,0}$ and the weighting matrix $S_{i,\rho}$ must be designed consequently in order to not amplify the effect of the noises.

It is worth pointing out that the conditions established in Theorem 4.3 are completely local. That is, since matrices $W_{j,\rho-1}$ must be computed once in a initialization phase, no more information is required from the neighboring agents in order to design the observer gains.

4.2. Design feasibility

The existence of a matrix $P_{i,\rho}$ solution of the Algebraic Riccati equation stated in (19) is straightforward from the controllability of pair $(I_{n_{i,\rho}}, W_{i,\rho}^\top A W_{i,\rho})$ (see for instance (Arnold & Laub, 1984)). It is a simple matter to check that regarding Popov-Belevitch-Hautus test (see, e.g., (Hespanha, 2009, Th.15.9)) the controllability is guaranteed due to the full rank of I_n matrix. Nevertheless, it left to prove the existence of gain matrices L_i and $N_{i,j,\rho}$ that fulfill expressions (17)-(18).

Theorem 4.4. *It is always possible, under Assumption 1, to find a set of matrices L_i and $N_{i,j,\rho}$ satisfying equations* (17)-(18).

Proof. Let us transpose expression (17) in order to obtain several systems of linear equations with the structure $Ax = b$ where the coefficient matrix A is $C_i W_{i,0} \varepsilon_{i,0} \varepsilon_{i,0}^\top W_{i,0}^\top C_i^\top$, the searched variable vector x is given by the row vectors of L_i and the matrix b corresponds to the row vectors of $(S_{i,0} + P_{i,0})^{-1} P_{i,0} W_{i,0}^\top A W_{i,0} \varepsilon_{i,0} \varepsilon_{i,0}^\top W_{i,0}^\top C_i$. Next, according to the Rouché-Capelli Theorem, the previous systems of equations are consistent if and only if the coefficient matrix A has full rank. Recall that this matrix is common for all the systems.

Now, from Definition 2.4 we have that the 0-hop observable subspace of agent i, $\mathscr{O}_{i,0} = Im(W_{i,0})$, is generated from pair (C_i, A) which directly implies that $m_i \leq n_{i,0}$. Consequently, it is clear that $C_i W_{i,0}$ is a full-rank matrix of dimension $m_i \times n_{i,0}$. Next, assuming the existence of a estimation error in the observer design phase, matrix $C_i W_{i,0} \varepsilon_{i,0}$ is a full-rank matrix. Finally, since the coefficient matrix A is the Grammian matrix of $C_i W_{i,0} \varepsilon_{i,0}$, both ranks are equal, that is, both matrices are full-rank matrices.

Concerning gains $N_{i,j,\rho}$, equation (18) can be rewritten as $(S_{i,\rho} + P_{i,\rho})^{-1} P_{i,\rho} W_{i,\rho}^\top A W_{i,\rho} = \bar{N}_{i,\rho} \Lambda_{i,\rho} W_{i,\rho}$, where $\Lambda_{i,\rho} = \mathrm{col}(W_{j,\rho-1}^\top)_{j \in \mathscr{N}_i}$ and $\bar{N}_{i,\rho} = \mathrm{col}(N_{i,j,\rho}^\top)_{j \in \mathscr{N}_i}^\top$. Next, by analogy to the L_i case, it is possible to transpose previous equation to obtain several linear systems of equations with the structure $Ax = b$, where the coefficient matrix A is $W_{i,\rho}^\top \Lambda_{i,\rho}^\top$ for all the systems. Finally, according to the properties of the innovation matrices and considering Lemma 2.5 (iii), it is clear that this is a full-rank matrix and therefore, by applying the Rouché-Capelli Theorem it is easy to see that all the systems of equations are consistent. $\square$

Remark 4.5. The design of local gains L_i established in Property 4.2 depends on the transformed estimation error of agent i at hop zero. However, when $m_i = n_{i,0}$ the equation can be simplified removing theses terms:

$$(S_{i,0}+P_{i,0})^{-1}P_{i,0}W_{i,0}^{\top}AW_{i,0}=L_iC_iW_{i,0}.$$

In this scenario, since $C_iW_{i,0}$ is a full rank square matrix and according to the Rouché-Capelli Theorem, the consistence of the equation is immediate.

4.3. Stability analysis for the perturbed scenario

When norm-bounded disturbances noises are affecting the system and measurements, it is well-known that exponential stability can no longer be guaranteed. In the following result, it is stated that when the perturbed scenario is considered, the estimation error is globally uniformly ultimately bounded and, additionally, the bound is directly modulated with the energy of the exogenous signals.

Assumption 4.6. *The norms of the noises $w(k)$ and $n_i(k)$ are upper-bounded as follows:*

$$||w(k)||<\delta_w,\ \ ||n_i(k)||<\delta_{n_i},\ \forall k,i\in\mathscr{V},$$

where $||\cdot||$ is a consistent norm and $\delta_w,\delta_{n_i}\in\mathbb{R}^+$ are the bounds.

Theorem 4.7. *Consider plant* (1) *observed by a set of agents that implements observation structure in* (8)*. Then, under Assumption 4.6, if the observer gains are designed following Property 4.2, then the estimation error of the system is Globally Uniformly Ultimately Bounded (GUUB), i.e., the estimates are attracted and restricted to lay within a small region around the plant state.*

Proof. Consider the following Lyapunov function for the transformed estimation error of agent i at hop ρ:

$$V_{i,\rho}(k)=\varepsilon_{i,\rho}(k)^{\top}P_{i,\rho}\varepsilon_{i,\rho}(k),$$

with $P_{i,\rho}$ obtained from Theorem 1. The increment of the function is given by

$$\Delta V_{i,\rho}(k)=\varepsilon_{i,\rho}(k+1)^{\top}P_{i,\rho}\varepsilon_{i,\rho}(k+1)-\varepsilon_{i,\rho}(k)^{\top}P_{i,\rho}\varepsilon_{i,\rho}(k). \tag{25}$$

Using the dynamics of the transformed observation error when $\rho=0$ given in (12), it turns out:

$$\begin{aligned}\Delta V_{i,0}(k) &= \varepsilon_{i,0}(k)^{\top}(E_{i,0}^{\top}P_{i,0}E_{i,0}-P_{i,0})\varepsilon_{i,0}(k)+w(k)^{\top}W_{i,0}P_{i,0}W_{i,0}^{\top}w(k)\\ &+ n_i^{\top}(k)L_i^{\top}P_{i,0}L_in_i(k)+2\varepsilon_{i,0}(k)^{\top}E_{i,0}^{\top}P_{i,0}\left(W_{i,0}^{\top}w(k)+L_in_i(k)\right),\\ &+ 2n_i^{\top}(k)L_i^{\top}P_{i,0}W_{i,0}^{\top}w(k)\end{aligned} \tag{26}$$

where for simplicity in the notation we have denoted $E_{i,0}\triangleq W_{i,0}^{\top}AW_{i,0}-L_iC_iW_{i,0}$. From Theorem 4.3 we know that

$$\varepsilon_{i,0}(k)^{\top}(E_{i,0}^{\top}P_{i,0}E_{i,0}-P_{i,0})\varepsilon_{i,0}(k)=\varepsilon_{i,0}(k)^{\top}\tilde{\kappa}_{i,0}\varepsilon_{i,0}(k)=-\kappa_{i,0}(\varepsilon_{i,0},u_{i,0},k).$$

Using a consistent norm, equation (26) can be bounded as

$$\begin{aligned}\Delta V_{i,0}(k) &\leq -\lambda_{min}(\tilde{\kappa}_{i,0})||\varepsilon_{i,0}(k)||^2 + ||W_{i,0}P_{i,0}W_{i,0}|| \cdot ||w(k)||^2 \\ &+ ||L_i^\top P_{i,0}L_i|| \cdot ||n_i(k)||^2 + 2||E_{i,0}^\top P_{i,0}W_{i,0}^\top|| \cdot ||\varepsilon_{i,0}(k)|| \cdot ||w(k)|| \\ &+ 2||E_{i,0}^\top P_{i,0}L_i|| \cdot ||\varepsilon_{i,0}(k)|| \cdot ||n_i(k)|| + 2||L_i^\top P_{i,0}W_{i,0}^\top|| \cdot ||n_i(k)|| \cdot ||w(k)||.\end{aligned}$$

The right side of this inequality is an algebraic second-order equation in $||\varepsilon_{i,\rho}(k)||$. Thus, if we impose that the right side of the equation is equal to zero, then $\Delta V_{i,\rho}(k) < 0$:

$$a||\varepsilon_{i,0}(k)||^2 + b||\varepsilon_{i,0}(k)|| + c = 0, \tag{27}$$

where

$$\begin{aligned} a &= -\lambda_{min}(\tilde{\kappa}_{i,0}), \\ b &= 2||E_{i,0}^\top P_{i,0}W_{i,0}^\top|| \cdot ||w(k)|| + 2||E_{i,0}^\top P_{i,0}L_i|| \cdot ||n_i(k)||, \\ c &= ||W_{i,0}P_{i,0}W_{i,0}^\top|| \cdot ||w(k)||^2 + ||L_i^\top P_{i,0}L_i|| \cdot ||n_i(k)||^2 \\ &\quad +2||L_i^\top P_{i,0}W_{i,0}^\top|| \cdot ||n_i(k)|| \cdot ||w(k)||, \end{aligned}$$

and then, the unique positive root of the equation is given by

$$||\varepsilon_{i,0}(k)|| = f_{i,0}(||w(k)||, ||n_i(k)||),$$

where $f_{i,0}$ is a function that solves the second-order equation (27). Thus, if we consider the extreme values of noise parameters, under Assumption 4.6, it is clear that $||\varepsilon_{i,0}(k)|| = f_{i,0}(\delta_w, \delta_{n_i})$ takes a finite value.

Suppose any initial condition for the estimation error, in such a way that $||\varepsilon_{i,0}(k_0)|| < \infty$. Let us denote $\alpha_{i,0} := f_{i,0}(\delta_w, \delta_{n_i})$. Assume that $||\varepsilon_{i,0}(0)|| > \alpha_{i,0}$. In this case the Lyapunov function decreases and this will imply $||\varepsilon_{i,0}(k)|| < \alpha_{i,0}$ for some $k > k_0$. Since the one step evolution of the estimation error $\varepsilon_{i,0}(k+1)$ in (12) is bounded provided that $||\varepsilon_{i,0}(k)|| < \alpha_{i,0}$, this finally proves that there exist a finite bound independent of time and initial conditions for $||\varepsilon_{i,0}(k)||$ for all $k > k_0$.

From (13), it can be seen that the evolution of the estimation error of agent i at hop ρ depends on the estimation error of that agent at the previous hops and the estimation error of the neighborhoods at hop $\rho - 1$, thus revealing a cascade structure. Hence, if we apply the same procedure recursively from $\rho = 0$ to $\rho = \ell_i$, we can reach to an algebraic second-order equation in $||\varepsilon_{i,\rho}(k)||$ whose coefficients depend on the solution of the second-order equations in the transformed estimation error of agent i and its neighborhood $j \in \mathcal{N}_i$ at the previous hop. It is clear that the solution of this equation is finite completing the proof. □

Please, note that the bounds of the local transformed estimation error norm, $||\varepsilon_{i,0}||$, depends on the amplitude of the noise terms w and n_i, and on the gain matrix L_i. Additionally, due to the cascade structure reveled in (13), the bound of the transformed estimation error at hop ρ, $||\varepsilon_{i,\rho}||$, depends on the amplitude of noise term w, on the gain matrices $N_{i,j,\rho}$ for every $j \in \mathcal{N}_i$ and on the transformed estimation error norm bounds at previous hops, $||\varepsilon_{i,s}||$ with $s \in \{0, \ldots, \rho - 1\}$.

Corollary 4.8. *If the measurements of the agents are not affected by noise, the infinity norm of the local transformed estimation error is decreasing as long as*

$$||\varepsilon_{i,0}(k)|| > \mu_{i,0}\delta_w,$$

where

$$\mu_{i,0} = \frac{||E_{i,0}^{\top}P_{i,0}W_{i,0}^{\top}|| + \sqrt{\left(||E_{i,0}^{\top}P_{i,0}W_{i,0}^{\top}||^2 + \lambda_{min}(\kappa_{i,0})||W_{i,0}P_{i,0}W_{i,0}^{\top}||\right)}}{\lambda_{min}(\tilde{\kappa}_{i,0})}.$$

Proof. The proof is based on solving (27) when $||n_i(k)|| = 0$. □

4.4. Tuning procedure proposed

Selecting the weighting matrices to design the distributed observer is usually a complicated task. The complication is even greater if the design must be carried out by a person that ignore the theory under the algorithm. Thus, this section presents a general method to design the weighting matrices $U_{i,\rho}$ and $S_{i,\rho}$. The aim of the method is to design the weighting matrices to each agent $i \in \mathcal{V}$ according to the distance between agent i and the agent which constitutes the source of the information used by i to reconstruct the unobservable subspace. Additionally, two scalar parameters (γ_i and λ_i) are introduced in order to allow the user to change the proportionality between matrices $U_{i,\rho}$ and $S_{i,\rho}$ regarding their experience with the process. Thus, this paper proposes the following values for the weighting matrices:

$$\begin{aligned} U_{i,0} &= \gamma_i I_{n_{i,0}}, && (28)\\ U_{i,\rho} &= \lambda_i I_{n_{i,\rho}}, \quad \forall\rho \in \{1,\dots,\ell_i\}, && (29)\\ S_{i,\rho} &= 10^{\rho+1} I_{n_{i,\rho}} \quad \forall\rho \in \{1,\dots,\ell_i\}, && (30)\end{aligned}$$

where $\gamma_i \in \mathbb{R}$ and $\lambda_i \in \mathbb{R}$. Thus, the design of the observer has been reduced to a problem in which it is only necessary to fix the value of two scalars. Note that meanwhile γ_i weights the confidence in the agents' measurements, λ_i weights the influence of the neighboring information.

This design method trades off the reliability of the model and the accuracy of the measurements according to the value of γ_i and λ_i. Thus, under Assumption 2.3 and Assumption 4.6, the method has the advantage that can be tuned based on the knowledge on the process, preserving the stability in the estimation for any value of γ_i and λ_i.

If weighting matrices $U_{i,\rho}$ and $S_{i,\rho}$ are designed following equations (28)-(30), $\gamma_i = 1$ and $\lambda_i = 1$ for all $i \in \mathcal{V}$, the matrices are chosen in order to weight the relative distance to the agent. Note that the design proposed in (28)-(30) gives a higher value to matrix $S_{i,\rho}$ as long as the distance between agent i and the agent which constitutes the source of the information increments. In this way, a more aggressive feedback signal is imposed to the local corrections and it is becoming softer as ρ arises.

5. Simulation results

In order to show the robustness of the distributed design of the observer some simulation examples are driven in this section. Consider the following system where there is one state with a stable dynamics, a pair of conjugated imaginary poles and a state with an unstable dynamic:

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}^+ = \begin{bmatrix} 0.95 & 0 & 0 & 0 \\ 0 & 0.8606 & -1.3368 & 0 \\ 0 & 0.0941 & 0.9315 & 0 \\ 0 & 0 & 0 & 1.015 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix},$$

and is observed by a set of four agents ($y_1 = x_1$, $y_2 = x_2$, $y_3 = x_3$, $y_4 = x_4$) with the network topology defined in Figure 2.

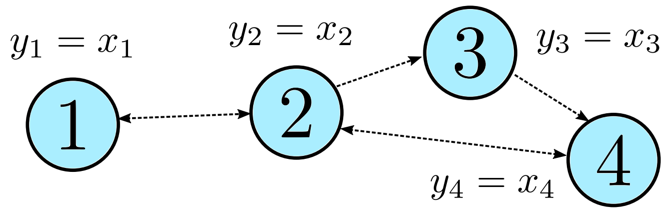

Figure 2. Network topology considered.

The basis vectors of the observable subspace for each agent can be easily obtained as:

$$W_{1,0} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad W_{2,0} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad W_{3,0} = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad W_{4,0} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

Note that agents 2 and 3 have the same observable subspace and, therefore, they will estimate states x_2 and x_3 based only on their local measurements.

Example 5.1. In this example the performance of the estimation is shown. According to Assumption 4.6, consider that the infinity norm of the noise terms $w(k)$ and $n_i(k)$ for all $i \in \mathscr{V}$ and for all time k are upper-bounded and the bounds are given by:

$$\delta_w = 0, \quad \delta_{n_1} = 0.8, \quad \delta_{n_2} = 0.9, \quad \delta_{n_3} = 0.7, \quad \delta_{n_4} = 0.6.$$

Consider a value of $\gamma_i = 1$ and $\lambda_i = 1$ for all $i \in \mathscr{V}$. In Figure 3 the evolution of the estimation error for agent 4, e_4, is shown. It is worth pointing out that the estimation error of state x_4, which belongs to the observable subspace of the agent, decreases drastically achieving a short convergence time. Errors $e_4(2)$ and $e_4(3)$, that according to the graph correspond to the innovation introduced at hop $\rho = 1$, start decreasing when $e_4(1)$ reaches the steady state. Lastly, state x_1, has a convergence rate slower than the others due to the fact that this state belongs to the innovation at $\rho = 2$. Thus, it is easy to see the cascade structure of the observer.

Figure 4 depicts the evolution of maximum value of the $||e_i(k)||_\infty$ for every agent $i \in \mathscr{V}$ and for the different observers modifying the value of γ_i and λ_i. Note that, for high values of these

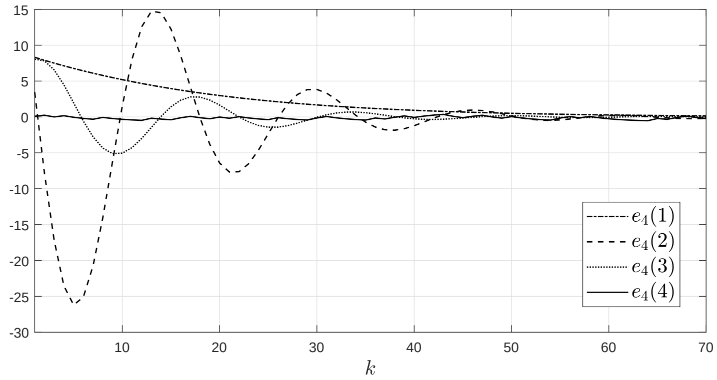

Figure 3. Estimation error evolution for agent 4.

parameters for all $i \in \mathscr{V}$, the feedback action is more aggressive achieving a lower settling time than when the value of them increase. However, the noise rejection for low values of γ_i and λ_i work better in the steady state. Recall that, for $\gamma_i = \lambda_i = 10$, the observers trade off between a good convergence rate and a good noise rejection in steady state, achieving a good performance according to both parameters.

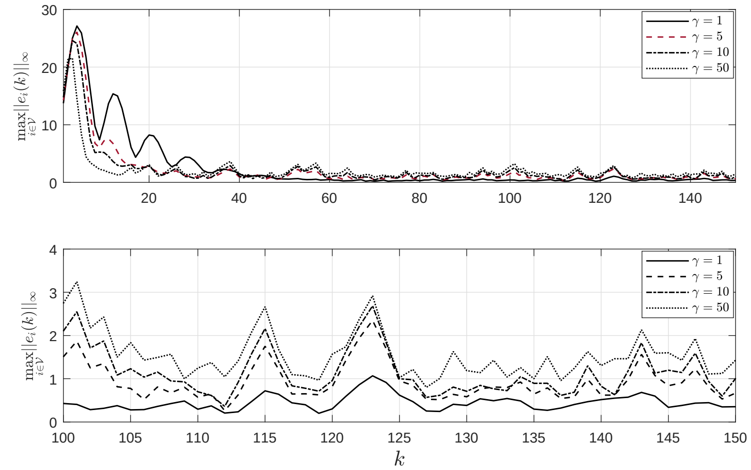

Figure 4. Evolution of the maximum value of the $||e_i(k)||_\infty$ for every agent $i \in \mathscr{V}$ according to γ_i.

Example 5.2. In this example we will consider the same system and network topology than in Example 5.1 but this time, let us consider the sequel noise bounds:

$$\delta_w = 0.6, \quad \delta_{n_i} = 0, \quad \forall i \in \mathscr{V}.$$

Figure 5 depicts the evolution of the maximum value of $||e_i(k)||_\infty$ for every agent $i \in \mathscr{V}$ for different values of γ_i and λ_i. Note that, in this second example, the noises are only introduced in the system dynamics. This, yields to a situation in which if $U_{i,\rho}$ is greater than $S_{i,\rho}$, the cost function (14) is weighting higher the estimation error, relying more on the measurements taken by the agents than on the system model. Thus, for higher values of γ_i and λ_i a better performance in steady state is obtained.

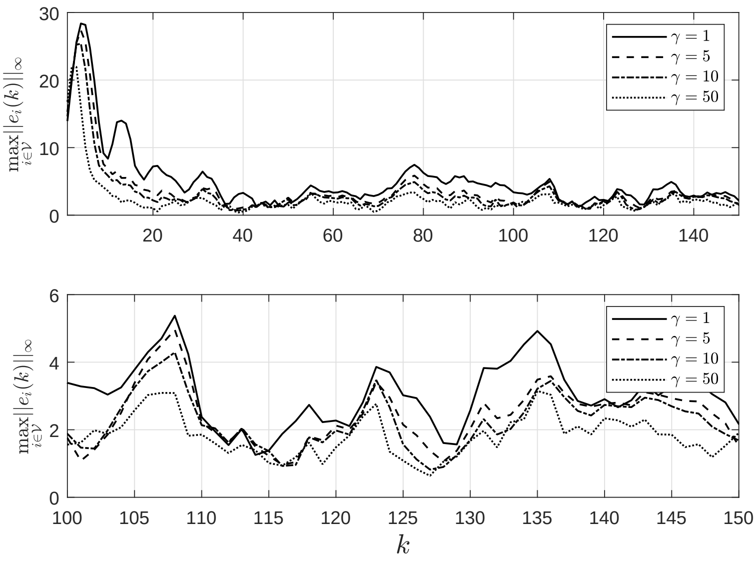

Figure 5. Evolution of the maximum value of the $||e_i(k)||_\infty$ for every agent $i \in \mathcal{V}$ according to γ_i.

6. Conclusions

By exploiting a novel observer structure, a distributed LQ-based design has been introduced in which, adjusting some weighting matrices, the performance of the estimation can be tuned. A method to tune these weighting matrices has been presented in such a way that it is only necessary to adjust the value of two parameters: γ_i and λ_i. The stability of the presented observer structure has been proven under the unperturbed and perturbed scenario. Some simulation examples have been introduced in order to show the effectiveness of the algorithm.

References

Arnold, W. F., & Laub, A. J. (1984, December). Generalized eigenproblem algorithms and software for algebraic Riccati equations. *Proceedings of the IEEE*, *72*(12), 1746-1754.

Das, S., & Moura, J. (2015). Distributed Kalman Filtering with Dynamic Observations Consensus. *IEEE Transactions on Signal Processing*, *63*(17), 4458-4473.

Droge, G., Kawashima, H., & Egerstedt, . B. (2014). Continuous-time proportional-integral distributed optimisation for networked systems. *Journal of Control and Decision*, *1*(3), 191-213.

Garin, F., & Schenato, L. (2010). A survey on distributed estimation and control applications using linear consensus algorithms. In A. Bemporad, M. Heemels, & M. Johansson (Eds.), *Networked control systems* (p. 75-107). London: Springer.

Hespanha, J. P. (2009). *Linear systems theory*. Princeton University Press.

Kar, S., & Moura, J. M. F. (2011). Gossip and distributed Kalman filtering: Weak consensus under weak detectability. *IEEE Transactions on Signal Processing*, *59*(4), 1766-1784.

Khan, U. A., Kar, S., Jadbabaie, A., & Moura, J. M. F. (2010, December). On connectivity, observability, and stability in distributed estimation. In *49th IEEE conference on decision and control (CDC)* (p. 6639-6644). Atlanta, USA.

Kim, T., Shim, H., & Cho, D. D. (2016, December). Distributed Luenberger observer design. In *55th IEEE conference on decision and control (CDC)* (p. 6928-6933). Las Vegas, USA.

Kishor, N., Singh, S., & Raghuvanshi, A. (2006). Dynamic simulations of hydro turbine and its state estimation based LQ control. *Energy Conversion and Management*, *47*(18), 3119 - 3137.

Li, W., Wang, Z., Wei, G., Ma, L., Hu, J., & Ding, D. (2015). A survey on multisensor fusion and consensus filtering for sensor networks. *Discrete Dynamics in Nature and Society*, *2015*.

Lin-Zhang, L., & Wen-Wei, L. (1993). An iterative algorithm for the solution of the discrete-time algebraic riccati equation. *Linear Algebra and Its Applications*, *188*, 465–488.

Mitra, A., & Sundaram, S. (2018). Distributed observers for LTI systems. *IEEE Transactions on Automatic Control*, *63*(11), 3689-3704.

Nedic, A., & Ozdaglar, A. (2009). Distributed subgradient methods for multi-agent optimization. *IEEE Transactions on Automatic Control*, *54*(1), 48-61.

Olfati-Saber, R. (2007, December). Distributed Kalman filtering for sensor networks. In *46th ieee conference on decision and control* (p. 5492-5498). New Orleans, USA.

Orihuela, L., Gómez-Estern, F., & Rubio, F. R. (2014). Scheduled communication in sensor networks. *IEEE Transactions on Control Systems Technology*, *22*(2), 801-808.

Park, S., & Martins, N. C. (2017). Design of distributed LTI observers for state omniscience. *IEEE Transactions on Automatic Control*, *62*(2), 561-576.

Qiu, J., Sun, K., Wang, T., & Gao, H. (2019). Observer-based fuzzy adaptive event-triggered control for pure-feedback nonlinear systems with prescribed performance. *IEEE Transactions on Fuzzy Systems*, 1-1.

Rego, F. F. C., Aguiar, A. P., Pascoal, M., & Jones, C. N. (2017, December). A Design Method for Distributed Luenberger Observers. In *56th IEEE Conference on Decision and Control (CDC)* (p. 3374-3379). Melbourne, Australia.

Rodríguez del Nozal, A., Millán, P., Orihuela, L., Seuret, A., & Zaccarian, L. (2019). Distributed estimation based on multi-hop subspace decomposition. *Automatica*, *99*, 213-220.

Rodríguez del Nozal, A., Orihuela, L., & Millán, P. (2017). Distributed consensus-based Kalman filtering considering subspace decomposition. *IFAC-PapersOnLine*, *50*(1), 2494–2499.

Sayed, A. H. (2014, April). Adaptive networks. *Proceedings of the IEEE*, *102*(4), 460-497.

Shi, C., & Yang, G. (2018). Augmented Lagrange algorithms for distributed optimization over multi-agent networks via edge-based method. *Automatica*, *94*, 55 - 62.

Sun, K., Mou, S., Qiu, J., Wang, T., & Gao, H. (2018). Adaptive fuzzy control for non-triangular structural stochastic switched nonlinear systems with full state constraints. *IEEE Transactions on Fuzzy Systems*, 1-1.

Zhang, H., Lewis, F. L., & Das, A. (2011). Optimal design for synchronization of cooperative systems: state feedback, observer and output feedback. *IEEE Transactions on Automatic Control*, *56*(8), 1948–1952.