Automatica 155 (2023) 111116

Contents lists available at ScienceDirect

Automatica

journal homepage: www.elsevier.com/locate/automatica

Distributed source-user estimation over directed graphs ☆

Fabrizio Caccavale *, Francesco Pierri

School of Engineering, University of Basilicata, 85100 Potenza, Italy

ARTICLE INFO

Article history:
Received 6 June 2022
Received in revised form 15 February 2023
Accepted 30 April 2023
Available online 14 June 2023

Keywords:
Multi-agent systems
Networked systems
Distributed estimation

ABSTRACT

This paper tackles an estimation problem in networked systems. A given time-varying signal is assumed to be measured or computed by only a subset of agents, named *sources*, while the other agents, the *users*, are required to estimate this signal in a distributed way. Each agent communicates only with a subset of neighbouring mates and the communication topology is described by a directed graph with relatively weak connectivity properties. The problem is solved for two class of signals with, respectively, null or bounded derivative of a given order, by resorting to a bank of distributed estimators of the signal and its derivatives run by each user agent. Convergence properties and noise rejection capabilities are investigated. Simulations are run to show the effectiveness of the approach and assess its performance.

1. Introduction

Applications of networked multi-agent systems, such as networks of sensing devices, distributed surveillance systems or cooperative multi-robot systems, require that each agent shares a suitable amount of information with the mates (Ren, Beard, & Atkins, 2007). A key problem in this domain is to design control and/or estimation algorithms by exploiting only local communications among the neighbouring agents (Ge, Han, Ding, Zhang, & Ning, 2018). Thus, in the last two decades, huge research efforts have been devoted to develop distributed control and estimation approaches, with applications in several domains, such as robotics, smart distribution grids, logistics and transportation systems. In several of such applications, it is required to estimate a given variable, known or measured by only a subset of agents in the system. This is the case when, in distributed sensor networks (Akyildiz, Su, Sankarasubramaniam, & Cayirci, 2002), a given physical variable is measured by only a subset of sensing nodes and has to be propagated across the whole network, as often required, e.g., in environmental sensing networks (Hart & Martinez, 2006). On the other hand, the goal of several distributed control schemes for multi-agent systems (e.g., teams of mobile robots) is to track a given reference trajectory known by only a subset of agents. This is the case when the agents in the system are required to track the state of a (virtual) leader agent which, in turn, communicates with only a subset of agents, as, e.g., in Hong, Chen, and Bushnell (2008), Miao, Liu, Wang, Yi, and Fierro (2018) and Ren and Sorensen (2008). A common feature of the above mentioned applications is that a given information, usually represented by a time-varying signal, has to be propagated across a networked system. Frailties of the centralized approaches, where a single central unit transmits the signal to each agent, could be partially overcome by resorting to *flooding* across the network, i.e., each agent that receives the signal forwards it to its neighbours in the communication network. However, this approach shares many drawbacks of centralized solutions and works reasonably well for relatively small-size networks (Kia et al., 2019). Thus, an alternative is distributed estimation of the variable, based only on local interactions among agents.

Distributed estimation problems usually require each agent in the system to compute a local estimate of the state of a dynamical system, based on local measurements and/or estimates from neighbouring agents, as, e.g., in Carli, Chiuso, Schenato, and Zampieri (2008), Li, Dong, Li, and Wang (2019) or Antonelli, Arrichiello, Caccavale, and Marino (2014) and Smith and Hadaegh (2006), where the collective state of the multi-agent system has to be estimated. In Calafiore and Abrate (2009), the problem of estimating an unknown constant parameter from noisy measurements collected by a network of sensors is solved via a distributed consensus diffusion scheme. Distributed state estimation has been exploited to achieve fault diagnosis in networked systems as well (Arrichiello, Marino, & Pierri, 2015; Marino, Pierri, & Arrichiello, 2017). Often, approaches based on dynamic consensus (Olfati-Saber & Jalalkamali, 2012; Olfati-Saber & Shamma, 2005; Spanos, Olfati-Saber, & Murray, 2005; Xu, Li, Xie, & Lum, 2011), where the agents are required to achieve consensus on

☆ The material in this paper was partially presented at the 61st IEEE Conference on Decision and Control, Dec. 6–9, 2022, Cancún, Mexico. This paper was recommended for publication in revised form by Associate Editor Luca Schenato under the direction of Editor Christos G. Cassandras.

* Corresponding author.
E-mail addresses: fabrizio.caccavale@unibas.it (F. Caccavale), francesco.pierri@unibas.it (F. Pierri).

https://doi.org/10.1016/j.automatica.2023.111116
0005-1098/

a given time-varying signal, are conveniently exploited to build distributed estimation and filtering schemes. In Olfati-Saber and Jalalkamali (2012), the estimation problem is related to the state of a moving target, where all the sensing agents measure the state of the target via different partial-state noisy measurements. In Olfati-Saber and Shamma (2005), all the agents measure the same signal (but with a different noise superimposed to it) and the goal is that of reaching ε-consensus on the signal, assumed to have a uniformly bounded rate. In Spanos et al. (2005), a signal is associated to each agent and the problem is the asymptotic tracking of the average of such signals. In Xu et al. (2011), consensus-based dynamic output feedback protocols are designed for consensus and formation control problems, even in the presence of time-varying topologies.

However, differently from the above discussed contributions, the problem tackled in this paper is not related to average consensus but to the estimation of a variable known (i.e., computed or measured) by only a subset of agents, called *sources*, where the variable is not necessarily the state of a given dynamical system. The other agents, the *users*, are required to estimate the signal in a decentralized way, without resorting to flooding strategies. Such a problem can be encountered in coordinated control of multi-robot systems, where the goal is to track the reference trajectories for a set of variables defined at team level, that encode the task to be executed by the robotic system (Antonelli et al., 2014; Ren, Sosnowski, & Hirche, 2020). Often, solutions to this class of problems (see, e.g., Antonelli et al. (2014)) require that each agent (robot) knows the reference trajectories for the task variables and their time derivatives up to a given order. However, in some application scenarios (e.g., exploration of large unstructured environments, limited onboard computing resources), this requirement cannot be met and only a few robotic agents receive from a central task planner (or are able to compute) the reference task variables. Thus, the other robots are required to estimate the reference task variables by resorting only to local interactions with neighbouring teammates. In other applications, the signal to be estimated can represent a physical variable, measured by only a subset of agents, that shall be shared with all the other agents in the system. Such problems could arise in networked sensing systems, where only some nodes are able to measure certain physical variables or compute a function of them (e.g., an indicator), which have to be propagated across the network without resorting to flooding or global broadcasting strategies. The above outlined estimation problem can be tackled via containment control approaches, whose objective is to constrain the state of a subset of agents (followers) in the convex hull of the states of another subset of agents (leaders) (Cao, Ren, & Egerstedt, 2012; Liu, Xie, & Wang, 2012; Yang, Ren, & Liu, 2014). If the leaders play the role of sources and their states coincide with the signal to be estimated, the convex hull coincides with the signal itself. However, containment control strategies can be effectively adopted to solve the estimation problem for relatively narrow classes of signals (e.g., signals with bounded first-order derivative) and require suitable conditions on the gains of the estimators.

Here, the goal is achieved by adopting a bank of cascaded consensus-based estimators, which allows each user agent to track the signal and its time derivatives up to a given order, under suitable assumptions on the topology of the communication network, described by a directed graph. Noticeably, the adopted requirement for the connectivity of the communication graph is considerably weaker than assumptions usually adopted in the literature (e.g., graph strongly connected or containing a directed spanning tree). It is shown that estimation errors converge exponentially for the class of time-varying signals with null derivatives of a given order. Convergence proofs are based on a result related to subgraphs of a directed graphs, which might be of general interest for future work on multi-agent systems. The performance of the estimation scheme are assessed for a more general class of signals having a bounded derivative of a certain order, by devising a bound on the estimation errors. The filtering features of the developed estimation scheme are then investigated to analyse its robustness with respect to measurement noise and disturbances. In addition, the performance of the developed scheme are investigated in the presence of switching topologies as well. Finally, simulation case studies are developed to assess the performance of the proposed estimation strategy.

2. Mathematical preliminaries and problem formulation

Let $\boldsymbol{I}_n$ denote the $(n \times n)$ identity matrix, $\boldsymbol{O}_n$ the $(n \times n)$ null matrix, $\boldsymbol{0}_n$ the $(n \times 1)$ null vector, $\boldsymbol{1}_n$ the $(n \times 1)$ vector of all ones, $\|\cdot\|$ the 2-norm for vectors and matrices and $\otimes$ the Kronecker product (Brewer, 1978). If $\mathcal{S}$ is a finite set, $|\mathcal{S}| < \infty$ denotes its cardinality.

Let $\boldsymbol{\sigma}(t) : t \in \mathbb{R} \to \boldsymbol{\sigma} \in \mathbb{R}^m$ be a smooth vector function of time. Let $\boldsymbol{\sigma}^{(h)}(t)$ denote its time derivative of order h; when useful to simplify notation, the widely adopted symbols $\boldsymbol{\sigma}(t)$, $\dot{\boldsymbol{\sigma}}(t)$ and $\ddot{\boldsymbol{\sigma}}(t)$ will be used in lieu of $\boldsymbol{\sigma}^{(0)}(t)$, $\boldsymbol{\sigma}^{(1)}(t)$ and $\boldsymbol{\sigma}^{(2)}(t)$, respectively. Of course, it is $\boldsymbol{\sigma}^{(h+1)}(t) = \dot{\boldsymbol{\sigma}}^{(h)}(t)$.

It is worth reporting some preliminary results related to input-to-state stability of a class of cascaded linear systems, as they will be exploited to investigate convergence of the estimation errors. Consider $p+1$ linear and time-invariant cascaded systems, with state vectors $\boldsymbol{x}_h \in \mathbb{R}^n$ $(h = p, p-1, \ldots, 1, 0)$

$$\dot{\boldsymbol{x}}_h(t) = \boldsymbol{A}\boldsymbol{x}_h(t) + \boldsymbol{B}_h\boldsymbol{u}_h(t) + \boldsymbol{x}_{h+1}(t), \tag{1}$$

where $\boldsymbol{x}_{p+1}(t) = \boldsymbol{0}_n$, $\forall t \geq 0$, $\boldsymbol{u}_h(t) \in \mathbb{R}^m$ is the input of the hth system, and $\boldsymbol{A}$, $\boldsymbol{B}_h$ are matrices of appropriate dimensions. Given any $\epsilon \geq 0$, there exists a $\rho \geq 1$ such that the transition matrix $\boldsymbol{\Phi}(t) = \exp(\boldsymbol{A}t)$ satisfies the well-known inequality (Kågström, 1977)

$$\|\boldsymbol{\Phi}(t)\| \leq \rho\, e^{-\alpha t}, \quad \forall t \geq 0, \tag{2}$$

where $\alpha = -(\lambda_M(\boldsymbol{A}) + \epsilon)$ and $\lambda_M(\cdot)$ denotes the maximal real part of the eigenvalues of a matrix (i.e., its spectral abscissa). Notice that, when $\boldsymbol{A}$ is a Hurwitz matrix, it is $\lambda_M(\boldsymbol{A}) < 0$ and ϵ can be always chosen small enough such that $-\alpha < 0$. On the other hand, ρ is the condition number of a matrix which depends on ϵ and the structure of the Jordan form of $\boldsymbol{A}$; it increases as ϵ decreases (Kågström, 1977). Thus, the bound (2) becomes more conservative as ϵ decreases, especially for small values of t, while it becomes closer to $\|\boldsymbol{\Phi}(t)\|$ for large values of t.

When the inputs $\boldsymbol{u}_h(t)$ are all identically null, the following result readily follows from the fact that each system (1) is either an autonomous exponentially stable linear time-invariant system (for $h = p$) or an exponentially stable linear time-invariant system with the asymptotically vanishing input $\boldsymbol{x}_{h+1}(t)$ (for $h < p$):

Lemma 1. *Consider the cascaded linear systems defined by (1) and assume that:*

- *$\boldsymbol{A}$ is Hurwitz,*
- *$\boldsymbol{u}_h(t) = \boldsymbol{0}_m$, $\forall h = p, p-1, \ldots, 1, 0$, $\forall t \geq 0$.*

Then, $\boldsymbol{x}_h(t)$ converges to $\boldsymbol{0}_n$ as $t \to \infty$, $\forall h = p, p-1, \ldots, 1, 0$.

Notice that the dynamics of $\boldsymbol{x} = \left[\boldsymbol{x}_p^{\mathrm{T}}\, \boldsymbol{x}_{p-1}^{\mathrm{T}} \ldots \boldsymbol{x}_1^{\mathrm{T}}\, \boldsymbol{x}_0^{\mathrm{T}}\right]^{\mathrm{T}}$ is characterized by a block-triangular matrix, with blocks along the diagonal all equal to $\boldsymbol{A}$. Hence, exponential convergence of $\boldsymbol{x}(t)$ to $\boldsymbol{0}_{n(p+1)}$ follows under the assumptions of Lemma 1.

On the other hand, when the inputs $\boldsymbol{u}_h(t)$ are all uniformly bounded, the following result, whose proof is reported in Appendix A, holds:

Lemma 2. *Consider the cascaded linear systems defined by (1) and assume that:*

- $\boldsymbol{A}$ *is Hurwitz,*
- *the inputs* $\boldsymbol{u}_h(t)$ *are bounded, i.e., there exist positive numbers,* $\bar{u}_h < \infty$*, such that* $\|\boldsymbol{u}_h(t)\| \leq \bar{u}_h$*,* $\forall h = p, p-1, \ldots, 1, 0$*,* $\forall t \geq 0$*.*

Then, $\boldsymbol{x}_h(t)$ *can be bounded as follows,* $\forall h = p, p-1, \ldots, 1, 0$*,* $\forall t \geq 0$*:*

$$\|\boldsymbol{x}_h(t)\| \leq \sum_{l=h}^{p} t^{l-h} e^{-\alpha t} \bar{x}_l(0) + \sum_{l=h}^{p} \bar{b}_l \bar{u}_l, \tag{3}$$

where

$$\bar{x}_l(0) = \frac{\rho^{l-h+1}}{(l-h)!} \|\boldsymbol{x}_l(0)\|, \quad \bar{b}_l = b_l \left(\frac{\rho}{\alpha}\right)^{l-h+1}, \tag{4}$$

and $b_l = \|\boldsymbol{B}_l\|$*, while* ρ *and* α *are the defined by (2).*

Clearly, the first term in (3) vanishes as $t \to \infty$, and, thus, the bound asymptotically reduces to $\sum_{l=h}^{p} \bar{b}_l \bar{u}_l$. Notice that Lemma 1 can be proven as a direct consequence of Lemma 2 and the first term in (3) can be seen as a time-varying upper bound on $\boldsymbol{x}_h(t)$ when all the external inputs are null.

2.1. Multi-agent system and communication topology

Consider a system composed of N agents, each identified by an index belonging to the set $\mathcal{N} = \{1, \ldots, N\}$. Let the topology of the communication flow among the agents be described by a *directed* graph $\mathcal{G} = \{\mathcal{E}, \mathcal{N}\}$, where the N vertices (nodes) of the graph represent the agents in the systems and $\mathcal{E} \subseteq \mathcal{N} \times \mathcal{N}$ is the set of edges (arcs) connecting the nodes. If all the communication links between the agents are bi-directional (i.e., $(i,j) \in \mathcal{E} \Rightarrow (j,i) \in \mathcal{E}$), the graph is *undirected*. Consider an edge $(i,j) \in \mathcal{E}$: i is called *parent node* and j *child node*. A *subgraph*, $\mathcal{G}' = \{\mathcal{E}', \mathcal{N}'\}$, of $\mathcal{G}$ is a graph such that $\mathcal{N}' \subseteq \mathcal{N}$ and $\mathcal{E}' \subseteq \mathcal{E} \cap (\mathcal{N}' \times \mathcal{N}')$.

The $(N \times N)$ *adjacency matrix* of $\mathcal{G}$ is given by

$$\boldsymbol{A} = \{a_{ij}\}_{i,j \in \mathcal{N}} : \ a_{ii} = 0, \ a_{ij} = \begin{cases} 1 & \text{if}\,(j,i) \in \mathcal{E} \\ 0 & \text{otherwise,} \end{cases} \tag{5}$$

whose element a_{ij} ($i \neq j$) is non-null if $j \in \mathcal{N}_i$, where $\mathcal{N}_i = \{j \in \mathcal{N} : (j,i) \in \mathcal{E}\}$ is the set of indexes of the neighbours (i.e., the parent nodes) of the ith agent, whose cardinality is denoted by $N_i = |\mathcal{N}_i|$.

The $(N \times N)$ *Laplacian matrix* of $\mathcal{G}$

$$\boldsymbol{L} = \{l_{ij}\}_{i,j \in \mathcal{N}} : \ l_{ii} = \sum_{j=1, j \neq i}^{N} a_{ij}, \ l_{ij} = -a_{ij}, \ i \neq j, \tag{6}$$

plays a key role in consensus problems over graphs (see, e.g., Olfati-Saber and Murray (2004)). The Laplacian matrix has all eigenvalues with non-negative real parts and at least a zero eigenvalue with $\boldsymbol{1}_N$ as the corresponding right eigenvector. Notice that $-\boldsymbol{L}$ is a Metzler matrix, i.e., its off-diagonal elements are all non-negative. A directed graph is called *strongly connected* (SC) if any two distinct nodes of the graph can be connected via a directed path, i.e., a path that follows the direction of the edges of the graph. Moreover, the adjacency and Laplacian matrices are both irreducible (i.e., it is not possible to put them in block-triangular form via permutation of rows and columns) if and only if the associated graph is SC. An undirected graph is called *connected* if there is an undirected path between every pair of distinct nodes. A node of a directed graph is *balanced* if its in-degree (i.e., the number of incoming edges) and its out-degree (i.e., the number of outgoing edges) are equal; a directed graph is called *balanced* if each node of the graph is balanced. Any undirected graph is balanced.

Consider a directed graph, $\mathcal{G} = \{\mathcal{N}, \mathcal{E}\}$, with Laplacian matrix $\boldsymbol{L}$, and let $\mathcal{N}$ be partitioned into two disjoint subsets, $\mathcal{N}_a$ (with $N_a = |\mathcal{N}_a|$) and $\mathcal{N}_b$ (with $N_b = |\mathcal{N}_b|$), i.e., $\mathcal{N}_b = \mathcal{N} \setminus \mathcal{N}_a$; hereafter, $\mathcal{G}[\mathcal{N}_a, \mathcal{N}_b]$ will denote a graph $\mathcal{G}$ with such a partition. Connectivity between the two subgraphs $\mathcal{G}_a = \{\mathcal{E} \cap (\mathcal{N}_a \times \mathcal{N}_a), \mathcal{N}_a\}$ and $\mathcal{G}_b = \{\mathcal{E} \cap (\mathcal{N}_b \times \mathcal{N}_b), \mathcal{N}_b\}$ can be characterized by the following property:

Property 1. *The directed graph* $\mathcal{G}[\mathcal{N}_a, \mathcal{N}_b]$ *is said to satisfy the subgraph connectivity (SGC) property (or, simply, the graph is SGC) if, for any node* $i \in \mathcal{N}_b$*, there exists a directed path from some node* $j \in \mathcal{N}_a$ *to* i*.*

This property can be interpreted in terms of information flow between the nodes in the two subgraphs, ensuring that the information carried by the nodes in $\mathcal{N}_a$ is conveyed to the nodes in $\mathcal{N}_b$. Clearly, the SGC property is weaker than requiring the graph to be SC.

Consider the $(N_b \times N_b)$ principal submatrix, $\bar{\boldsymbol{L}}$, obtained by deleting the rows and the columns of $\boldsymbol{L}$ with indexes belonging to $\mathcal{N}_a$. Let r_i ($i \in \bar{\mathcal{N}}_b = \{1, \ldots, N_b\}$) be the index of the ith node belonging to $\mathcal{N}_b$. It can be recognized that $\bar{\boldsymbol{L}}$ can be written as

$$\bar{\boldsymbol{L}} = \boldsymbol{L}_b + \boldsymbol{D}_{ba}, \tag{7}$$

where $\boldsymbol{L}_b$ is the Laplacian of the subgraph $\mathcal{G}_b$, $\boldsymbol{D}_{ba}$=diag $\{N_{a,r_i}\}_{r_i \in \mathcal{N}_b}$ and $N_{a,r_i} \geq 0$ is the number of parent nodes of r_i belonging to $\mathcal{N}_a$. Then, the following result, whose proof is reported in Appendix B, can be established:

Theorem 1. *If the directed graph* $\mathcal{G}[\mathcal{N}_a, \mathcal{N}_b]$ *satisfies the SGC property, then* $-\bar{\boldsymbol{L}}$ *is Hurwitz. If the graph does not satisfy the SGC property, the eigenvalues of* $-\bar{\boldsymbol{L}}$ *have all non-positive real parts and at least one of them is null.*

Consider two partitions of $\mathcal{G}$, $\mathcal{G}[\mathcal{N}'_a, \mathcal{N}'_b]$ and $\mathcal{G}[\mathcal{N}''_a, \mathcal{N}''_b]$, where $\mathcal{N}''_a$ and $\mathcal{N}''_b$ are obtained by removing a set of nodes, $\mathcal{N}_\delta \subseteq \mathcal{N}'_b$, from $\mathcal{N}'_b$ and adding them to $\mathcal{N}'_a$ (i.e., $\mathcal{N}''_a = \mathcal{N}'_a \cup \mathcal{N}_\delta$ and $\mathcal{N}''_b = \mathcal{N}'_b \setminus \mathcal{N}_\delta$), while the graph edges remain all unchanged. Notice that if $\mathcal{G}[\mathcal{N}'_a, \mathcal{N}'_b]$ satisfies the SGC property, the same holds for $\mathcal{G}[\mathcal{N}''_a, \mathcal{N}''_b]$. Then, the following result, whose proof is reported in Appendix C, holds:

Proposition 1. *Consider a graph* $\mathcal{G}$ *with the two partitions,* $\mathcal{G}[\mathcal{N}'_a, \mathcal{N}'_b]$ *and* $\mathcal{G}[\mathcal{N}''_a, \mathcal{N}''_b]$*, obtained as detailed above. Let* $\mathcal{G}[\mathcal{N}'_a, \mathcal{N}'_b]$ *satisfy the SGC property and* $\bar{\boldsymbol{L}}'$ *(respectively,* $\bar{\boldsymbol{L}}''$*) be the principal submatrix obtained from the graph Laplacian by removing the rows and columns with indexes belonging to* $\mathcal{N}'_a$ *(respectively,* $\mathcal{N}''_a$*). Then, the following inequality holds:*

$$\lambda_M(-\bar{\boldsymbol{L}}'') \leq \lambda_M(-\bar{\boldsymbol{L}}'), \tag{8}$$

which is strict if $\mathcal{G}'_b = \{\mathcal{E} \cap (\mathcal{N}'_b \times \mathcal{N}'_b), \mathcal{N}'_b\}$ *is SC.*

Consider now two directed graphs $\mathcal{G}'[\mathcal{N}_a, \mathcal{N}_b]$ and $\mathcal{G}''[\mathcal{N}_a, \mathcal{N}_b]$, which differ only for the edges connecting the nodes in $\mathcal{N}_a$ to those in $\mathcal{N}_b$. Then, it is $\bar{\boldsymbol{L}}' = \boldsymbol{L}_b + \boldsymbol{D}'_{ba}$ and $\bar{\boldsymbol{L}}'' = \boldsymbol{L}_b + \boldsymbol{D}''_{ba}$, respectively. The following result, whose proof is reported in Appendix D, holds:

Proposition 2. *If* $\mathcal{G}'[\mathcal{N}_a, \mathcal{N}_b]$ *and* $\mathcal{G}''[\mathcal{N}_a, \mathcal{N}_b]$ *satisfy the SGC property and* $\boldsymbol{D}''_{ba} \geq \boldsymbol{D}'_{ba}$*, where the matrix inequality is understood component-wise, the following inequality holds*

$$\lambda_M(-\bar{\boldsymbol{L}}'') \leq \lambda_M(-\bar{\boldsymbol{L}}'), \tag{9}$$

which is strict if $\boldsymbol{D}'_{ba} \neq \boldsymbol{D}''_{ba}$ *and* $\mathcal{G}_b$ *is SC.*

2.2. Estimation problem

It is assumed that only a subset of agents (called *sources*) in the system measures (or computes) the smooth time-varying vector signal $\boldsymbol{\sigma}(t) \in \mathbb{R}^m$. The indexes of source agents belong to the set $\mathcal{N}_s \subseteq \mathcal{N}$, with $N_s = |\mathcal{N}_s|$. The other agents (called *users*), whose indexes belong to the set $\mathcal{N}_u = \mathcal{N} \backslash \mathcal{N}_s$ (with $N_u = |\mathcal{N}_u|$), do not have access to $\boldsymbol{\sigma}(t)$ and are required to estimate it and, possibly, its time derivatives $\boldsymbol{\sigma}^{(h)}(t)$ up to the pth order. Notice that the communication graph can be partitioned as $\mathcal{G}[\mathcal{N}_s, \mathcal{N}_u]$ and the two subgraphs, $\mathcal{G}_s = \{\mathcal{E} \cap (\mathcal{N}_s \times \mathcal{N}_s), \mathcal{N}_s\}$ and $\mathcal{G}_u = \{\mathcal{E} \cap (\mathcal{N}_u \times \mathcal{N}_u), \mathcal{N}_u\}$, can be defined as in Section 2.1. Hereafter, $\mathcal{N}_{u,i} = \mathcal{N}_i \cap \mathcal{N}_u$ will denote the subset (with $N_{u,i} = |\mathcal{N}_{u,i}|$) of parent nodes of the ith agent which are users, while $\mathcal{N}_{s,i} = \mathcal{N}_i \cap \mathcal{N}_s$ is the subset (with $N_{s,i} = |\mathcal{N}_{s,i}|$) of parent nodes of the ith agent which are sources.

The estimate of the hth derivative computed by the ith user agent ($i \in \mathcal{N}_u$) will be denoted by ${}^i\widehat{\boldsymbol{\sigma}}_h(t)$, i.e., the left superscript will identify the agent and the right subscript the order of the estimated derivative.[1]

The signal $\boldsymbol{\sigma}(t)$ is required to satisfy the following assumption:

Assumption 1. There exists an integer $p \geq 0$ such that $\boldsymbol{\sigma}^{(p+1)}(t) = \mathbf{0}_m$, $\forall t \geq 0$.

In many applications, Assumption 1 would be too strong and a weaker condition will be considered as well:

Assumption 2. There exist an integer $p \geq 0$ and a constant, $\bar{\sigma}_{p+1} < \infty$, such that $\|\boldsymbol{\sigma}^{(p+1)}(t)\| \leq \bar{\sigma}_{p+1}$, $\forall t \geq 0$.

In some cases, e.g., when $\boldsymbol{\sigma}(t)$ represents a reference trajectory for a multi-robot system, fulfilment of Assumption 2 can be imposed with user-defined p and $\bar{\sigma}_{p+1}$. Otherwise, when the properties of $\boldsymbol{\sigma}(t)$ cannot be suitably imposed, Assumption 2 requires that it must be sufficiently slowly varying and the achievable performance of the estimation scheme will depend on the magnitude of the bound.

3. Distributed estimation scheme

If the ith agent is a source (i.e., $i \in \mathcal{N}_s$), its estimate coincides with $\boldsymbol{\sigma}^{(h)}(t)$, i.e.

$$ {}^i\widehat{\boldsymbol{\sigma}}_h(t) = \boldsymbol{\sigma}^{(h)}(t), \quad i \in \mathcal{N}_s, \quad h = p, p-1, \ldots, 1, 0. \tag{10}$$

If the ith agent is a user (i.e., $i \in \mathcal{N}_u$), it has to estimate $\boldsymbol{\sigma}(t)$ and its derivatives up to a given order (less than or equal to p). The estimation scheme is composed of a chain of $p+1$ cascaded estimators: each of them computes an estimate of the derivative of order h (with $h = p, p-1, \ldots, 1, 0$), ${}^i\widehat{\boldsymbol{\sigma}}_h(t)$, as follows

$$ {}^i\dot{\widehat{\boldsymbol{\sigma}}}_h(t) = k_o \sum_{j \in \mathcal{N}_i} \left({}^j\widehat{\boldsymbol{\sigma}}_h(t) - {}^i\widehat{\boldsymbol{\sigma}}_h(t)\right) + {}^i\widehat{\boldsymbol{\sigma}}_{h+1}(t), \tag{11}$$

where ${}^i\widehat{\boldsymbol{\sigma}}_{p+1}(t) = \mathbf{0}_m$, $\forall i \in \mathcal{N}_u$, $\forall t \geq 0$, and k_o is a positive gain. The first term on the right-hand side of (11) represents a classical consensus term (Olfati-Saber & Murray, 2004), which is needed to propagate the information across the network. The second term is aimed at compensating the effects of the higher-order derivatives of $\boldsymbol{\sigma}(t)$ on the estimates.

In view of (10), the dynamics of the estimation error, ${}^i\widetilde{\boldsymbol{\sigma}}_h = \boldsymbol{\sigma}^{(h)} - {}^i\widehat{\boldsymbol{\sigma}}_h$, can be written as

$$\begin{aligned} {}^i\dot{\widetilde{\boldsymbol{\sigma}}}_h(t) &= -k_o \sum_{j \in \mathcal{N}_i} \left({}^i\widetilde{\boldsymbol{\sigma}}_h(t) - {}^j\widetilde{\boldsymbol{\sigma}}_h(t)\right) + \\ &\quad \delta_{hp}\boldsymbol{\sigma}^{(p+1)}(t) + \bar{\delta}_{hp}\,{}^i\widetilde{\boldsymbol{\sigma}}_{h+1}(t) \\ &= -k_o N_i\,{}^i\widetilde{\boldsymbol{\sigma}}_h(t) + k_o \sum_{j \in \mathcal{N}_{u,i}} {}^j\widetilde{\boldsymbol{\sigma}}_h(t) + \\ &\quad \delta_{hp}\boldsymbol{\sigma}^{(p+1)}(t) + \bar{\delta}_{hp}\,{}^i\widetilde{\boldsymbol{\sigma}}_{h+1}(t), \end{aligned} \tag{12}$$

where $\bar{\delta}_{hp} = 1 - \delta_{hp}$, and δ_{hp} is the Kronecker delta (i.e., $\delta_{hp} = 1$, when $h = p$, otherwise $\delta_{hp} = 0$). Let $m_u = mN_u$ and let $\widetilde{\boldsymbol{\sigma}}_h = \left\{{}^i\widetilde{\boldsymbol{\sigma}}_h\right\}_{i \in \mathcal{N}_u} \in \mathbb{R}^{m_u}$ be the vector collecting the estimation errors for the user agents. Then, the dynamics of $\widetilde{\boldsymbol{\sigma}}_h(t)$ is given

$$ \dot{\widetilde{\boldsymbol{\sigma}}}_h(t) = -\bar{\boldsymbol{L}}_o\widetilde{\boldsymbol{\sigma}}_h(t) + \delta_{hp}\boldsymbol{\mu}(t) + \bar{\delta}_{hp}\widetilde{\boldsymbol{\sigma}}_{h+1}(t), \tag{13}$$

where $\boldsymbol{\mu}(t) = \mathbf{1}_{N_u} \otimes \boldsymbol{\sigma}^{(p+1)}(t)$,

$$ \bar{\boldsymbol{L}}_o = k_o\left(\bar{\boldsymbol{L}} \otimes \boldsymbol{I}_m\right), \tag{14}$$

and $\bar{\boldsymbol{L}}$ is the $(N_u \times N_u)$ matrix extracted from the graph Laplacian, $\boldsymbol{L}$, by deleting the rows and the columns with indexes belonging to $\mathcal{N}_s$.

It is worth noticing that, without the additional term ${}^i\widehat{\boldsymbol{\sigma}}_{h+1}(t)$ in (11), the dynamics (13) would become $\dot{\widetilde{\boldsymbol{\sigma}}}_h(t) = -\bar{\boldsymbol{L}}_o\widetilde{\boldsymbol{\sigma}}_h(t) + \boldsymbol{\sigma}^{(h+1)}(t)$, for any h. Hence, the behaviour of the estimation errors would be influenced by the (possibly unbounded) $(h+1)$th derivative of the signal, thus preventing convergence (respectively, boundedness) of $\widetilde{\boldsymbol{\sigma}}_h(t)$ when $\boldsymbol{\sigma}^{(h+1)}(t)$ is not null (respectively, not bounded).

3.1. Convergence of estimation errors

Matrix $-\bar{\boldsymbol{L}}_o$ satisfies the following Lemma:

Lemma 3. *If the communication graph $\mathcal{G}[\mathcal{N}_s, \mathcal{N}_u]$ satisfies the SGC property, then, matrix $-\bar{\boldsymbol{L}}_o$ is Hurwitz and, given any $\epsilon_o \geq 0$, there exists a $\rho_o \geq 1$ such that the corresponding transition matrix $\boldsymbol{\Phi}_o(t) = \exp\left(-\bar{\boldsymbol{L}}_o t\right)$ satisfies the inequality*

$$ \|\boldsymbol{\Phi}_o(t)\| \leq \rho_o\, e^{-\alpha_o t}, \quad \forall t \geq 0, \tag{15}$$

where $\alpha_o = -(\lambda_M(-\bar{\boldsymbol{L}}_o) + \epsilon_o)$ and $\lambda_M(-\bar{\boldsymbol{L}}_o) = \lambda_M(-k_o(\bar{\boldsymbol{L}} \otimes \boldsymbol{I}_m)) = k_o\lambda_M(-\bar{\boldsymbol{L}}) < 0$.

Proof. Since the communication graph is assumed to satisfy the SGC property, $\bar{\boldsymbol{L}}$ satisfies the assumptions of Theorem 1, and thus $-\bar{\boldsymbol{L}}$ is Hurwitz. Since any eigenvalue of the Kronecker product $-\bar{\boldsymbol{L}} \otimes \boldsymbol{I}_m$ arises as a product of eigenvalues of $-\bar{\boldsymbol{L}}$ and $\boldsymbol{I}_m$ (Brewer, 1978), $-k_o\left(\bar{\boldsymbol{L}} \otimes \boldsymbol{I}_m\right)$ is Hurwitz as well. Hence, (15) follows directly from (2). ■

Of course, as already noticed in Section 2, it is always possible to choose ϵ_o small enough such that $\alpha_o < 0$ (and so will be assumed in the following). On the other hand, the eigenvalues of $-\bar{\boldsymbol{L}}_o$ are those of $-(\bar{\boldsymbol{L}} \otimes \boldsymbol{I}_m)$, which are fixed for a given topology, scaled by k_o; thus, $-\lambda_M(-\bar{\boldsymbol{L}}_o)$ increases with k_o. Moreover, since the structure of the Jordan normal form of $-\bar{\boldsymbol{L}}_o$ is not influenced by k_o, ρ_o is not influenced by k_o. Therefore, since α_o increases with k_o, ρ_o/α_o decreases with k_o and, thus, for any ϵ_o, it is always possible to find a value of k_o large enough such that $\rho_o/\alpha_o < 1$.

When Assumption 1 holds, $\boldsymbol{\sigma}^{(p+1)}(t)$ is identically null and the dynamics (13) has the same cascaded structure defined by (1), with $\boldsymbol{A} = -\bar{\boldsymbol{L}}_o$ and $\boldsymbol{u}_h(t) = \mathbf{0}_{m_u}$, $\forall h = p, p-1, \ldots, 1, 0$ and $\forall t \geq 0$. Thus, the following result follows directly from Lemmas 1 and 3:

Theorem 2. *If Assumption 1 holds and the communication graph $\mathcal{G}[\mathcal{N}_s, \mathcal{N}_u]$ satisfies the SGC property, then $\widetilde{\boldsymbol{\sigma}}_h(t)$ converges to $\mathbf{0}_{m_u}$ as $t \to \infty$, $\forall h = p, p-1, \ldots, 1, 0$.*

[1] This notation, although a bit involved, is needed, since ${}^i\dot{\widehat{\boldsymbol{\sigma}}}_h \neq {}^i\widehat{\boldsymbol{\sigma}}_{h+1}$, and thus the use of the superscript (h) for the estimate of the hth derivative would be misleading.

Clearly, in view of (15), since $-\lambda_M(-\bar{\boldsymbol{L}}_o)$ increases with k_o, the larger is k_o the faster is the convergence of the estimation errors. Notice that the dynamics of $\widetilde{\boldsymbol{\sigma}} = \left[\widetilde{\boldsymbol{\sigma}}_p^{\mathrm{T}}\,\widetilde{\boldsymbol{\sigma}}_{p-1}^{\mathrm{T}}\,\ldots\,\widetilde{\boldsymbol{\sigma}}_1^{\mathrm{T}}\,\widetilde{\boldsymbol{\sigma}}_0^{\mathrm{T}}\right]^{\mathrm{T}}$ is characterized by a block-triangular matrix, with blocks along the diagonal all equal to $-\bar{\boldsymbol{L}}_o$. Hence, exponential convergence of $\widetilde{\boldsymbol{\sigma}}(t)$ to $\mathbf{0}_{m_u(p+1)}$ follows.

On the other hand, when Assumption 2 holds, the dynamics (13) has the same cascaded structure defined by (1), with $\boldsymbol{A} = -\bar{\boldsymbol{L}}_o$, $\boldsymbol{B}_h = \boldsymbol{I}_{m_u}$ and $\boldsymbol{u}_h(t) = \delta_{hp}\boldsymbol{\mu}(t)$. Thus, a time-varying bound on the collective estimation error can be computed as a direct application of Lemma 2, with $\bar{u}_p = \sqrt{N_u}\bar{\sigma}_{p+1}$ and, $\forall h < p$, $\bar{u}_h = 0$:

Theorem 3. *If Assumption 2 holds and the communication graph $\mathcal{G}[\mathcal{N}_s, \mathcal{N}_u]$ satisfies the SGC property, then $\widetilde{\boldsymbol{\sigma}}_h(t)$ can be bounded, $\forall h = p, p-1, \ldots, 1, 0$ and $\forall t \geq 0$, as follows*

$$\|\widetilde{\boldsymbol{\sigma}}_h(t)\| \leq \sum_{l=h}^{p} t^{l-h} e^{-\alpha_o t}\bar{\sigma}_l(0) + \bar{\beta}_p\sqrt{N_u}\bar{\sigma}_{p+1}, \tag{16}$$

where

$$\bar{\sigma}_l(0) = \frac{\rho_o^{l-h+1}}{(l-h)!}\,\|\widetilde{\boldsymbol{\sigma}}_l(0)\|\,, \quad \bar{\beta}_p = \left(\frac{\rho_o}{\alpha_o}\right)^{p-h+1}. \tag{17}$$

The first term in (16) vanishes as $t \to \infty$, and, thus, the bound asymptotically reduces to $\bar{\beta}_p\sqrt{N_u}\bar{\sigma}_{p+1}$. Notice that Theorem 2 can be proven as a direct consequence of Theorem 3 and the first term in (16) can be seen as an upper bound on $\widetilde{\boldsymbol{\sigma}}_h(t)$ when $\boldsymbol{\sigma}^{(p+1)}(t)$ is identically null.

Notice that, assuming k_o large enough such that $\rho_o/\alpha_o < 1$, Eq. (17) shows that the more the index h decreases (i.e., proceeding along the chain of estimators) the more the effect of $\boldsymbol{\sigma}^{(p+1)}$ on the estimation errors is reduced, since $(\rho_o/\alpha_o)^{p-h+1} < 1$ decreases as h approaches 0. Moreover, since ρ_o/α_o decreases for increasing values of k_o, $\bar{\beta}_p$ decreases as well. Of course, the above remarks are not be regarded as strict design guidelines, since inequality (15) could be rather conservative, especially for small values of t, but a way to remark that the estimation errors can be reduced by increasing k_o and that the effect of a non-null $\boldsymbol{\sigma}^{(p+1)}(t)$ is progressively reduced along the chain of estimators.

Finally, similarly to Tu and Sayed (2013), it is worth studying the effect of the number and distribution of sources in the network on the performance of the proposed distributed estimation scheme. To this aim, Proposition 1 can be exploited. Let the communication graph $\mathcal{G}\left[\mathcal{N}_s', \mathcal{N}_u'\right]$ satisfy the SGC property. Then, if a set of users $\mathcal{N}_\delta$ (with $N_\delta = |\mathcal{N}_\delta|$) is removed from $\mathcal{N}_u'$ and added to $\mathcal{N}_s'$, a new partition of the graph is generated, $\mathcal{G}\left[\mathcal{N}_s'', \mathcal{N}_u''\right]$, with $\mathcal{N}_s'' = \mathcal{N}_s' \cup \mathcal{N}_\delta$ and $\mathcal{N}_u'' = \mathcal{N}_u' \setminus \mathcal{N}_\delta$, which is still SGC and characterized by the $(N_s' - N_\delta) \times (N_s' - N_\delta)$ principal submatrix of the Laplacian, $\bar{\boldsymbol{L}}''$, obtained by deleting the rows and the columns corresponding to new set of source nodes. Then, according to (8), $\lambda_M(-\bar{\boldsymbol{L}}'')$ is not larger (respectively, smaller, if $\mathcal{G}_u'$ is SC) than $\lambda_M(-\bar{\boldsymbol{L}}')$. When $\mathcal{G}_u$ is not SC, since inequality (8) is not strict, it is not guaranteed that any arbitrary users/sources shift leads to a decrease of the spectral abscissa of $-\bar{\boldsymbol{L}}$. Notice that, however, such a monotonic behaviour is guaranteed only if $\mathcal{N}_s'' = \mathcal{N}_s' \cup \mathcal{N}_\delta$ and $\mathcal{N}_u'' = \mathcal{N}_u' \setminus \mathcal{N}_\delta$, i.e., if the set of sources is enriched by shifting some users to sources.

Let r_i ($i \in \bar{\mathcal{N}}_u = \{1, \ldots, N_u\}$) be the index of the ith node belonging to $\mathcal{N}_u$ and $N_{s,r_i} \geq 0$ the number of parent nodes of r_i belonging to $\mathcal{N}_s$. It can be recognized that $\bar{\boldsymbol{L}}$ can be expressed as in (7) and Proposition 2 can be invoked to conclude that $\lambda_M(-\bar{\boldsymbol{L}})$, for a given fixed partition, $\mathcal{G}[\mathcal{N}_s, \mathcal{N}_u]$, does not increase (respectively, decreases, if $\mathcal{G}_u$ is SC) when at least one N_{s,r_i} increases. In other words, when $\mathcal{G}_u$ is SC, $\lambda_M(-\bar{\boldsymbol{L}})$ becomes more negative as new source-user communication links are added, which can happen, e.g., when new users enter the communication range of one or more sources. Otherwise, when $\mathcal{G}_u$ is not SC, such a strict monotonic behaviour is not guaranteed.

3.2. Effect of noise and disturbances

When $\boldsymbol{\sigma}(t)$ and its derivatives are measured by source nodes, they can be corrupted by noise, which can be further amplified when only $\boldsymbol{\sigma}(t)$ is measured and its derivatives are computed numerically. Thus, it is important to assess the filtering capabilities of the proposed estimation scheme. To this aim, assume that the hth derivative of $\boldsymbol{\sigma}(t)$ transmitted by the jth source is affected by an additional disturbance input, ${}^j\boldsymbol{\upsilon}_h(t)$

$${}^j\widehat{\boldsymbol{\sigma}}_h(t) = \boldsymbol{\sigma}^{(h)}(t) + {}^j\boldsymbol{\upsilon}_h(t), \qquad j \in \mathcal{N}_s, \tag{18}$$

which is assumed, as in He, Zhou, Cheng, Shi, and Chen (2016), to be bounded, i.e., $\forall h = p, p-1, \ldots, 1, 0$ and $\forall j \in \mathcal{N}_s$, it is

$$\left\|{}^j\boldsymbol{\upsilon}_h(t)\right\| \leq {}^j\bar{\upsilon}_h < \infty, \quad \forall t \geq 0. \tag{19}$$

The estimation error dynamics, for $i \in \mathcal{N}_u$, is given by

$${}^i\dot{\widetilde{\boldsymbol{\sigma}}}_h(t) = -k_o N_i\,{}^i\widetilde{\boldsymbol{\sigma}}_h(t) + k_o \sum_{j\in\mathcal{N}_{u,i}} {}^j\widetilde{\boldsymbol{\sigma}}_h(t) + \delta_{hp}\boldsymbol{\sigma}^{(p+1)}(t) + \bar{\delta}_{hp}\,{}^i\widetilde{\boldsymbol{\sigma}}_{h+1}(t) - k_o\boldsymbol{\upsilon}_{h,i}(t), \tag{20}$$

where

$$\boldsymbol{\upsilon}_{h,i}(t) = \sum_{j\in\mathcal{N}_{s,i}} {}^j\boldsymbol{\upsilon}_h(t).$$

Let $\boldsymbol{\upsilon}_h = \left\{\boldsymbol{\upsilon}_{h,i}\right\}_{i\in\mathcal{N}_u} \in \mathbb{R}^{m_u}$, the dynamics of the collective estimation error for the users becomes

$$\dot{\widetilde{\boldsymbol{\sigma}}}_h(t) = -\bar{\boldsymbol{L}}_o\widetilde{\boldsymbol{\sigma}}_h(t) + \delta_{hp}\boldsymbol{\mu}(t) - k_o\boldsymbol{\upsilon}_h(t) + \bar{\delta}_{hp}\widetilde{\boldsymbol{\sigma}}_{h+1}(t), \tag{21}$$

where, in view of (19), $\boldsymbol{\upsilon}_h$ can be bounded as follows

$$\|\boldsymbol{\upsilon}_h(t)\| \leq \sum_{i\in\mathcal{N}_u}\sum_{j\in\mathcal{N}_{s,i}} \left\|{}^j\boldsymbol{\upsilon}_h(t)\right\| \leq \bar{\upsilon}_h, \tag{22}$$

with $\bar{\upsilon}_h = \sum_{i\in\mathcal{N}_u}\sum_{j\in\mathcal{N}_{s,i}} {}^j\bar{\upsilon}_h$.

Therefore, the dynamics (21) has the same cascaded structure defined by (1), with $\boldsymbol{A} = -\bar{\boldsymbol{L}}_o$, $\boldsymbol{B}_h = \boldsymbol{I}_{m_u}$ and $\boldsymbol{u}_h(t) = \delta_{hp}\boldsymbol{\mu}(t) - k_o\boldsymbol{\upsilon}_h(t)$. When Assumption 1 holds, by invoking Lemma 2, with $\bar{u}_h = k_o\bar{\upsilon}_h$ ($\forall h = p, p-1, \ldots, 1, 0$), boundedness of the estimation errors follows:

$$\|\widetilde{\boldsymbol{\sigma}}_h(t)\| \leq \sum_{l=h}^{p} t^{l-h}e^{-\alpha_o t}\bar{\sigma}_l(0) + k_o\sum_{l=h}^{p}\bar{\beta}_l\bar{\upsilon}_l, \tag{23}$$

where $\bar{\beta}_l = (\rho_o/\alpha_o)^{l-h+1}$. As expected, convergence of estimation errors is not achieved, due to the additional term $k_o\sum_{l=h}^{p}\bar{\beta}_l\bar{\upsilon}_l$ in (23), accounting for the effect of noise and disturbances. Notice that $\bar{\beta}_l$ decreases when h approaches 0, i.e., proceeding along the chain of estimators, hence mitigating the effect of noise and disturbances on the estimates. On the other hand, when Assumption 2 holds, since it is $\bar{u}_p = \sqrt{N_u}\bar{\sigma}_{p+1} + k_o\bar{\upsilon}_p$ and $\bar{u}_h = k_o\bar{\upsilon}_h$ (for $h < p$), the upper bound on $\|\widetilde{\boldsymbol{\sigma}}_h(t)\|$ contains the additional term $\bar{\beta}_p\sqrt{N_u}\bar{\sigma}_{p+1}$ with respect to that in (23).

It is worth noticing that system (21) is characterized by the transfer matrix ($s \in \mathbb{C}$ denotes the complex variable)

$$\boldsymbol{\Phi}_o(s) = \left(s\boldsymbol{I}_{m_u} + \bar{\boldsymbol{L}}_o\right)^{-1}, \tag{24}$$

having poles all strictly negative. Thus, it represents the transfer matrix of a proper stable low-pass filter. Due to the cascaded structure of the estimation filters, $k_o\boldsymbol{\upsilon}_h$ affects $\widetilde{\boldsymbol{\sigma}}_h$ after being filtered by the MIMO low-pass filter with transfer matrix $\boldsymbol{\Phi}_o(s)$, while it affects $\widetilde{\boldsymbol{\sigma}}_l$ (for $l < h$) after being filtered by $h - l + 1$ identical MIMO low-pass filter with transfer matrix $\boldsymbol{\Phi}_o(s)$. In other words, the filtering effect is strengthened by proceeding along the chain of estimators, thanks to the cascade of low-pass filtering of the upstream signals. On the other hand, by increasing k_o, the bandwidth of each filter is increased and, thus, the filtering effect of high-frequency noise components is weakened.

3.3. Extension to switching topology

In several applications, e.g. networks of moving agents, existing communication links can disappear and/or new links between agents can be created. In terms of the network topology, this implies that some edges are added or removed from the graph at some, unpredictable, times. Thus, in such cases, the network is described by a *switching topology*, i.e., a finite collection of K graphs with N nodes, $\Gamma = \{\mathcal{G}_1, \ldots, \mathcal{G}_K\}$. Each graph in Γ is characterized by its adjacency matrix, $\boldsymbol{A}_k$, and its Laplacian matrix, $\boldsymbol{L}_k$, with $k \in \mathcal{K} = \{1, \ldots, K\}$. For any $\boldsymbol{L}_k$, the corresponding matrix, $\bar{\boldsymbol{L}}_k$, extracted from $\boldsymbol{L}_k$ by deleting the rows and the columns corresponding to the source nodes in $\mathcal{G}_k \in \Gamma$, can be defined. The switching rule, which determines, at each time, the index of the active topology in Γ, can be expressed via the scalar piecewise constant function of time $s(t) : t \in [t_0, +\infty) \rightarrow \mathcal{K}$, where t_0 is the initial time. Thus, all the above defined matrices can be modelled as piecewise constant functions of time: $\boldsymbol{A}_{s(t)} \in \{\boldsymbol{A}_1, \ldots, \boldsymbol{A}_K\}$, $\boldsymbol{L}_{s(t)} \in \{\boldsymbol{L}_1, \ldots, \boldsymbol{L}_K\}$ and $\bar{\boldsymbol{L}}_{s(t)} \in \left\{\bar{\boldsymbol{L}}_1, \ldots, \bar{\boldsymbol{L}}_K\right\}$.

The estimation error dynamics (13) can be written, $\forall\ h = p, p-1, \ldots, 1, 0$, as

$$\dot{\tilde{\boldsymbol{\sigma}}}_h(t) = -\bar{\boldsymbol{L}}_{o,s(t)}\tilde{\boldsymbol{\sigma}}_h(t) + \delta_{hp}\boldsymbol{\mu}(t) + \bar{\delta}_{hp}\tilde{\boldsymbol{\sigma}}_{h+1}(t), \tag{25}$$

where the constant matrix $\bar{\boldsymbol{L}}_o$ is replaced by the time-varying piece-wise constant matrix $\bar{\boldsymbol{L}}_{o,s(t)} = k_o\left(\bar{\boldsymbol{L}}_{s(t)} \otimes \boldsymbol{I}_m\right)$. If each $\mathcal{G}_k \in \Gamma$ is SGC and balanced, by using the same arguments in Olfati-Saber and Murray (2004) and Antonelli et al. (2014), it can be shown that the autonomous counterpart of (25),

$$\dot{\tilde{\boldsymbol{\sigma}}}_h(t) = -\bar{\boldsymbol{L}}_{o,s(t)}\tilde{\boldsymbol{\sigma}}_h(t), \tag{26}$$

is a globally exponentially stable system and, thus, there exist two constants, $\rho_o > 0$ and $\alpha_o > 0$, such that the state transition matrix $\boldsymbol{\Phi}_o(t, t_0) = \exp\left(-\bar{\boldsymbol{L}}_{o,s(t)}(t-t_0)\right)$ can be bounded, for any $t_0 \in \mathbb{R}$, as follows (Khalil, 1996)

$$\|\boldsymbol{\Phi}_o(t, t_0)\| \le \rho_o\, e^{-\alpha_o(t-t_0)}, \quad \forall t \ge t_0. \tag{27}$$

Therefore, under Assumption 1, each system defined by (25) is either an autonomous exponentially stable linear time-varying system (for $h = p$) or an exponentially stable linear time-varying system with the asymptotically vanishing input $\tilde{\boldsymbol{\sigma}}_{h+1}(t)$ (for $h < p$), and, thus, Theorem 2 can be extended to the case of switching topology as well, with the additional assumption that each graph in Γ is balanced. Similarly, when Assumption 2 holds, looking at the proof of Lemma 2, it can be easily recognized that inequality (27) allows to extend Theorem 3 to the case of switching topology, with the additional assumption that each graph in Γ is balanced.

The assumptions that each graph belonging to Γ is SGC and balanced might be rather restrictive in some application scenarios and can be relaxed, as proposed in Jadbabaie, Lin, and Morse (2003) and Ren and Beard (2005) for consensus problems. However, if in Γ there are graphs that do not satisfy the SGC property, it can be shown, via simple counterexamples, that exponential stability is not guaranteed for arbitrary switching sequences. Hence, the assumptions on Γ can be relaxed if some restrictions on the switching rule are considered. In detail, assume that at least one of the graphs in Γ is SGC (but not necessarily balanced). Notice that, when a graph $\mathcal{G}_k \in \Gamma$ is not SGC, according to Theorem 1, the corresponding matrix $-\bar{\boldsymbol{L}}_{o,k} = -k_o\left(\bar{\boldsymbol{L}}_k \otimes \boldsymbol{I}_m\right)$ is not Hurwitz, as it admits at least m null eigenvalues. Then, the set $\bar{\mathcal{L}} = \left\{-\bar{\boldsymbol{L}}_{o,1}, \ldots, -\bar{\boldsymbol{L}}_{o,K}\right\}$ can be partitioned in the two disjoint subsets, $\bar{\mathcal{L}}_H \neq \emptyset$ and $\bar{\mathcal{L}}_N$, collecting, respectively, the matrices $-\bar{\boldsymbol{L}}_{o,k} \in \bar{\mathcal{L}}$ that are Hurwitz (corresponding to graphs satisfying the SGC property) and not Hurwitz (corresponding to graphs not satisfying the SGC property); the index set $\mathcal{K}$ can be partitioned accordingly, i.e., $\mathcal{K} = \mathcal{K}_H \cup \mathcal{K}_N$. According to (15), for any $k \in \mathcal{K}$ and any $\epsilon_{o,k} \ge 0$, there exists a constant $\rho_{o,k} \ge 1$ such that

$$\left\|\exp\left(-\bar{\boldsymbol{L}}_{o,k}t\right)\right\| \le \rho_{o,k}\, e^{-\alpha_{o,k}t}, \quad \forall t \ge t_0,$$

where $\epsilon_{o,k} \ge 0$ can be chosen such that $-\alpha_{o,k} = \lambda_M\left(-\bar{\boldsymbol{L}}_{o,k}\right) + \epsilon_{o,k}$ is negative, if $k \in \mathcal{K}_H$, while it is non-negative, if $k \in \mathcal{K}_N$. Let $a = \max_{k\in\mathcal{K}}\left\{\ln\left(\rho_{o,k}\right)\right\} \ge 0$, $\alpha_H = \max_{k\in\mathcal{K}_H}\left\{-\alpha_{o,k}\right\} < 0$ and $\alpha_N = \max_{k\in\mathcal{K}_N}\left\{-\alpha_{o,k}\right\} \ge 0$. Let the integer $S(t_0, t)$ denote the number of times $s(t)$ switches over the interval $[t_0, t)$ and $T_H(t_0, t)$ (respectively, $T_N(t_0, t)$) the total activation time of the SGC (respectively, not SGC) topologies in the interval $[t_0, t)$. Assume that there exists a given $\alpha' \in (0, \alpha_H)$, such that

$$\frac{T_H(t_0, t)}{T_N(t_0, t)} \ge \frac{\alpha_N + \alpha'}{\alpha_H - \alpha'}, \quad \forall t \ge t_0, \tag{28}$$

and, for some $S_0 \ge 0$ and $\alpha \in \left(0, \alpha'\right)$, the switching signal $s(t)$ satisfies

$$S(t_0, t) \le S_0 + \frac{t - t_0}{\tau'_d}, \quad \forall t \ge t_0, \tag{29}$$

where $\tau'_d = a/(\alpha' - \alpha)$, which can be interpreted as an *average dwell time* (i.e., the average interval between consecutive switching events) and S_0 is the so-called *chatter bound* (Hespanha & Morse, 1999; Zhai, Hu, Yasuda, & Michel, 2001). Then, by using the same arguments in Zhai et al. (2001), it can be shown that the system (26) is globally exponentially stable, i.e., letting $\rho_o = e^{a(S_0+1)}$ and $\alpha_o = \alpha$, (27) holds for all $s(t)$ satisfying (29), for any average dwell time $\tau_d \ge \tau'_d$ and any $S_0 \ge 0$. Therefore, Theorems 2 and 3 can be extended to the case of switching topology, under the assumptions that at least a graph in Γ is SGC and conditions (28), (29) hold. In the case $\bar{\mathcal{L}}_N = \emptyset$, i.e., all the graphs in Γ are SGC, but not necessarily balanced, condition (28) is trivially satisfied, as $T_N(t_0, t) = 0$, and, thus, only condition (29) has to be satisfied.

4. Simulation results

In order to validate the proposed estimation scheme, numerical simulations are carried out. In all the case studies, the estimators (11) are implemented via a discrete-time first-order approximation with a fixed time step of $T = 0.005$ s.

4.1. First case study

A multi-agent system composed of $N = 6$ agents (2 sources and 4 users) is in charge of estimating a scalar 3rd-order polynomial $\sigma(t)$ (i.e., $m = 1$, $p = 3$). Since $\sigma^{(4)}(t) = 0$, the signal satisfies Assumption 1 with $p+1 = 4$; thus, 4 decentralized estimators are adopted. The communication graph, which is SGC but not SC, is reported in Fig. 1. The agents 1 and 6 are the sources. A white noise has been superimposed to $\sigma(t)$, with standard deviation equal to 0.05, while the noise superimposed to its non-null time derivatives of order $h = 1, 2, 3$ has standard deviation equal to 0.1, 0.15 and 0.2, respectively. The time histories of the signal and its non-null time derivatives are reported in Fig. 2. The gain of the estimation filters has been set to $k_o = 10$.

Fig. 3 reports the estimation errors for the users. As expected, all the errors are exponentially decaying to zero and the effect of noise is progressively reduced as h decreases. The filtering capabilities of the estimation scheme are confirmed by the results in Fig. 4, which reports the singular values of the transfer matrix $\boldsymbol{\Phi}_o(s)$.

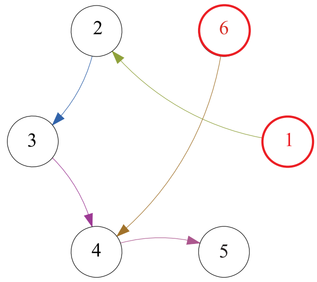

Fig. 1. First case study: topology of the communication graph. The nodes in red are the sources, the nodes in black are the users. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

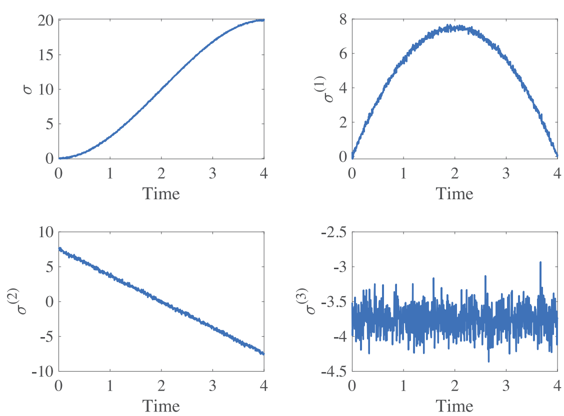

Fig. 2. First case study: signal $\sigma(t)$ and its time derivatives.

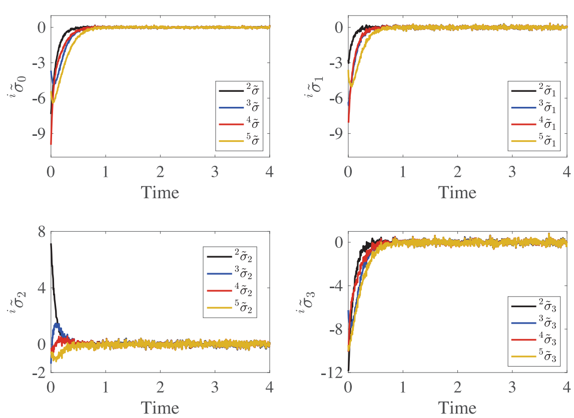

Fig. 3. First case study: estimation errors for the user agents.

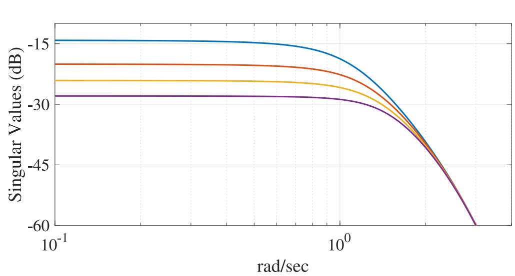

Fig. 4. First case study: singular values of the transfer matrix $\boldsymbol{\Phi}_o(s)$.

Table 1
Root mean square estimation errors for agent 2.

	$h = 1$	$h = 0$
$k_o = 5$	$2.73 \cdot 10^{-1}$	$5.45 \cdot 10^{-2}$
$k_o = 10$	$1.38 \cdot 10^{-1}$	$1.41 \cdot 10^{-2}$
$k_o = 50$	$2.17 \cdot 10^{-2}$	$6.23 \cdot 10^{-4}$
$k_o = 100$	$1.38 \cdot 10^{-2}$	$1.73 \cdot 10^{-4}$

4.2. Second case study

In the second case study, a different system of $N = 5$ agents is in charge of estimating a new scalar signal $\sigma(t)$ given by the sum of a 2nd-order polynomial and a sinusoidal function with frequency of 2π Hz. Since $\ddot{\sigma}(t) \leq \bar{\sigma}_2 = 2$, but non-null, only Assumption 2 is satisfied. The time histories of the signal and its time derivatives are reported in Fig. 5. The decentralized estimators have been designed by assuming $p = 1$. Agent 1 is the only source in the system. The communication network is represented by the SC graph in Fig. 6.

In Fig. 7 the estimation errors for the user agent with index 2, for different values of the gain k_o are reported. It can be noticed that, as expected, the errors do not converge to zero but are only bounded; it can be recognized that the effect of $\ddot{\sigma}(t)$ is reduced when h decreases and k_o increases, as confirmed by the results in Table 1, which reports the root mean square errors of the two estimators run by agent 2 for different values of k_o. The behaviour of the other users' estimators is similar and is not reported here for the sake of compactness.

The proposed estimation scheme is compared to a different approach, based on the containment control scheme developed in Liu et al. (2012) with dynamic leaders. Namely, the source agents play the role of leaders, the state of each coincides with the signal $\sigma(t)$. The user agents play the role of followers. According to Liu et al. (2012), assuming that $\mathcal{G}$ contains a directed spanning forest, if each user ($i \in \mathcal{N}_u$) runs the following estimation filter

$$\begin{cases} {}^i\dot{\hat{\sigma}}_0(t) = {}^i\hat{\sigma}_1(t) \\ {}^i\dot{\hat{\sigma}}_1(t) = k_{o,0} \sum_{j \in \mathcal{N}_i} \left({}^j\hat{\sigma}_0(t) - {}^i\hat{\sigma}_0(t)\right) + \\ \qquad\qquad k_{o,1} \sum_{j \in \mathcal{N}_i} \left({}^j\hat{\sigma}_1(t) - {}^i\hat{\sigma}_1(t)\right), \end{cases} \tag{30}$$

then, for $p \leq 1$ and under suitable assumptions on the positive gains $k_{o,0}$ and $k_{o,1}$, ${}^i\hat{\sigma}_0(t)$ asymptotically converges to the convex hull spanned by the sources (leaders), which, of course, coincides with $\sigma(t)$. In order to carry out a fair comparison, $k_{o,0}$ and $k_{o,1}$ have been chosen equal to k_o (i.e., $k_o = k_{o,0} = k_{o,1} = 50$). The results obtained by adopting the estimation filter (30), implemented via the sampled-data protocol adopted in Liu et al. (2012), are reported in Fig. 8. It can be recognized that both schemes lead to bounded errors on $\dot{\sigma}(t)$ of the same order of magnitude. However, the proposed method allows to obtain an estimation of $\sigma(t)$ much more accurate, since the effect of $\ddot{\sigma}(t)$ on the estimation errors is reduced thanks to the cascaded structure of the estimation scheme. Moreover, it must be noticed that (30) does not lead to convergence of the estimation errors when $p > 1$ (even for signals with null $p + 1$th derivative).

4.3. Third case study

In this case study, a system of $N = 5$ agents is in charge of estimating the signal $\sigma(t)$ in Fig. 2. The directed communication graph switches among the topologies in Fig. 9: at time $t_1 = 0.5$ s the graph switches from the SC and balanced topology in Fig. 9a to that in Fig. 9b, which is not SGC, at $t_2 = 1.5$ s it switches again to the topology in Fig. 9c, neither SC nor balanced, but SCG. It can be verified that there exist α', S_0 and τ'_d such that conditions (28)

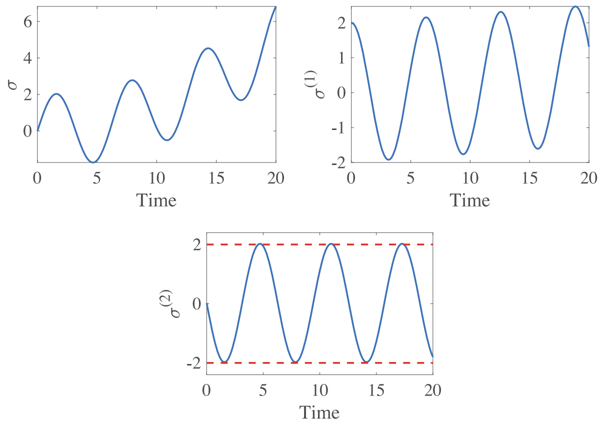

Fig. 5. Second case study: signal $\sigma(t)$ and its time derivatives.

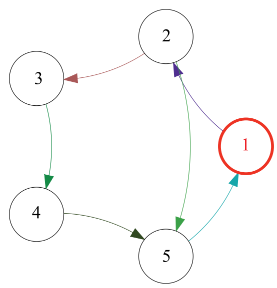

Fig. 6. Second case study: topology of the communication graph. The node in red represents the source, the nodes in black represent the users. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

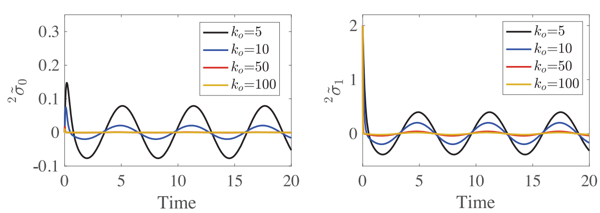

Fig. 7. Second case study: estimation errors for agent 2 with different values of the gain k_o. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

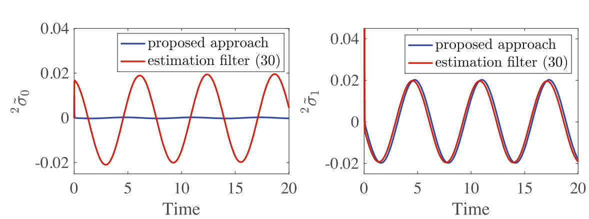

Fig. 8. Second case study: comparison between the estimation errors for agent 2 obtained by using the approach proposed in this paper (blue line) and the estimation filter in Eq. (30) proposed in Liu et al. (2012) (red line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and (29) are satisfied for all $t \geq 0$. The gain of the estimation filters has been set to $k_o = 10$.

Fig. 10 reports the estimation errors for the users. It can be noticed that, when the non-SCG topology is active, the errors of the fourth agent drift away. On the same way, also the estimates

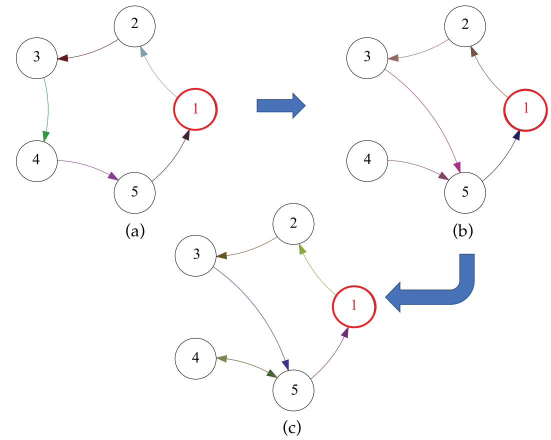

Fig. 9. Third case study: topologies of the communication graphs: (a) SC topology in the time interval [0,0.5]; (b) non-SCG topology in the time interval (0.5,1.5]; (c) SCG topology in the time interval (1.5,4]. The node in red represents the source, the nodes in black represent the users. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

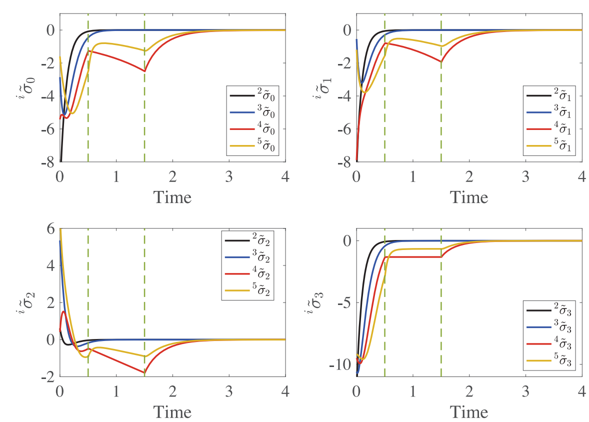

Fig. 10. Third case study: estimation errors for the user agents. The dashed lines represent the instant of topology switching. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

of agent 5 are affected by errors, due to the fact that the estimates computed by agent 4 act as a disturbance on the error dynamics for agent 5. However, after the last switch, all the errors converge exponentially to zero.

4.4. Fourth case study

In this case study, a system composed of $N = 50$ agents is considered. Two different topologies are considered for the communication graph: a simple directed cycle and an undirected SC graph. Then, for each topology, starting from a configuration in which only one node plays the role of source, at each step a new set of sources is selected (such that $N_s = |\mathcal{N}_s|$ is increased by one unit) and the corresponding spectral abscissa, $\lambda_M(-\bar{\boldsymbol{L}})$, is computed. In a first numerical experiment (Fig. 11, top), the new set of sources is selected randomly among all the agents: it can be recognized that, in the case of the undirected SC graph, $\lambda_M(-\bar{\boldsymbol{L}})$ does not decrease monotonically with N_s. Noticeably, for the cycle graph, the subgraph $\mathcal{G}_s$ is never SC and $\lambda_M(-\bar{\boldsymbol{L}})$ is constant, regardless the number and distribution of sources. On the other hand, if, at each step, the new set of sources is obtained

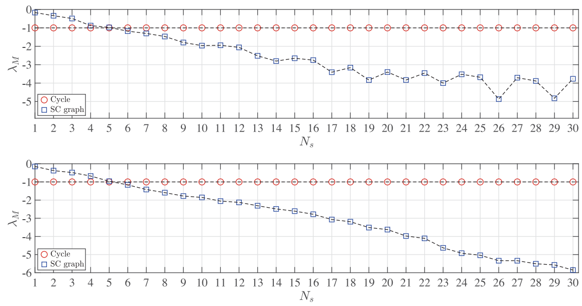

Fig. 11. Fourth case study: spectral abscissa of $-\bar{\boldsymbol{L}}$ for a cycle (red circles) and a SC undirected graph (blue squares) as a function of the number of sources. Top: new sources are randomly selected among all agents. Bottom: new sources are randomly selected among only user agents.

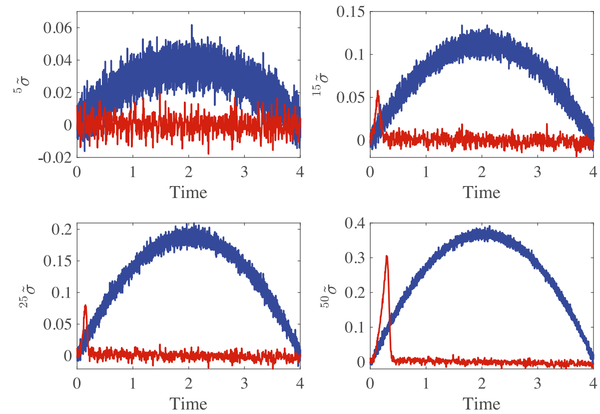

Fig. 12. Fourth case study: estimation errors of the signal $\sigma(t)$ for 4 different user agents in the presence of one source (red line) compared with the estimation error obtained with a simple flooding approach (blue line). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

by adding to the old one a randomly selected user agent, such that the assumptions of Proposition 1 are satisfied, for the undirected SC graph, $\lambda_M(-\bar{\boldsymbol{L}})$ is a strictly decreasing function of N_s (Fig. 11, bottom), except in one case (i.e., when N_s goes from 26 to 27) in which the subgraph $\mathcal{G}_s$ is not SC. On the other hand, again, in the case of the cycle, $\lambda_M(-\bar{\boldsymbol{L}})$ remains constant.

Then, the proposed estimation method is compared with a simple flooding approach, where, starting from the sources, each node transmits $\sigma(t)$ to its neighbours. The chosen graph topology is the directed cycle with a single source, which is a quite critical configuration when transmission delays are present. In order to carry out a fair comparison, the information exchange between an agent and its neighbours is characterized by a transmission delay of $\tau_t = 0.001$ s, while the estimation filters (11) are run with a time step of $T = 0.005$ s (assuming that the computation of the output takes a time $T_c = T - \tau_t$) and a gain $k_o = 150$. The signal $\sigma(t)$ to be estimated is the same 3rd-order polynomial considered in the first case study. The results are shown in Fig. 12, reporting the estimation errors for $\sigma(t)$ of four different user agents. It can be recognized that the longer is the path from the source to the user, the larger is the peak estimation error for both the approaches. However, for the proposed estimation scheme, after an initial peak in the transient phase, the error converges to zero. On the other hand, for the flooding approach, the estimation errors do not exhibit a convergent behaviour, as they are not determined by an asymptotically stable dynamics, and are characterized by larger root mean square values, which depend on the transmission delays and the derivatives of $\sigma(t)$; moreover, the noise is transmitted unaltered to all the users. In sum, the results confirm that flooding approaches perform poorly for large networks with a relatively small number of sources.

5. Conclusions and future work

In this paper, a solution to a distributed estimation problem for multi-agent networked systems has been developed and analysed. The approach is aimed at estimating a time-varying vector signal known by only a subset of *source* agents. The proposed estimation scheme is based on a bank of cascaded estimators run by the agents which do not have access to the signal (the *user* agents). Convergence properties of the estimation errors are investigated by assuming that the inter-agent communication graph satisfies a suitable connectivity property, weaker than strong connectivity, and under two different assumptions on the signal to be estimated. Simulations have been carried out to assess the performance of the proposed estimation scheme. The developed estimation approach can be conveniently adopted in different application fields. In fact, ongoing work is focused on the use of the estimation scheme in the framework of centroid and formation control of groups of mobile robots.

Appendix A. Proof of Lemma 2

For $h = p$, Eq. (1) describes the dynamics of a linear and time-invariant system with the bounded input $\boldsymbol{u}_p(t)$. Thus, $\boldsymbol{x}_p(t)$ can be expressed as

$$\boldsymbol{x}_p(t) = e^{-\boldsymbol{A}t}\boldsymbol{x}_p(0) + \int_0^t e^{-\boldsymbol{A}(t-\tau)}\boldsymbol{B}_p\boldsymbol{u}_p(\tau)d\tau .$$

By taking into account (2), the following chain of inequalities can be devised

$$\begin{aligned}\left\|\boldsymbol{x}_p(t)\right\| &\le \rho e^{-\alpha t}\left\|\boldsymbol{x}_p(0)\right\| + b_p\rho\bar{u}_p\int_0^t e^{-\alpha(t-\tau)}d\tau \\ &\le \rho e^{-\alpha t}\left\|\boldsymbol{x}_p(0)\right\| + b_p\frac{\rho}{\alpha}\bar{u}_p,\end{aligned}$$

i.e., $\left\|\boldsymbol{x}_p(t)\right\|$ can be upper bounded by the function $\bar{x}_p(t) = \rho e^{-\alpha t}\left\|\boldsymbol{x}_p(0)\right\| + b_p\left(\rho/\alpha\right)\bar{u}_p$.

For $h = p - 1$, Eq. (1) describes a linear and time-invariant system with the bounded inputs $\boldsymbol{u}_{p-1}(t)$ and $\boldsymbol{x}_p(t)$. Thus, $\boldsymbol{x}_{p-1}(t)$ can be expressed as

$$\begin{aligned}\boldsymbol{x}_{p-1}(t) = e^{-\boldsymbol{A}t}\boldsymbol{x}_{p-1}(0) + \\ \int_0^t e^{-\boldsymbol{A}(t-\tau)}\left(\boldsymbol{B}_{p-1}\boldsymbol{u}_{p-1}(\tau) + \boldsymbol{x}_p(\tau)\right)d\tau ,\end{aligned}$$

which, by virtue of (2) and the bound on $\boldsymbol{x}_p(t)$ computed above, leads to the following chain of inequalities

$$\begin{aligned}\left\|\boldsymbol{x}_{p-1}(t)\right\| &\le \rho e^{-\alpha t}\left\|\boldsymbol{x}_{p-1}(0)\right\| + \\ &\quad b_{p-1}\frac{\rho}{\alpha}\bar{u}_{p-1} + \rho\int_0^t e^{-\alpha(t-\tau)}\bar{x}_p(\tau)d\tau \\ &\le \rho e^{-\alpha t}\left\|\boldsymbol{x}_{p-1}(0)\right\| + \rho^2 te^{-\alpha t}\left\|\boldsymbol{x}_p(0)\right\| \\ &\quad b_{p-1}\frac{\rho}{\alpha}\bar{u}_{p-1} + b_p\left(\frac{\rho}{\alpha}\right)^2\bar{u}_p,\end{aligned}$$

i.e., $\left\|\boldsymbol{x}_{p-1}(t)\right\|$ can be upper bounded by the function $\bar{x}_{p-1}(t) = \rho e^{-\alpha t}\left\|\boldsymbol{x}_{p-1}(0)\right\| + te^{-\alpha t}\rho^2\left\|\boldsymbol{x}_p(0)\right\| + b_{p-1}\left(\rho/\alpha\right)\bar{u}_{p-1} + b_p\left(\rho/\alpha\right)^2\bar{u}_p$.

The same argument can be adopted for any $h < p$, since

$$\boldsymbol{x}_h(t) = e^{-\boldsymbol{A}t}\boldsymbol{x}_h(0) + \int_0^t e^{-\boldsymbol{A}(t-\tau)}(\boldsymbol{B}_h\boldsymbol{u}_h(\tau) + \boldsymbol{x}_{h+1}(\tau))\,d\tau ,$$

and

$$\|\boldsymbol{x}_h(t)\| \le \rho e^{-\alpha t}\|\boldsymbol{x}_h(0)\| + b_h\frac{\rho}{\alpha}\bar{u}_h + \rho\int_0^t e^{-\alpha(t-\tau)}\bar{x}_{h+1}(\tau)d\tau,$$

from which (3) and (4) can be devised by induction.

Appendix B. Proof of Theorem 1

It is worth reporting the following result on diagonally dominant matrices (Shivakumar & Chew, 1974), which will be exploited in the proof:

Theorem 4. *A complex $(N \times N)$ matrix $\boldsymbol{M} = \{m_{ij}\}_{i,j\in\mathcal{N}}$ is non-singular if:*

- *$\boldsymbol{M}$ is diagonally dominant: $|m_{ii}| \ge \sum_{j\neq i}|m_{ij}|$, $\forall i \in \mathcal{N}$;*
- *there exists an index set $\mathcal{I} \subseteq \mathcal{N}$ such that $|m_{ii}| > \sum_{j\neq i}|m_{ij}|$, $\forall i \in \mathcal{I}$;*
- *for each $i \notin \mathcal{I}$, there exists a sequence of indexes $i_1, i_2, \ldots, i_q$, such that $m_{ii_1}, m_{i_1i_2}, \ldots, m_{i_{q-1}i_q}$ are all non-null and $i_q \in \mathcal{I}$.*

Let $r_i \in \mathcal{N}_b$ denote the index of the ith node in $\mathcal{N}_b$, where $i \in \bar{\mathcal{N}}_b = \{1, \ldots, N_b\}$. Since the graph is assumed to satisfy the SGC property, the set of indexes

$$\mathcal{N}_{ba} = \left\{r_i \in \mathcal{N}_b : N_{a,r_i} > 0\right\}, \tag{B.1}$$

is non-empty, i.e., there exists at least one node $r_i \in \mathcal{N}_b$ which has a parent node belonging to $\mathcal{N}_a$. On the other hand, the set $\mathcal{N}_{bb} = \mathcal{N}_b \setminus \mathcal{N}_{ba}$ is empty when $\mathcal{N}_b = \mathcal{N}_{ba}$, i.e., all nodes in $\mathcal{N}_b$ have at least one parent node in $\mathcal{N}_a$. Let $\bar{\mathcal{N}}_{ba} = \left\{i \in \bar{\mathcal{N}}_b : r_i \in \mathcal{N}_{ba}\right\}$, $\bar{\mathcal{N}}_{bb} = \left\{i \in \bar{\mathcal{N}}_b : r_i \in \mathcal{N}_{bb}\right\}$, $N_{ba} = \left|\bar{\mathcal{N}}_{ba}\right| = |\mathcal{N}_{ba}|$ and $N_{bb} = \left|\bar{\mathcal{N}}_{bb}\right| = |\mathcal{N}_{bb}|$. Then, since $\boldsymbol{L}_b$ is a Laplacian matrix, in view of (6) and (7), matrix $\bar{\boldsymbol{L}} = \left\{\bar{l}_{ij}\right\}_{i,j\in\bar{\mathcal{N}}_b}$ is diagonally dominant and

$$|\bar{l}_{ii}| > \sum_{j\neq i}|\bar{l}_{ij}|, \qquad \forall i \in \bar{\mathcal{N}}_{ba}. \tag{B.2}$$

If $\bar{\mathcal{N}}_{ba} = \bar{\mathcal{N}}_b$, then $\bar{\boldsymbol{L}}$ is strictly dominant and, thus, it is a full-rank matrix. When $\bar{\mathcal{N}}_{ba} \subset \bar{\mathcal{N}}_b$ (i.e., $\bar{\mathcal{N}}_{bb}$ is not empty), if, starting from any row of $\bar{\boldsymbol{L}}$ with index $i_0 \in \bar{\mathcal{N}}_{bb}$ (i.e., $r_{i_0} \in \mathcal{N}_{bb}$), a sequence of indexes, $i_0, i_1, \ldots, i_p$, such that $\bar{l}_{i_0i_1}, \bar{l}_{i_1i_2}, \ldots, \bar{l}_{i_{q-1}i_q}$ are all non-null and $i_q \in \bar{\mathcal{N}}_{ba}$ (i.e., $r_{i_q} \in \mathcal{N}_{ba}$) exists, then, according to Theorem 4, $\bar{\boldsymbol{L}}$ is full rank. Given the definition of graph Laplacian, such a sequence exists if and only if there exists a directed path in $\mathcal{G}_b$, with edges $\left(r_{i_q}, r_{i_{q-1}}\right), \ldots, \left(r_{i_2}, r_{i_1}\right), \left(r_{i_1}, r_{i_0}\right)$, connecting some node $r_{i_q} \in \mathcal{N}_{ba}$ to the node $r_{i_0} \in \mathcal{N}_{bb}$. Since $\mathcal{G}$ is assumed to be SGC, such a path exists. In fact, since $r_{i_0} \in \mathcal{N}_{bb}$, by definition, it cannot be child of a node in $\mathcal{N}_a$; however, according to the SGC property, it must be reached by a path with origin in $\mathcal{N}_a$. Thus, r_{i_0} must be necessarily reached by a path in $\mathcal{N}_b$ originated from a node $r_{i_q} \in \mathcal{N}_{ba}$. Hence, $\bar{\boldsymbol{L}}$ is a full-rank matrix even in the case $\bar{\mathcal{N}}_{ba} \subset \bar{\mathcal{N}}_b$.

According to the Geršgorin disc theorem (Horn & Johnson, 1990), the eigenvalues of $-\bar{\boldsymbol{L}}$ are located in the union of the N_b discs centred in $-\bar{l}_{ii}$ with radii $d_i = \sum_{j\neq i}|\bar{l}_{ij}|$ $(i \in \mathcal{N}_b)$. Since $\bar{l}_{ii} \in \mathbb{R}$ and $\bar{l}_{ii} > 0$, $\forall i \in \mathcal{N}_b$, the discs are all centred on the negative real axis and, due to diagonal dominance, are contained in the left-half complex plane and cannot intersect the imaginary axis but in the origin. On the other hand, since $-\bar{\boldsymbol{L}}$ is full rank, and thus cannot admit null eigenvalues, all its eigenvalues have strictly negative real parts. This proves the first statement of the Theorem.

The Geršgorin disc theorem implies that all the eigenvalues of $-\bar{\boldsymbol{L}}$ have non-positive real parts. Thus, in order to prove the second statement of the Theorem, it must be shown that at least one of its eigenvalues is null when $\mathcal{G}$ is not SGC. To this aim, notice that $\mathcal{G}$ is not SGC if at least one of the following conditions holds:

(i) $\mathcal{N}_{ba} = \emptyset$;
(ii) any node in $\mathcal{N}_{bb}$ has not parents in $\mathcal{N}_{ba}$;
(iii) at least one node in $\mathcal{N}_{bb}$ has not parents.

In the first case, since $\boldsymbol{D}_{ba} = \boldsymbol{O}_{N_b}$, it is $-\bar{\boldsymbol{L}} = -\boldsymbol{L}_b$ and, thus, at least one of its eigenvalues is null, being $\boldsymbol{L}_b$ a Laplacian. In the second case, if $i \in \bar{\mathcal{N}}_{bb}$, it is $\bar{l}_{ij} = 0$ for any $j \in \bar{\mathcal{N}}_{ba}$. Hence, it is always possible, through permutation of rows and columns of $\bar{\boldsymbol{L}}$, rewrite $\bar{\boldsymbol{L}}$ such that the first N_{ba} rows (columns) correspond to the nodes in $\mathcal{N}_{ba}$ and the last N_{bb} rows (columns) correspond to the nodes in $\mathcal{N}_{bb}$. The resulting matrix is block-triangular and the lower-right $(N_{bb} \times N_{bb})$ block is the Laplacian of the subgraph $\mathcal{G}_{bb}\left(\mathcal{E} \cap (\mathcal{N}_{bb} \times \mathcal{N}_{bb}), \mathcal{N}_{bb}\right)$. Hence, this block (and $\bar{\boldsymbol{L}}$) admits at least a null eigenvalue. Finally, in the third case, the rows in $\bar{\boldsymbol{L}}$ corresponding to the nodes without parents are null; thus, also in this case at least one eigenvalue of $-\bar{\boldsymbol{L}}$ is null.

Appendix C. Proof of Proposition 1

The proof is based on the following property of the spectral abscissa of Metzler matrices (Bullo, 2022, Ch. 10):

Lemma 4. *Let $\boldsymbol{A}$ and $\boldsymbol{B}$ be Metzler matrices of the same dimensions:*

- *if $\boldsymbol{A} \ge \boldsymbol{B}$, then $\lambda_M(\boldsymbol{A}) \ge \lambda_M(\boldsymbol{B})$;*
- *if, in addition, $\boldsymbol{A} \neq \boldsymbol{B}$ and $\boldsymbol{A}$ is irreducible, then $\lambda_M(\boldsymbol{A}) > \lambda_M(\boldsymbol{B})$,*

where the matrix inequality is understood component-wise.

First, notice that $-\bar{\boldsymbol{L}}'$ is a Metzler matrix. Let $N_b'' = |\mathcal{N}_b''|$ and $N_\delta = |\mathcal{N}_\delta|$. Then, since $\mathcal{N}_a'' = \mathcal{N}_a' \cup \mathcal{N}_\delta$ and $\mathcal{N}_b'' = \mathcal{N}_b' \setminus \mathcal{N}_\delta$, matrix $-\bar{\boldsymbol{L}}'$ can be written, eventually via a permutation of rows and columns, as follows:

$$-\bar{\boldsymbol{L}}' = \begin{bmatrix} -\bar{\boldsymbol{L}}'' & \boldsymbol{C} \\ \boldsymbol{R} & -\bar{\boldsymbol{L}}_\delta' \end{bmatrix} \ge -\check{\boldsymbol{L}}' = \begin{bmatrix} -\bar{\boldsymbol{L}}'' & \boldsymbol{O} \\ \boldsymbol{R} & -\bar{\boldsymbol{L}}_\delta' \end{bmatrix},$$

where the matrix inequality is understood component-wise, $\boldsymbol{R}$ and $\boldsymbol{C}$ are, respectively, $(N_\delta \times N_b'')$ and $(N_b'' \times N_\delta)$ non-negative matrices, $\boldsymbol{O}$ is the $(N_b'' \times N_\delta)$ null matrix and $-\bar{\boldsymbol{L}}_\delta'$ is the $(N_\delta \times N_\delta)$ matrix collecting the elements $-\bar{l}_{ij}'$ for $i, j \in \mathcal{N}_\delta$. Then, the first part of the claim follows directly from Lemma 4, by noting that $-\check{\boldsymbol{L}}'$ is Metzler and $\lambda_M(-\check{\boldsymbol{L}}') = \max\{\lambda_M(-\bar{\boldsymbol{L}}_\delta'), \lambda_M(-\bar{\boldsymbol{L}}'')\}$. The second part follows from the fact that, if $\mathcal{G}_b'$ is SC, $\bar{\boldsymbol{L}}'$ is irreducible (since it can expressed as the sum of an irreducible matrix, the Laplacian of $\mathcal{G}_b'$, and a diagonal matrix, $\boldsymbol{D}_{ba}'$) and $\bar{\boldsymbol{L}}' \neq \check{\boldsymbol{L}}'$ (since $\boldsymbol{C}$ is non-null, otherwise $\bar{\boldsymbol{L}}'$ would be in block-triangular form).

Appendix D. Proof of Proposition 2

The claim follows from Lemma 4, since both $-\bar{\boldsymbol{L}}'$ and $-\bar{\boldsymbol{L}}''$ are Metzler matrices and $-\bar{\boldsymbol{L}}' \ge -\bar{\boldsymbol{L}}''$, being $-\boldsymbol{D}_{ba}' \ge -\boldsymbol{D}_{ba}''$. On the other hand, if $\mathcal{G}_b$ is SC, then $\boldsymbol{L}_b$ is irreducible and, thus, $-\bar{\boldsymbol{L}}'$ and $-\bar{\boldsymbol{L}}''$ are both irreducible (each being the sum of an irreducible matrix and a diagonal matrix), this implies that inequality (9) holds strictly if $\boldsymbol{D}_{ba}' \neq \boldsymbol{D}_{ba}''$.

References

Akyildiz, I. F., Su, Weilian, Sankarasubramaniam, Y., & Cayirci, E. (2002). A survey on sensor networks. *IEEE Communications Magazine*, *40*(8), 102–114.
Antonelli, G., Arrichiello, F., Caccavale, F., & Marino, A. (2014). Decentralized time-varying formation control for multi-robot systems. *International Journal of Robotics Research*, *33*, 1029–1043,.
Arrichiello, F., Marino, A., & Pierri, F. (2015). Observer-based decentralized fault detection and isolation strategy for networked multirobot systems. *IEEE Transactions on Control Systems Technology*, *23*(4), 1465–1476.
Brewer, J. W. (1978). Kronecker products and matrix calculus in system theory. *IEEE Transactions on Circuits and Systems*, *25*(9), 772–781.
Bullo, F. (2022). *Lectures on network systems* (1.6 ed.). Kindle Direct Publishing.

Calafiore, G. C., & Abrate, F. (2009). Distributed linear estimation over sensor networks. *International Journal of Control*, *82*(5), 868–882.
Cao, Y., Ren, W., & Egerstedt, M. (2012). Distributed containment control with multiple stationary or dynamic leaders in fixed and switching directed networks. *Automatica*, *48*(8), 1586–1597.
Carli, R., Chiuso, A., Schenato, L., & Zampieri, S. (2008). Distributed Kalman filtering based on consensus strategies. *IEEE Journal on Selected Areas in Communications*, *26*(4), 622–633.
Ge, X., Han, Q.-L., Ding, D., Zhang, X.-M., & Ning, B. (2018). A survey on recent advances in distributed sampled-data cooperative control of multi-agent systems. *Neurocomputing*, *275*, 1684–1701.
Hart, J. K., & Martinez, M. (2006). Environmental sensor networks: A revolution in the earth system science? *Earth-Science Reviews*, *78*(3), 177–191.
He, Jianping, Zhou, Mengjie, Cheng, Peng, Shi, Ling, & Chen, Jiming (2016). Consensus under bounded noise in discrete network systems: An algorithm with fast convergence and high accuracy. *IEEE Transactions on Cybernetics*, *46*(12), 2874–2884.
Hespanha, J. P., & Morse, A. S. (1999). Stability of switched systems with average dwell-time. In *Proc. of the 38th IEEE conference on decision and control* (pp. 2655–2660). Phoenix, AZ, USA.
Hong, Y., Chen, G., & Bushnell, L. (2008). Distributed observers design for leader-following control of multi-agent networks. *Automatica*, *44*(3), 846–850.
Horn, R. A., & Johnson, C. R. (1990). *Matrix analysis*. Cambridge Univ Pr.
Jadbabaie, A., Lin, J., & Morse, A. S. (2003). Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, *48*(6), 988–1001.
Kågström, Bo (1977). Bounds and perturbation bounds for the matrix exponential. *BIT Numerical Mathematics*, *17*(1), 39—57.
Khalil, H. K. (1996). *Nonlinear systems* (2nd ed.). Upper Saddle River, New Jersey: Prentice-Hall.
Kia, S. S., Van Scoy, B., Cortes, J., Freeman, R. A., Lynch, K. M., & Martinez, S. (2019). Tutorial on dynamic average consensus: The problem, its applications, and the algorithms. *IEEE Control Systems Magazine*, *39*(3), 40–72.
Li, C., Dong, H., Li, J., & Wang, F. (2019). Distributed Kalman filtering for sensor network with balanced topology. *Systems & Control Letters*, *131*, Article 104500.
Liu, H., Xie, G., & Wang, L. (2012). Necessary and sufficient conditions for containment control of networked multi-agent systems. *Automatica*, *48*(7), 1415–1422.
Marino, A., Pierri, F., & Arrichiello, F. (2017). Distributed fault detection isolation and accommodation for homogeneous networked discrete-time linear systems. *IEEE Transactions on Automatic Control*, *62*(9), 4840–4847.
Miao, Z., Liu, Y.-H., Wang, Y., Yi, G., & Fierro, R. (2018). Distributed estimation and control for leader-following formations of nonholonomic mobile robots. *IEEE Transactions on Automation Science and Engineering*, *15*(4), 1946–1954.
Olfati-Saber, R., & Jalalkamali, P. (2012). Coupled distributed estimation and control for mobile sensor networks. *IEEE Transactions on Automatic Control*, *57*(10), 2609–2614.
Olfati-Saber, R., & Murray, R. M. (2004). Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, *49*(9), 1520–1533.
Olfati-Saber, R., & Shamma, J. S. (2005). Consensus filters for sensor networks and distributed sensor fusion. In *Proc. of the 44th IEEE conference on decision and control* (pp. 6698–6703).
Ren, W., & Beard, R. W. (2005). Consensus seeking in multiagent systems under dynamically changing interaction topologies. *IEEE Transactions on Automatic Control*, *50*(5), 655–661.
Ren, W., Beard, R. W., & Atkins, E. M. (2007). Information consensus in multivehicle cooperative control. *IEEE Control Systems Magazine*, *27*(2), 71–82.
Ren, W., & Sorensen, N. (2008). Distributed coordination architecture for multi-robot formation control. *Robotics and Autonomous Systems*, *56*, 24–333.
Ren, Y., Sosnowski, S., & Hirche, S. (2020). Fully distributed cooperation for networked uncertain mobile manipulators. *IEEE Transactions on Robotics*, *36*(4), 984–1003.
Shivakumar, P. N., & Chew, K. H. (1974). A sufficient condition for nonvanishing of determinants. *Proceedings of the American Mathematical Society*, *43*(1), 63–66.
Smith, R. S., & Hadaegh, F. Y. (2006). A distributed parallel estimation architecture for cooperative vehicle formation control. In *Proc. of the american control conference*.
Spanos, D. P., Olfati-Saber, R., & Murray, R. M. (2005). Dynamic consensus on mobile networks. In *IFAC world congress*.
Tu, Sheng-Yuan, & Sayed, Ali H. (2013). On the influence of informed agents on learning and adaptation over networks. *IEEE Transactions on Signal Processing*, *61*(6), 1339–1356.
Xu, Jun, Li, Tao, Xie, Lihua, & Lum, Kai Yew (2011). Dynamic consensus and formation: Fixed and switching topologies. In *Proc. of the 18th IFAC world congress* (pp. 9188–9193). IFAC.
Yang, D., Ren, W., & Liu, X. (2014). Fully distributed adaptive sliding-mode controller design for containment control of multiple Lagrangian systems. *Systems & Control Letters*, *72*, 44–52.
Zhai, G., Hu, B., Yasuda, K., & Michel, A. N. (2001). Stability analysis of switched systems with stable and unstable subsystems: an average dwell time approach. *International Journal of Systems Science*, *32*(8), 1038–1052.

Fabrizio Caccavale received the Laurea degree and the Ph.D. degree in Electronic Engineering from the University of Naples Federico II in 1993 and 1997, respectively. From 1999 to 2001 he has been Assistant Professor at the University of Naples Federico II. He is currently Full Professor at University of Basilicata. His research interests include robot control, multi-robot systems, fault diagnosis, control of aerial robots. He has published more than 130 journal and conference papers, book chapters edited books and monographs. He has been PI or local PI of several national and international research projects. He has been in the program committee of several international conferences and workshops. He has been Associate Editor of the journals *Robotica* and *IEEE Transactions on Control Systems Technology*. He is Senior Member of IEEE and founding member of the *Italian Institute of Robotics and Intelligent Machines* (I-RIM).

Francesco Pierri received the Laurea degree and the Ph.D. degree from the University of Basilicata in 2003 and 2007, respectively. From 2022 he is Associate Professor of Automatic Control and Robotics at the School of Engineering, University of Basilicata. His research interests include: interaction control for robotic manipulators, control and fault diagnosis of multi-robot systems and control of aerial manipulators. From 2008 to 2022 he has been Assistant Professor at the School of Engineering, University of Basilicata. From 2013 to 2018 he has been Associate Editor of the *International Journal of Robotics and Automation*. From 2018 he is Associate Editor of the *IEEE Robotics and Automation Letters*. He has published about 70 papers on international journals and international conference proceedings and he is co-author of the book ‘‘Control and Monitoring of Chemical Batch Reactors’’.