Leader-following consensus with non-uniform and large communication delays

Stefano Battilotti, *Senior Member, IEEE*, Filippo Cacace, *Member, IEEE*, Claudia Califano,*Senior Member, IEEE*, and Massimiliano d'Angelo, *Member, IEEE*

***Abstract*— In this paper we study the leader following consensus problem for linear multi-agent systems and a class of nonlinear ones with bounded disturbances and non-uniform and arbitrarily large communication delays on both directed and undirected graphs. In the linear case we solve for the first time the leader-following consensus problems in presence of arbitrarily large, non-uniform and time-varying communications delay on generic connected graphs. The approach is fully distributed and it is based on a suitable weighting modification of the network links and on an output feedback chain of predictors.**

***Index Terms*— leader-following consensus; delay systems; nonlinear systems.**

I. Introduction

This work addresses the leader-following control problem of linear agents and a class of nonlinear agents over static networks in the presence of norm bounded additive disturbances and communication delays, possibly heterogeneous and time-varying. The problem is solved by means of a stabilizing distributed predictor-based control law.

The cooperative control of a group of agents has been gaining great attention in the last years due to its high potential in many applications in particular in robotics and sensor networks such as vehicle formation [1], autonomous vehicles [2], robotic systems [3], sensor networks [4], target tracking [5], and synchronization [6]. The main objective can be summarized in designing a distributed network protocol, which takes into account the interactions between neighbors, that drives the group of agents to agree on certain variables of interest as time increases

For linear agents in both the cases of fixed and switching topologies, the leader-following problem has been addressed and solved in [7]–[13]. Consensus based on output feedback was instead considered in [14]–[16], while in [17]–[22] measurement noises were introduced. More recently the consensus problem with agents described by nonlinear dynamics was investigated in [23]–[26], while hybrid consensus for multi-agent systems with data-driven jumps, multi-consensus and clustering partitions have been investigated in [27]–[29].

S. Battilotti, Claudia Califano and M. d'Angelo are with DIAG, Università "Sapienza" di Roma, Rome, Italy (e-mail: {battilotti,califano,mdangelo}@diag.uniroma1.it).

F. Cacace is with Faculty of Engineering, University Campus Bio-Medico, Rome, Italy (e-mail: f.cacace@ieee.org).

In this distributed context, when communication delays are considered, the problem is further involved.

Whereas the problems of additive noise [17], [22], [30]–[32] and input delay [33]–[36] are relatively simple to address by extending approaches used in the single agent case, the problem of communications delays (see [37]–[51] and the references therein) is much more challenging, since the delay may undermine the stability of the consensus dynamics across the network. Most of the mentioned works refer to special systems such as scalar systems, or integrators, in some cases in presence of switching topology [12] or time-varying delays. Only a few methods are available for the case of general linear systems, either deterministic or stochastic. In detail, [39] solves the mean square consensus problem of single-integrator systems with measurement noise and communication delays under strongly connected and balanced digraphs, [43] considers first-order integrators under switching topology and communication delay, [44] solves the consensus problem for a tracking problem on integrators, [48] considers stochastic single integrators with delay, [51] considers a network of integrators with communication delays and additive as well as multiplicative noise and [50] provides an approach for second-order systems with multiple and time-varying delays. For general linear systems, [47] solves the problem of delay through a predictor with an integral term. Since this approach results in significant complexity of the implementation, which is particularly critical for agents without large computing resources, [49] proposes a truncated predictor approach for the case of deterministic agents, that however requires that the open-loop dynamics is not exponentially unstable. The approach in [46] addresses general linear systems in the deterministic framework with possibly nonlinear disturbance and constant and uniform communications delay. This approach is based on a novel extended state predictor, and the solution is found by means of LMIs. Finally, more recently [52] resorts to a new approach based on the scalar Lambert equation and obtaining constructive design, while [37] solves an event-triggered consensus problem for heterogeneous deterministic multi-agent systems with nonuniform delays with a LMI-based approach.

In this paper we follow a different approach that exploits one fundamental feature of the leader-following problem, namely the fact that the flow of information is oriented from the leader to the followers. Since the leader-following problem boils down to estimating locally the disagreement with the leader, we aim at reducing the problem to a local estimation problem with time-varying delays in the measurements. We observe that the main obstacle to this end is the fact that for generic networks the information is routed through the cycles of the graph. In presence of delays on the individual edges, each node will receive both recent and outdated information about the leader's state, thus complicating the task of the estimation algorithm. In particular, this prevents the use of centralized predictor designs that have been proposed in the literature for the state estimation of linear or nonlinear systems in presence of measurement delays. Based on this observation, we claim that the difficulties disappear if each node runs an algorithm to determine the relative position of the neighbors with respect to the leader, in such a way that only the most recent information is used to estimate the leader's state. We show that this approach allows to recover the solutions used in the cases of single systems with measurement delay, and in particular that simple and robust designs are possible even in presence of heterogeneous and time-varying delays on the links of the communication network. This approach has been first proposed in [53], in the context of linear stochastic agents with uniform constant delay, and here it is further developed to the case of nonlinear agents with heterogeneous and arbitrarily large time-varying delays. In practice, the improvement with respect to existing proposals is that the hypotheses are less restrictive, as we admit heterogeneous and time-varying delays, the design is simple and it yields a computationally cheap controller that does not involve distributed terms and a non conservative delay bound.

The paper is organized as follows. In Section II we recall some notions on directed and undirected graphs that are used to characterize the network. We then formally define the leader following problem with linear agents. In Section III we propose a distributed algorithm to associate to the given network a directed acyclic graph (DAG) rooted in the leader. In Section IV the solution to the leader following problem is derived first by considering small delays and then by generalizing it to the case of large delays. In Section V the same problem is solved for nonlinear systems with disturbances on each agent dynamics and measurements, thus generalizing the results obtained in the previous section. An example adapted from [46] is discussed in Section VI. The proposed methodology is tested by comparing the performance with the results presented in [46].

Notation: $\mu(M)$ denotes the spectral abscissa of the matrix M. If $\mu(M) < 0$ M is said to be Hurwitz. $\|.\|$ is the Euclidean norm unless otherwise specified.

II. Preliminaries and problem statement

To the leader and the N homogeneous agents we associate an unweighted simple graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where $\mathcal{V} = \{0, 1, 2, \ldots, N\}$ is the set of vertices representing the agents and $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ is the set of edges of the graph. The node 0 represents the leader. Edge (i, j) indicates that agent j can send information to agent i. In this case j is a neighbor of i. The set of neighbors of node j is denoted by $\mathcal{N}^j$. The connections graph is represented through the adjacency matrix $\mathcal{A} = [a_{ij}] \in \{0,1\}^{(N+1)\times(N+1)}$, where $a_{ij} = 1$ if $(i, j) \in \mathcal{E}$ while it is 0 if $(i, j) \notin \mathcal{E}$.

The $N + 1$ agents are initially considered linear and described, for $k = 0, \ldots, N$, by

$$\dot{X}_k(t) = AX_k(t) + BU_k(t) \tag{1}$$

$$Y_k(t) = CX_k(t), \tag{2}$$

where $X_k(t) \in \mathbb{R}^n$ denotes the state of the k-th agent and $Y_k(t) \in \mathbb{R}^q$ is the associated output. For the leader following problem it is assumed that the leader behavior is not influenced by other agents (*i.e.* the leader disregards any information from other nodes even when $\mathcal{N}^0 \neq \emptyset$) and the leader's control input is zero [9], [24], [54], i.e. $U_0 \equiv 0$. Depending on whether or not the graph $\mathcal{G}$ is oriented we have two cases.

Undirected graphs: in this case $\mathcal{A}$ is symmetric, $a_{ij} = a_{ji}$. The number of connections of each node is represented through the degree matrix $\mathcal{D}$ which is a diagonal matrix with $\mathcal{D}_{ii} = |\mathcal{N}^i|$. The Laplacian $\mathcal{L} = [\ell_{ij}] \in \mathbb{R}^{(N+1)\times(N+1)}$ is defined as $\mathcal{L} = \mathcal{D} - \mathcal{A}$. $\mathcal{G}$ is *connected* if there is a path (*e.g.* a sequence of connected edges) between every pair of vertices. An undirected graph $\mathcal{G}$ is connected if and only if $\mathcal{L}$ has a simple 0 eigenvalue [55], [43], [12], [26]. In case of *undirected graphs* we make the following assumption.

Assumption 1: The undirected graph $\mathcal{G}$ is connected.

Directed graphs: The graph is said to be *directed* if $(i, j) \in \mathcal{E}$ does not necessarily imply $(j, i) \in \mathcal{E}$. A directed graph $\mathcal{G}$ is *strongly connected* if between any pair of distinct nodes i and j in $\mathcal{G}$, there exists a directed path from i to j, $i, j \in \mathcal{N}$. A directed graph $\mathcal{G}$ *contains a directed spanning tree* if there exists a root node that has directed paths to all other nodes. The Laplacian $\mathcal{L} \in \mathbb{R}^{(N+1)\times(N+1)}$ is defined as $\mathcal{L} := [\ell_{i,j}] = \mathcal{M} - \mathcal{A}$ where the i-th diagonal entry of the diagonal matrix $\mathcal{M}$ is given by $m_i = \sum_{j=0}^{N} a_{i,j}$. By construction $\mathcal{L}$ has a zero eigenvalue with an associated eigenvector $\mathbf{1}_{N+1}$ (i.e. such that $\mathcal{L}\mathbf{1}_{N+1} = 0$) and if the graph is strongly connected all the other eigenvalues lie in the open right-half complex plane. In the case of *directed graphs*, the following assumption will be considered.

Assumption 2: A directed spanning tree is contained in $\mathcal{G}$ with the leader as the root node.

Notice that for undirected graphs, Assumption 1 implies the existence of a spanning tree with the leader as the root node.

Definition 1: Leader Following Problem (LFP): Given a graph topology $\mathcal{G}$ associated to (1), find $U_k(t)$ for each agent k so that $X_0(t) - X_k(t)$ is asymptotically stable.

The presence of time-varying communications delays is modeled by the set $\{\delta_{jk} : \mathbb{R}_+ \to [0, \bar{\delta}_{jk}]\}$ for $(j, k) \in \mathcal{E}$, that indicates that at time t the communication from agent k to agent j is affected by the delay $\delta_{jk}(t) \in [0, \bar{\delta}_{jk}]$.

Definition 2: Leader Following Problem with Non-uniform Communication Delays (LFPND): Given a graph topology $\mathcal{G}$ associated to (1), and the communications delays $\{\delta_{jk}\}$ among the nodes, find $U_k(t)$ for each agent k so that $X_0(t) - X_k(t)$ is asymptotically stable.

III. Changing the Weight of the Communication Links

The essential idea is to restrict the communications to a DAG rooted in the leader. This avoids unnecessary communications, simplifies the stability analysis and allows for non-uniform delays. The communication can be restricted to a DAG rooted in the leader by replacing the links represented by ℓ_{kj} with the new weights $\bar{\ell}_{kj}$ defined by Algorithm 1.

The essential idea is to set to 0 the weight of all the neighbors whose distance from the leader, measured by the number of edges of the shortest path, is larger or equal to the distance of the node itself. Algorithm 1 achieves this in a distributed and iterative way. At each iteration a node sets its own distance f_k as the minimum distance of its neighbors plus 1. When an incoming link originates from a node farther away from the leader, the corresponding weight is set to 0. At steady state, Algorithm 1 ensures that information arriving at each node k originates from nodes j that are closer than k to the leader. Thus, $\mathcal{N}^0 = \emptyset$ for the leader, $\mathcal{N}^j = \{0\}$ for all the nodes j initially having the leader as neighbor, etc.

Algorithm 1 Distributed algorithm for node k to restrict the communications to a DAG rooted in the leader

Ensure: A DAG rooted in the leader with Laplacian matrix entries $\bar{\ell}_{kj}$
if $k = 0$ **then**
  $f_k \leftarrow 0$
else
  $f_k \leftarrow \infty$
end if
loop ▷ forever
  send f_k to the neighbors
  if $\min_{j\in\mathcal{N}^k}\{f_j\} < \infty$ **then**
    $f_k = \min_{j\in\mathcal{N}^k}\{f_j\} + 1$
    for $j \in \mathcal{N}^k$ **do**
      if $f_j < f_k$ **then**
        $\bar{\ell}_{kj} \leftarrow \ell_{kj}$
      else
        $\bar{\ell}_{kj} \leftarrow 0$
      end if
    end for
  end if
  $\bar{\ell}_{kk} \leftarrow -\sum_{j\in\mathcal{N}^i} \bar{\ell}_{kj}$
end loop

The idea behind the Algorithm 1 is illustrated in Fig. 1 that shows the original (top) and resulting (bottom) graphs. The edge $(1,4)$ suppressed in the modified topology actually does not convey any useful information to node 1, since the relevant information for the leader-following task comes from the leader 0. Actually, the incoming information from node 4 is a disturbance for node 1. Whereas in the delay-less case the communications are instantaneous and the dynamics of this disturbance is dissipative, in presence of delays the stability of the consensus error can no longer be guaranteed. This is the reason why there is no known method to compensate large delays over arbitrary graphs, and even the case of bounded delays is difficult to manage: the delay builds up at each edge as the information spreads through the network, thus an agent may receive information with arbitrarily long delay even if the delay across the individual edges is small.

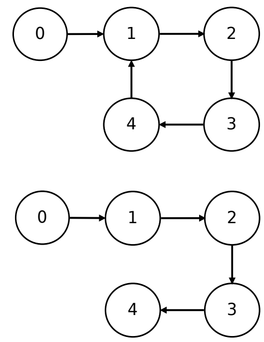


Fig. 1. Communication topology for the example in Section VI (top). The graph resulting from weighting the edges according to the algorithm of Section III (bottom).

In the sequel, we assume that the links in the communication graph has been changed according to the above algorithm and we show that in this case it is possible to compensate arbitrarily long delays on individual links.

IV. Linear Agents with Non-uniform Communication Delays

We recall two simple sufficient conditions to solve the leader following problem with linear agents in absence of delay.

Lemma 1: If (A, B) is a stabilizable pair, and F_k, $k = 1,\dots,N$, is such that $A + BF_k$ is Hurwitz stable, then the control

$$U_k(t) = F_k(X_k(t) - X_0(t)),\ k = 1,\dots,N \tag{3}$$

solves the LFP.

Proof. Let $\eta_k(t) = X_k(t) - X_0(t)$, $k = 1,\dots,N$. We have

$$\dot{\eta}_k(t) = (A + BF_k)\eta_k(t) \tag{4}$$

Since $A + BF_k$ is Hurwitz stable, it follows that $\eta_k(t)$ is exponentially stable and therefore (3) solves the LFP. η_k is the disagreement of the agent k with respect to the leader, and clearly $\eta_0 \equiv 0$. □

Lemma 2: If (A, B) is a stabilizable pair, and F_k, $k = 1,\dots,N$, is such that $A + BF_k$ is Hurwitz stable, while $\widehat{X}_{0,k}(t)$ is an estimate of $X_0(t)$ available at agent k such that $\|X_0(t) - \widehat{X}_{0,k}(t)\| \to 0$, then the control

$$U_k(t) = F_k(X_k(t) - \widehat{X}_{0,k}(t)),\ k = 1,\dots,N, \tag{5}$$

solves the LFP.

Proof. The disagreement $\eta_k = X_k(t) - X_0(t)$ obeys to

$$\dot{\eta}_k(t) = (A + BF_k)\eta_k(t) + BF_k\varepsilon_k(t), \tag{6}$$

where $\varepsilon_k(t) = X_0(t) - \widehat{X}_{0,k}(t)$ is the estimation error. Since $A + BF_k$ is Hurwitz stable and $\widehat{X}_{0,k}(t) \to X_0(t)$, it follows that $\eta_k(t) \to 0$ and (3) solves the LFP. □

A. LFPND with small delays

The LFPND can be solved by the control (5) with an estimate such that $\widehat{X}_{0,k}(t) \to X_0(t)$ exponentially.

Assumption 3: The functions $\{\delta_{kj} : \mathbb{R}_+ \to [0, \bar{\delta}_{kj}]\}$ are differentiable with $\dot{\delta}_{kj}(t) < 1$ almost everywhere. At each t the values $\delta_{kj}(t)$, $\dot{\delta}_{kj}(t)$, $j \in \mathcal{N}^k$ are known at the agents k.

Assumption 3 is not too demanding since, for instance, we may assume that some form of clock synchronization is present in the network and a timestamp is associated to the messages from j to k. Moreover, the derivatives of δ_{jk} can be computed with precision in any digital implementation based on small integration steps, since it reduces to the difference between the delay value across the integration step.

Theorem 1: If:

1) Assumption 1 or 2 hold together with Assumption 3.
2) The connection topology is modified according to Algorithm 1.
3) (A, B) is controllable and the gains F_k are such that $A + BF_k$, $k = 1, \ldots, N$, is Hurwitz stable.
4) (A, C) is an observable pair and at each node k the gains L_{kj} are chosen such that with, $L_k = \sum_{j\in\mathcal{N}^k} \bar{\ell}_{kj} L_{kj}$, $\overline{A}_k = A - L_k C$, $k = 1, \ldots, N$, is Hurwitz stable, *i.e.* $\mu(\overline{A}_k) < 0$.
5) For all k and for some $\alpha > 0$ such that $|\mu(\overline{A}_k)| > \alpha$ the bounds on the communications delay $\{\bar{\delta}_{kj}\}$ satisfy

$$\sum_{j\in\mathcal{N}^k} \int_0^{\bar{\delta}_{kj}} \left\| C e^{(\overline{A}_k + \alpha I_n)\theta} L_{kj} \right\| \, \mathrm{d}\theta \; < \; 1, \tag{7}$$

where $k = 1, \ldots, N$, then the solution for $t \geq 0$ of

$$\begin{aligned} \dot{\widehat{X}}_{0,k}(t) =& A\widehat{X}_{0,k}(t) - \sum_{j\in\mathcal{N}^k} (1 - \dot{\delta}_{kj}(t)) \bar{\ell}_{kj} e^{\overline{A}_k \delta_{kj}} L_{kj} \\ & \cdot \Big(C\widehat{X}_{0,k}(t - \delta_{kj}(t)) - Y_j(t - \delta_{kj}(t)) \Big) \end{aligned} \tag{8}$$

with $\widehat{X}_{0,k}(\tau) = \varphi_k(\tau)$ for $\tau \leq 0$, $\|\varphi_k\|_\infty < \infty$, is an exponential estimate of $X_0(t)$ and the control (5) with $\widehat{X}_{0,k}$ given by (8) solves the LFPND.

Remark 1: The predictor (8) depends on the communications delays, and condition (7) provides a bound for the delays $\{\bar{\delta}_{kj}\}$. The parameter α represents the desired rate of convergence of $\widehat{X}_{0,k}$ to X_0. In general, a larger α will yield a smaller set of tolerable delays $\{\bar{\delta}_{kj}\}$.

Proof. Let $\varepsilon_k(t) = X_0(t) - \widehat{X}_{0,k}(t)$ be the estimation error at node k. By adding and subtracting in each term of the summation on the right side of (8) the term $CX_0(t - \delta_{kj}(t))$, we obtain, $k = 1, \ldots, N$,

$$\begin{aligned} \dot{\varepsilon}_k(t) =& A\varepsilon_k(t) - \sum_{j\in\mathcal{N}^k} (1 - \dot{\delta}_{kj}(t)) \bar{\ell}_{kj} e^{\overline{A}_k \delta_{kj}(t)} \\ & \cdot L_{kj} C \left(\varepsilon_k(t - \delta_{kj}(t)) + \eta_j(t - \delta_{kj}(t)) \right). \end{aligned} \tag{9}$$

We now proceed inductively. Since $\eta_0 \equiv 0$, for the nodes k such that $\mathcal{N}^k = \{0\}$, (9) becomes

$$\dot{\varepsilon}_k(t) = A\varepsilon_k(t) - (1 - \dot{\delta}_{k0}(t)) e^{\overline{A}_k \delta_{k0}(t)} L_k C \varepsilon_k(t - \delta_{k0}(t)). \tag{10}$$

In order to prove that (10) is exponentially stable, let $\varepsilon_k^\alpha(t) = e^{\alpha t}\varepsilon_k(t)$ for $t \geq 0$, thus

$$\begin{aligned} \dot{\varepsilon}_k^\alpha(t) = (A + \alpha I_n)\varepsilon_k^\alpha(t) - (1 - \dot{\delta}_{k0}(t)) e^{(\overline{A}_k + \alpha I_n)\delta_{k0}(t)} \\ \cdot L_k C \varepsilon_k^\alpha(t - \delta_{k0}(t)). \end{aligned} \tag{11}$$

Notice that $\overline{A}_k + \alpha I_n$ is Hurwitz since, by hypothesis, $\alpha + \mu(\overline{A}_k) < 0$. Equation (11) admits the integral representation

$$\varepsilon_k^\alpha(t) = \int_{t-\delta_{k0}(t)}^t e^{(\overline{A}_k + \alpha I_n)(t-\tau)} L_k C \varepsilon_k^\alpha(\tau) \, \mathrm{d}\tau + a_k, \tag{12}$$

where a_k depends on the initial condition φ_k of $\widehat{X}_{0,k}$ in $[-\bar{\delta}_{k0}, 0]$. It is immediate to verify that differentiating (12) one gets (11). Pre-multiplying the LHS of (12) by C and taking norms,

$$\begin{aligned} \|C\varepsilon_k^\alpha(t)\| \leq \int_0^{\bar{\delta}_{k0}} \left\| C e^{(\overline{A}_k + \alpha I_n)\theta} L_k \right\| \, \mathrm{d}\theta \cdot \sup_{\tau \in [t - \bar{\delta}_{k0}, t]} \|C\varepsilon_k^\alpha(\tau)\| \\ + \|Ca_k\|, \end{aligned} \tag{13}$$

The condition (7) implies that $\|C\varepsilon_k^\alpha(t)\|$ is uniformly bounded, that is, $\|C\varepsilon_k(t)\| \to 0$ exponentially. Since (A, C) is observable, then $\|\varepsilon_k(t)\| \to 0$ exponentially and $\eta_k \to 0$ exponentially with rate α in view of Lemma 2. We now consider the generic node k with the inductive hypothesis that $\|\eta_j(t)\| \to 0$ for the agents j that are closer than k to the leader. Since in (9) the terms $\eta_j(t)$ are therefore exponentially vanishing, the dynamics of ε_k tends asymptotically to

$$\begin{aligned} \dot{\varepsilon}_k(t) =& A\varepsilon_k(t) - \sum_{j\in\mathcal{N}^k} (1 - \dot{\delta}_{kj}(t)) \bar{\ell}_{kj} e^{\overline{A}_k \delta_{kj}(t)} \\ & \cdot L_{kj} C \varepsilon_k(t - \delta_{kj}(t)). \end{aligned} \tag{14}$$

We can introduce $\varepsilon_k^\alpha(t) = e^{\alpha_k t}\varepsilon_k(t)$ for $t \geq 0$, with $\alpha_k < \alpha_j < \alpha$. Notice that α_k, that can be arbitrarily close to α, is not a design parameter but it is used only to prove the convergence. Moreover, if the inequality (7) is satisfied by α then it is satisfied by any $\alpha_k < \alpha$. Proceeding as in the previous case we arrive at (notice that $|\ell_{kj}| = 1$)

$$\begin{aligned} \|C\varepsilon_k^\alpha(t)\| \leq \sum_{j\in\mathcal{N}^k} \int_0^{\bar{\delta}_{kj}} \left\| C e^{(\overline{A}_k + \alpha_k I_n)\theta} L_k \right\| \, \mathrm{d}\theta \\ \cdot \sup_{\tau \in [t - \bar{\delta}_{kj}, t]} \|C\varepsilon_k^\alpha(\tau)\| + \|Ca_k\|, \end{aligned} \tag{15}$$

and as above we conclude that (14) is exponentially stable thanks to (7) and $\eta_k \to 0$ exponentially with rate α_k. □

Theorem 1 yields the distributed design procedure for the gains F_k and L_k shown in Algorithm 2. Here we suppose that the desired rate α of convergence to 0 of the leader's state estimation error at each node k is a design parameter. This value determines also the rate of convergence to the leader's state of X_k when the control (5) is used. Algorithm 2 sets $\alpha + \mu(A - L_k C) = 0$, but the hypothesis $\alpha + \mu(A - L_k C) < 0$ of Theorem 1 is satisfied by any value less than α, thus the algorithm determines the largest α that satisfies (7), and it may fail if this is not possible. We notice that this may happen only when A is not Hurwitz, because otherwise the choice $L_k = 0$ and $\alpha = -\mu(A) + \epsilon_\alpha$ with arbitrarily small $0 < \epsilon_\alpha < -\mu(A)$

Algorithm 2 Distributed control design algorithm at each node k with non-uniform small communications delays

Require: Assumption 1 or 2, and Assumption 3
Require: Systems matrices A, B, C
Require: delay bounds $\bar{\delta}_{kj}$ for $j \in \mathcal{N}_k$
Require: desired exponential rate $\alpha > 0$ of convergence to 0 of the leader's state estimation error
Require: adaptation step $0 < \delta_\alpha < \alpha$.
Ensure: controller gain F_k, observer gain L_k that solve LFPND and attainable rate α_k or **fail**.

```
compute F_k such that μ(A + BF_k) = −α
compute L_k such that μ(A − L_kC) = −α
α_k ← α
while α_k > 0 & (7) not satisfied do
    α_k ← α_k − δ_α
    compute L_k such that μ(A − L_kC) = −α_k
end while
if α_k > 0 then
    success
else
    fail
end if
```

always satisfies (7). We are therefore left with the problem to cope with large delays when the matrix A is simply stable or exponentially unstable, which is the topic of the next section.

Remark 2: The convergence analysis of Theorem 1 is based on the exact knowledge of $\delta_{kj}(t)$ and its derivative. We notice that a small systematic bias on $\delta_{kj}(t)$ has no effect on $\dot{\delta}_{kj}(t)$. A noisy measurement of $\delta_{kj}(t)$ reflects on $\dot{\delta}_{kj}(t)$ and it introduces a multiplicative state noise in the equation of the estimation error. When there is a significant noise on $\dot{\delta}_{kj}(t)$ it is possible to drop it in the observer/predictor, *i.e.* to replace $(1 - \dot{\delta}_{kj}(t))$ with 1 in (8). In this way one obtains the observer/predictor described in [56], which, in the case of constant delays, has the same delay bound.

B. LFPND with large delays

The design of chains of predictors is a consolidated technique for single systems (see for instance the more recent [57]) but its application to networks of multi-agent systems has so far proven elusive. By taking advantage of the DAG structure of the communication network provided by Algorithm 1 the solution is easy to find. With the aim of adapting the observer design of Theorem 1 to cope with large delays we adopt a slightly different philosophy: the last element of the chain is an observer that estimates $X_0(t - \bar{\delta}_k)$ from the measurement of the neighbors that refer to $t - \bar{\delta}_k$, and the remaining elements are predictors over a fraction of $\bar{\delta}_k$ that eventually yield an estimate of $X_0(t)$. Consequently, the chain includes at least two elements and the communications delays can be arbitrarily large. Since the observer and the predictors are fed with information from the neighbors having constant delay the delay derivatives are no longer needed. Finally, the delay bound for the individual elements of the chain is in general less restrictive than (7). On the negative side, an array of at least two predictors is needed, as well as a buffer to align temporally the information coming from the neighbors. Assumption 3 is replaced by the following milder version.

Assumption 4: The functions $\{\delta_{kj} : \mathbb{R}_+ \to [0, \bar{\delta}_{kj}]\}$ are continuous and the values $\delta_{kj}(t)$, $j \in \mathcal{N}^k$, are known at the agents k.

At the agent k a chain of m_k+1 predictors is made up by the estimators $\widehat{X}_{i,k}$, $i = 0, \ldots, m_k$. The chain length is the smallest integer $m_k > 0$ such that, with $\bar{\delta}_k = \max_{j\in\mathcal{N}^k}\{\bar{\delta}_{kj}\}/m_k$,

$$\int_0^{\bar{\delta}_k} \left\| C e^{(\overline{A}_k + \alpha I_n)\theta} L_k \right\| \mathrm{d}\theta < 1, \tag{16}$$

for $L_k = \sum_{j\in\mathcal{N}^k} \bar{\ell}_{kj} L_{kj}$ and arbitrary α such that $\alpha + \mu(\overline{A}_k) < 0$. The estimator $\widehat{X}_{i,k}(t)$ aims at estimating $X_0(t - i\bar{\delta}_k)$, thus $\widehat{X}_{0,k}$ estimates $X_0(t)$ and $\widehat{X}_{m_k,k}$ estimates $X_0(t - m_k\bar{\delta}_k)$. The last estimator, with index $i = m_k$, is an observer with equation, for $t \geq 0$,

$$\begin{aligned}\dot{\widehat{X}}_{m_k,k}(t) &= A\widehat{X}_{m_k,k}(t) \\ &\quad - \sum_{j\in\mathcal{N}^k} \bar{\ell}_{kj} L_{kj} \left(C\widehat{X}_{m_k,k}(t) - y_j(t - m_k\bar{\delta}_k) \right).\end{aligned} \tag{17}$$

Notice that (17) doesn't use the most recent value of y_j but, since δ_{jk} is continuous, the value $y_j(t - m_k\bar{\delta}_k)$ is available, for example by buffering the past output values of the neighbors in $[t - m_k\bar{\delta}_k,\, t]$. The i-th predictor, $0 \leq i < m_k$, has equation

$$\dot{\widehat{X}}_{i,k}(t) = A\widehat{X}_{i,k}(t) - e^{\overline{A}_k\bar{\delta}_k} L_k C \left(\widehat{X}_{i,k}(t - \bar{\delta}_k) - \widehat{X}_{i+1,k}(t) \right). \tag{18}$$

All the estimators are initialized to an arbitrary norm-bounded function at negative times. As before, the controls are designed as a feedback from the estimate $\widehat{X}_{0,k}(t)$, that is,

$$U_k(t) = F_k(X_k(t) - \widehat{X}_{0,k}(t)),\ k = 1, \ldots, N, \tag{19}$$

Theorem 2: In the hypothesis of Theorem 1, with Assumption 3 replaced by Assumption 4, assume that, for each k, m_k satisfies (16) for some $\alpha > 0$ such that $\alpha + \mu(\overline{A}_k) < 0$. Then the control (19) with $\widehat{X}_{0,k}$ solution of (18) solves the LFPND.

Proof. The estimation errors are $\varepsilon_{i,k} = X_0(t - i\bar{\delta}_k) - \widehat{X}_{i,k}(t)$. For $i = m_k$, that is, the last predictor, the error satisfies, $t \geq 0$,

$$\dot{\varepsilon}_{m_k,k}(t) = (A - L_kC)\varepsilon_{m_k,k}(t) - \sum_{j\in\mathcal{N}^k} \bar{\ell}_{kj} L_{kj} C \eta_j(t - m_k\bar{\delta}_k), \tag{20}$$

and, in the inductive hypothesis $\eta_j \to 0$, $\varepsilon_{m_k,k}$ is exponentially stable. For $i < m_k$ the error equation satisfies

$$\dot{\varepsilon}_{i,k}(t) = A\varepsilon_{i,k}(t) - e^{\overline{A}_k\bar{\delta}_k} L_k C \left(\varepsilon_{i,k}(t - \bar{\delta}_k) + \varepsilon_{i+1,k}(t) \right), \tag{21}$$

and since $\varepsilon_{i+1,k}(t) \to 0$, by proceeding as in Theorem 1 the delay condition (16) is sufficient to prove $\varepsilon_{j,k}(t) \to 0$ and $\widehat{X}_{0,k}(t) \to X_0(t)$. □.

Remark 3: The prediction could also be obtained with only m_k, rather than $m_k + 1$, predictors, see for example [57], but the structure in (17)–(18) is simpler. Notice that condition (16) is in general less restrictive than (7). For example, a chain of two estimators (*i.e.* $m = 1$) would yield a less conservative

Algorithm 3 Distributed control design algorithm at each node k for arbitrarily large non-uniform communications delays

Require: Assumption 1 or 2, and Assumption 4
Require: Systems matrices A, B, C
Require: delay bounds $\bar{\delta}_{kj}$ for $j \in \mathcal{N}_k$
Require: desired exponential rate $\alpha > 0$ and margin $\epsilon_\alpha > 0$ of convergence to 0 of the leader's state estimation error
Ensure: controller gain F_k, observer gain L_k, chain length $m_k + 1$ that solve LFPND.
compute F_k such that $\mu(A + BF_k) = -\alpha$
compute L_k such that $\mu(A - L_kC) = -\alpha - \epsilon_\alpha$
compute the largest $\bar{\delta}_k$ that satisfies (16) for given A, C, L_k, α
$m_k \leftarrow \lceil \frac{\max_{j \in \mathcal{N}^k}\{\ \bar{\delta}_{kj}\}}{\bar{\delta}_k} \rceil$

delay bound than the single observer (8) when there is more than one neighbor.

Theorem 2 yields the distributed design procedure shown in Algorithm 3 to compute at each agent k the gains F_k, L_k and the chain length $m_k + 1$. The gains F_k, L_k are computed by each agent through Algorithm 3 and they are used to implement the local observers (17), (18) and the local controller (19). Also in this case we suppose that the desired rate α convergence to 0 of the leader's state estimation error at each node k is a design parameter. The main difference with respect to Algorithm 2 is that in this case α can be freely chosen, and the convergence to 0 of the leader's state estimation error is ensured for any delay and any location of the eigenvalues of A. Clearly, larger $\mu(A)$ and larger delays yield a larger value of the chain length $m_k + 1$ and a larger number of chained observers at each agent.

V. Nonlinear agents with Non-uniform Communication delays and disturbances

We can extend the results of Theorem 1 to the case of nonlinear systems with additive disturbances. To this aim, consider the agents' structure, $k = 0, \dots, N$,

$$\dot{X}_k(t) = AX_k(t) + BU_k(t) + \psi(X_k(t)) + w_k(t) \tag{22}$$
$$Y_k(t) = CX_k(t) + v_k(t), \tag{23}$$

where, as usual $U_0 \equiv 0$, $w_k \in \mathbb{R}^n$ and $v_k \in \mathbb{R}$ are disturbances, $\psi : \mathbb{R}^n \to \mathbb{R}^n$ is smooth and the following assumptions hold.

Assumption 5:

- The disturbances are norm bounded, $\|w_i\|_\infty = \bar{w} < \infty$, $\|v_i\|_\infty = \bar{v} < \infty$, $i = 0, 1, \dots, N$.
- $\gamma := \sup_{X \in \mathbb{R}^n} \|\frac{\partial \psi}{\partial X}(X)\| < \infty$ (i.e. ψ is globally Lipschitz).

Moreover, we introduce an assumption for stabilizing through the input U_k the disagreement $X_k - X_0$.

Assumption 6: There exist positive definite $P, Q \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m\times m}$ such that for all $X \in \mathbb{R}^n$

$$P\left(A + \frac{\partial \psi}{\partial X}(X)\right) + \left(A + \frac{\partial \psi}{\partial X}(X)\right)^\top P - PBR^{-1}B^\top P + Q \leq 0 \tag{24}$$

When (A, B) is in Brunowski form and $\psi(X) \equiv B\phi(X)$ (i.e. input matching condition), it is possible to construct the matrices P, R and Q satisfying (24) (see Lemma 6 in the Appendix). Boundedness of $\|\frac{\partial \psi}{\partial X}(X)\|$ (Assumption 5) is to a certain extent necessary for the existence of constant matrices P, Q and R satisfying (24).

Definition 3: Leader Following Problem with Non-uniform Communication Delays and Disturbances (LFPNDD): Given a graph topology $\mathcal{G}$ associated to (22), and the communications delays $\{\delta_{kj}\}$ among the nodes, find $U_k(t)$ for each agent k so that $\|X_0(t) - X_k(t)\|$ is uniformly bounded.

First of all, we propose a simple control law which stabilizes the disagreement $\eta_k := X_k - X_0$ and guarantees its boundedness under bounded input and disturbance in absence of delay.

Lemma 3: Under Assumption 6, the control input

$$U_k(t) = -R^{-1}B^\top P(X_k(t) - X_0(t)) \tag{25}$$

solves the LFP. Moreover, for each $k = 1, \dots, N$ the system

$$\dot{\eta}_k(t) = (A - R^{-1}B^\top P)\eta_k(t) + B\Delta U_k(t) + w_k(t) + \psi(\eta_k(t) + X_0(t)) - \psi(X_0(t)) + \Delta w_k(t) \tag{26}$$
$$\dot{X}_0(t) = AX_0(t) + \psi(X_0(t)) + w_0(t) \tag{27}$$

has bounded state $\eta_k(t)$ under arbitrary but bounded signals $\Delta U_k(t)$ and $\Delta w_k(t)$. Moreover, $\eta_k(t) \to 0$ as $\Delta U_k(t)$, $\Delta w_k(t)$, $w_k(t) \to 0$.

The proof easily follows from Assumption 6 and it is omitted. With the help of Lemma 3 we can give the following solution to the LFPNDD (the proof is reported in the Appendix).

Theorem 3: If:

1) Assumption 1 or Assumption 2 hold together with Assumptions 3, 5 and 6.
2) The connection topology is modified according to Algorithm 1.
3) The pair (C, A) is observable and at each node k the gains L_{kj} are chosen such that, with $L_k = \sum_{j \in \mathcal{N}^k} \bar{\ell}_{kj} L_{kj}$, the matrix $\overline{A}_k = A - L_kC$ is Hurwitz stable.
4) For all k the bounds $\{\bar{\delta}_{kj}\}$ on the communications delay are such that

$$\sum_{j \in \mathcal{N}^k} \int_0^{\bar{\delta}_{kj}} \|e^{\overline{A}_k \theta} L_{kj} C\| \, \mathrm{d}\theta + \int_0^\infty \gamma \|e^{\overline{A}_k \tau}\| \, \mathrm{d}\tau < 1, \tag{28}$$

where $k = 1, \dots, N$, then the control input U_k

$$U_k(t) = -R^{-1}B^\top P(X_k(t) - \widehat{X}_{0,k}(t)) \tag{29}$$

where $\widehat{X}_{0,k}(t)$ is the solution for $t \geq 0$ of

$$\dot{\widehat{X}}_{0,k}(t) = A\widehat{X}_{0,k}(t) + B\psi(\widehat{X}_{0,k}(t)) - \sum_{j \in \mathcal{N}^k} (1 - \dot{\delta}_{kj}(t)) \cdot \bar{\ell}_{kj} e^{\overline{A}_k \delta_{kj}(t)} L_{kj} \cdot \left(C\widehat{X}_{0,k}(t - \delta_{kj}(t)) - y_j(t - \delta_{kj}(t))\right) \tag{30}$$

solves the LFPNDD.

It is possible to provide conditions that ensure that the delay condition (28) is satisfied by a suitable choice of L_k. For example, when $\psi(s)$ satisfies the input matching condition and the agents' matrices (A, B, C) are a Brunowski triple (A_b, B_b, C_b) there always exists a suitable gain L_k and sufficiently small positive delay bounds $\bar{\delta}_{kj}$ such that condition (28) is satisfied. To show this we recall two known technical results.

Lemma 4: Let $A_L = A_b - LC_b$ and $\sigma(A_L) = \bar{\lambda} = \{\lambda_1, \dots, \lambda_n\}$ be a set of n distinct negative reals. $V(\bar{\lambda})$ denotes the Vandermonde matrix whose i-th row is $[\lambda_i^{n-1} \ \cdots \ \lambda_i \ 1]$. Then, the change of coordinates represented by $V(\bar{\lambda})$ makes A_L diagonal, i.e. $A_L = V^{-1}(\bar{\lambda})\Lambda V(\bar{\lambda})$ with $\Lambda = \mathrm{diag}_i(\lambda_i)$. Moreover, if $\lambda_i = -w^i$, $w > 0$, $i = 1, \dots, n$, then

$$\lim_{w\to\infty} \|V^{-1}(\bar{\lambda})\| = 1. \tag{31}$$

Proof: See Lemma 2.2 in [58]. ■

Lemma 5: For any $\epsilon > 0$ there exists L such that, with $A_L = A_b - LC_b$,

$$\int_0^\infty \left\|e^{A_L s} B_b\right\| \mathrm{d}s \ < \epsilon. \tag{32}$$

Proof: Choose $\lambda_i = -w^i$, $w > 0$. Since $\|V(\bar{\lambda})B_b\| = \|\mathrm{col}_{i=1}^n(1)\| = \sqrt{n}$,

$$\lim_{t\to\infty} \int_0^t \left\|e^{A_L s} B_b\right\| \mathrm{d}s \le \lim_{t\to\infty} \left\|V^{-1}(\bar{\lambda})\right\| \int_0^t e^{-ws}\sqrt{n}\,\mathrm{d}s = \frac{\left\|V^{-1}(\bar{\lambda})\right\|\sqrt{n}}{w} \tag{33}$$

and in view of Lemma 4 it is always possible to choose w such that $\left\|V^{-1}(\bar{\lambda})\right\|\sqrt{n}/w < \epsilon$. ■

Remark 4: The trade-off between the Lipschitz constant of the non-linearities and the maximum communication delays is clearly visible in the structure of condition (28). In order to make the second term in the left side of (28) less than 1 it is necessary that the eigenvalues of $\overline{A}_k$ are sufficiently negative which is possible if L_k is sufficiently large. A large L_k increases the norm in the first integral, making $\bar{\delta}_{kj}$ smaller.

When there are only additive norm bounded disturbances, i.e. $\psi(X, U) \equiv 0$, we recover the delay bound (7).

Corollary 4: In the hypotheses of Theorem 3, if $\psi(X, U) \equiv 0$ and the bounds $\{\bar{\delta}_{kj}\}$ on the communications delays satisfy (7) then the controls (29) solve LFPNDD.

Theorem 3 yields the distributed design procedure shown in Algorithm 4 to compute at each agent k the gains $-R^{-1}B^\top P$ and L_k that are used to implement the local observer (30) and the local controller (29). The design parameters are the matrices R and Q of (24) and the desired spectral abscissa α of the closed-loop estimation error dynamical matrix $A - L_k C$. The algorithm determines the largest α_k that satisfies (28).

Finally, when disturbances vanish in the origin we recover asymptotic consensus.

Corollary 5: In the hypotheses of Theorem 3, if $\bar{w} = \bar{v} = 0$ and $\psi(0, 0) = 0$, if L_{kj} and the bounds $\{\bar{\delta}_{kj}\}$ on the communications delays satisfy, for any $\alpha > 0$,

$$\sum_{j\in\mathcal{N}^k} \int_0^{\bar{\delta}_{kj}} \left\|e^{(\overline{A}_k+\alpha I_n)\theta} L_{kj} C\right\| \mathrm{d}\theta + \gamma \int_0^\infty \left\|e^{(\overline{A}_k+\alpha I_n)\tau}\right\| \mathrm{d}\tau < 1, \tag{34}$$

then the controls (29) solve the LFPNDD, *i.e.* $X_0(t) - X_k(t) \to 0$.

Algorithm 4 Distributed control design algorithm for nonlinear agents at each node k with non-uniform small communications delays

Require: Assumptions 1, 2, 3, 5, 6
Require: systems matrices A, B, C
Require: delay bounds $\bar{\delta}_{kj}$ for $j \in \mathcal{N}_k$
Require: gain parameters Q, R^{-1}
Require: Lipschitz constant γ
Require: desired spectral abscissa $-\alpha$ of the closed-loop estimation error dynamics, adaptation step $0 < \delta_\alpha < \alpha$.
Ensure: controller gain $-R^{-1}B^\top P$, observer gain L_k that solve LFPNDD or **fail**.

```
compute P by solving the Riccati equation (24)
compute L_k such that μ(A − L_k C) = −α
α_k ← α
while α_k > 0 & (28) not satisfied do
    α_k ← α_k − δ_α
    compute L_k such that μ(A − L_k C) = −α_k
end while
if α_k > 0 then
    success
else
    fail
end if
```

VI. Example

We consider an example adapted from [46] with heterogeneous constant delays. A network of unmanned aerials vehicles (UAVs) is composed by one leader and 4 followers with state $X(t) \in \mathbb{R}^2$, input $U_k(t) \in \mathbb{R}^2$, and dynamics, $k = 0, \dots, 4$,

$$\dot{X}_k(t) = AX_k(t) + BU_k(t) + \psi(X_k(t)), \tag{35}$$

$$Y_k(t) = CX_k(t), \tag{36}$$

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \ B = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}, \ C = \begin{pmatrix} 1 & 0 \end{pmatrix}$$

$$\psi(X) = \beta \begin{pmatrix} \sin(X_1) \\ \sin(X_2) \end{pmatrix}. \tag{37}$$

Since $\|\psi(X)\|$ in (35) is uniformly bounded and $\sup_X \|\frac{\partial\psi}{\partial X}\| < 2\beta$, Assumption 5 and Assumption 6 hold true and we use the representation (22) with $w_k(t) = v_k(t) \equiv 0$. Condition (34) of Corollary 5 provides a delay bound for asymptotic consensus.

The network topology is shown in Figure 1 (top). The agents are strongly connected, but by using the communication links according to the algorithm of Section III the connection is restricted to the DAG in Figure 1 (bottom). Notice that, by disregarding the link from node 4, the estimation task of node 1 becomes simpler and node 4 is not affected.

We set $\beta = 3 \cdot 10^{-2}$ as in [46]. The control gain $F_k = -R^{-1}B^\top P$ is the same for all the agents ans it is obtained by solving $PA + A^\top P - PBR^{-1}B^\top P + Q = 0$ with $Q = I_2$ and $R = 1$. It is readily verified that this choice satisfies (24) of Assumption 6 and $\sigma(A + BF_k) = \{-1.06 \pm 0.94j\}$.

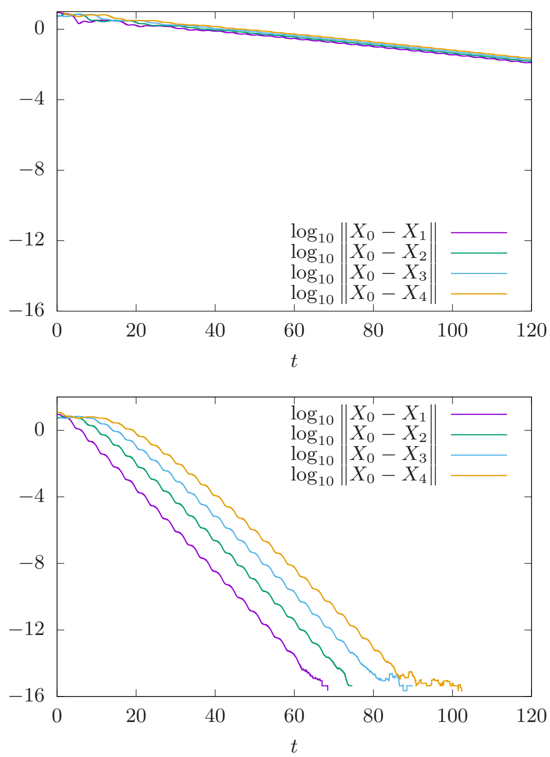


Fig. 2. Logarithm of the norm of the consensus error of the agents for the original network (top) and the modified network (bottom).

The observer gain $L_k = L_{kj}$ is also uniform and it assigns the eigenvalues $\{-0.6 \pm j\}$ to $A - L_k C$ (that is, $\alpha = 0.6$). In this configuration the communications delay bound (34) is $\bar{\delta}_{kj} = 0.508$ for all the links. In the simulations we set the communications delays $\delta_{10} = 0.50$, $\delta_{21} = 0.45$, $\delta_{32} = 0.35$, $\delta_{43} = 0.40$, $\delta_{14} = 0.20$. Fig. 2 (bottom) plots the log-norm of the consensus error of the 4 agents with the network topology of Fig. 1 (bottom), that shows exponential convergence to 0 of the consensus error. In contrast, when using the original topology of Fig. 1 (top) the consensus error convergence is much slower, as shown in Fig. 2 (top), even if the delay of the extra link $4 \to 1$ is small.

Notice that all the followers converge to the leader's trajectory with the same exponential rates (the slope of the coloured lines in Fig. 2, bottom), however the instantaneous disagreement values are different because nodes closer to the leader start converging earlier due to communications delays. However, all trajectories eventually become identical when $\|X_k - X_0\| \simeq 0$. The plateau at 10^{-16} is the limit of numerical precision.

These results, although limited to the consensus algorithm described in this paper, illustrate the de-stabilizing effect of network cycles in presence of delays and support our claim that a suitable weighting of the network connections is the most effective way of dealing with delays in the Leader-Follower context. Finally, we remark that these results compare favourably with those reported in [46], where the total delay, which in that case includes input delay, is $\delta = 0.1$ and practical consensus is reached for $t > 150s$. In our case, practical consensus is reached for $t > 25s$ with a much larger delay. Although the comparison is only partial because the scheme in [46] includes additional modeled disturbances, the proposed method displays better performance in dealing with delays.

VII. Conclusions

In this paper we have provided some evidence for the claim that the presence of delays on networks can be more effectively handled by modifying the network topology. In particular, we have provided the first solution of the leader-following consensus problem of linear systems in presence of arbitrarily large and time-varying communications delay. It should be remarked that this is possible in the leader-following context thanks to the fact that information flows from the leader to the followers, whereas it is less obvious how this approach can be adapted to the case of consensus among agents. When the network topology is modified by following this approach, the convergence of the estimators to the leader's state can be proven locally, thus any other estimator/predictor from delayed measurement can replace the one we have used here. It is therefore simple to extend these results to the case of stochastic or heterogeneous systems.

Appendix

Proof of Theorem 3. Let $\eta_k = X_k - X_0$, $\varepsilon_k = X_0 - \widehat{X}_{0,k}$ and $F = -R^{-1}B^\top P$. With the control input (25) the dynamics of η_k is

$$\begin{aligned}\dot{\eta}_k(t) =& (A + BF)\eta_k(t) + B\Delta U_k(t) + \psi(\eta_k(t) + X_0(t)) \\ & - \psi(X_0(t)) + w_k(t) + \Delta w_k(t),\end{aligned}$$

with $\Delta U_k = F\varepsilon_k$ and $\Delta w_k = -w_0$. Moreover, if $\|\varepsilon_k(t)\|$ is uniformly bounded then $\|\Delta U_k(t)\|$ is uniformly bounded. By Lemma 3 and since $\|\Delta w_k\|_\infty \leq \bar{w} < \infty$, if $\|\varepsilon_k(t)\|$ is uniformly bounded then $\|\eta_k(t)\|$ is uniformly bounded. In order to prove that $\|\varepsilon_k(t)\|$ is uniformly bounded, by subtracting (30) from (22) we obtain, $t \geq 0$,

$$\begin{aligned}\dot{\varepsilon}_k(t) =& A\varepsilon_k(t) + \tilde{\psi}(X_0(t), \varepsilon_k(t)) + w_0(t) \\ & - \sum_{j \in \mathcal{N}^k} (1 - \dot{\delta}_{kj}(t))\bar{\ell}_{kj} e^{\overline{A}_k^0 \delta_{kj}(t)} L_{kj} \left(C\varepsilon_k(t - \delta_{kj}(t))\right. \\ & \left. - C\eta_j(t - \delta_{kj}(t)) + v_j(t - \delta_{kj}(t))\right),\end{aligned}$$

where, for the sake of concision, $\tilde{\psi}(X_0, \varepsilon_k) = \psi(X_0) - \psi(X_0 - \varepsilon_k)$ and $\varepsilon_k(\tau) = X_0(\tau) - \widehat{X}_{0,k}(\tau)$ for $\tau \in [-\bar{\delta}, 0]$. Thanks to the graph structure produced by the algorithm of Section III we can adopt the same inductive procedure as in Theorem 1. We first suppose that the disagreements $\eta_j \equiv 0$, that holds for the immediate followers of the leader. In this case we have the following integral representation for $\varepsilon_k(t)$, valid for $t \geq 0$,

$$\begin{aligned}\varepsilon_k(t) =& \sum_{j \in \mathcal{N}^k} \int_{t-\delta_{kj}(t)}^{t} e^{\overline{A}_k(t-\tau)} \bar{\ell}_{kj} L_{kj} \left(C\varepsilon_k(\tau) + v_j(\tau)\right) \mathrm{d}\tau \\ & + \int_0^t e^{\overline{A}_k(t-\tau)} \Big(\tilde{\psi}(X_0(\tau), \varepsilon_k(\tau)) - v_j(\tau) + w_0(\tau)\Big) \mathrm{d}\tau + b_k.\end{aligned} \quad (38)$$

The value of b_k depends on the pre-shape functions and it can be chosen to make ε_k continuous in $t=0$. Let us define,

$$m_1 = \int_0^\infty \left\| e^{\overline{A}_k s} \right\| \, \mathrm{d}s \; < \; \infty,$$

$$m_2 = \sum_{j\in\mathcal{N}^k} \int_0^{\bar{\delta}_{kj}} \left\| e^{\overline{A}_k \tau} L_{kj} C \right\| \, \mathrm{d}\tau,$$

$$m_3 = \sum_{j\in\mathcal{N}^k} \int_0^{\bar{\delta}_{kj}} \left\| e^{\overline{A}_k s} L_{kj} \right\| \, \mathrm{d}s \; < \; \infty.$$

By taking norms in (38) we get, $t>0$,

$$\begin{aligned}\|\varepsilon_k(t)\| \leq & (\gamma m_1 + m_2) \sup_{\tau\in[0,t]} \|\varepsilon_k(\tau)\| + \|b_k\| + m_1 \bar{w} \\ & + (m_1+m_3)\bar{v}\end{aligned}$$

Since (28) implies that $\gamma m_1 + m_2 < 1$, by taking the sup in $[0,t]$ we obtain the bound

$$\sup_{\tau\in[0,t]} \|\varepsilon_k(\tau)\| \leq \frac{\|b_k\| + m_1\bar{w} + (m_1+m_3)\bar{v}}{1-\gamma m_1 - m_2} \; < \; \infty.$$

This implies that $\|\varepsilon_k(t)\|$ is uniformly bounded in time, and so it is $\|\eta_k(t)\|$, as shown above. The inductive step is proved as above by considering the additional presence of the disagreements η_j, that are however uniformly bounded in norm. ■

Lemma 6: Assume that (A,B) is in Brunowski form, i.e.

$$A = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ldots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix}, B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix},$$

and let $\psi : \mathbb{R}^n \to \mathbb{R}$ be such that $\sup_{X\in\mathbb{R}^n} \|\frac{\partial\psi}{\partial X}(X)\| < +\infty$. There exist positive definite $P, Q \in \mathbb{R}^{n\times n}$ and $R \in \mathbb{R}^{m\times m}$ such that for all $X \in \mathbb{R}^n$

$$\begin{aligned} P\left(A + B\frac{\partial\psi}{\partial X}(X)\right) + \left(A + B\frac{\partial\psi}{\partial X}(X)\right)^\top P \\ - PBR^{-1}B^\top P + Q \leq 0. \end{aligned} \quad (39)$$

Proof. Let I_n denote the $n \times n$ identity matrix. By [59] and since (A,B) is controllable and (I_n, A) is observable, we have the existence of $\Pi \in \mathbb{R}^{n\times n}$ such that

$$\Pi A + A^\top \Pi - \Pi B B^\top \Pi + 2I_n = 0. \quad (40)$$

Define

$$\begin{aligned} P(\varepsilon) := \Gamma^{-1}(\varepsilon)\Pi\Gamma^{-1}(\varepsilon), \Gamma(\varepsilon) := \mathrm{diag}\{\varepsilon^{-1}, \cdots, \varepsilon^{-n}\}, \\ Q(\varepsilon) := \varepsilon^{-1}\Gamma^{-2}(\varepsilon), R(\varepsilon) := \varepsilon^{-2n+1} I_n \end{aligned}$$

where $\varepsilon > 0$ is a parameter to be selected, and consider the matrix

$$\begin{aligned} P(\varepsilon)\left(A + B\frac{\partial\psi}{\partial X}(X)\right) + \left(A + B\frac{\partial\psi}{\partial X}(X)\right)^\top P(\varepsilon) \\ - P(\varepsilon)BR^{-1}(\varepsilon)B^\top P(\varepsilon) + Q(\varepsilon). \end{aligned} \quad (41)$$

By multiplying both sides of (41) by $\Gamma(\varepsilon)$ and since (A,B) is in Brunowski form, we obtain the matrix

$$\begin{aligned} \varepsilon^{-1}\Big\{ \Pi\left(A + \varepsilon B\frac{\partial\psi}{\partial X}(X)S(\varepsilon)\right) \\ + \left(A + \varepsilon B\frac{\partial\psi}{\partial X}(X)S(\varepsilon)\right)^\top \Pi - \Pi B B^\top \Pi + I_n \Big\} \end{aligned}$$

where $S(\varepsilon) := \mathrm{diag}\{\varepsilon^{n-1}, \cdots, \varepsilon, 1\}$. By eq. (40), we have

$$\begin{aligned} & \varepsilon^{-1}\Big\{ \Pi\left(A + \varepsilon B\frac{\partial\psi}{\partial X}(X)S(\varepsilon)\right) \\ & \quad + \left(A + \varepsilon B\frac{\partial\psi}{\partial X}(X)S(\varepsilon)\right)^\top \Pi - \Pi B B^\top \Pi + I_n \Big\} \\ & \leq \Pi B\frac{\partial\psi}{\partial X}(X)S(\varepsilon) + \left(B\frac{\partial\psi}{\partial X}(X)S(\varepsilon)\right)^\top \Pi - \varepsilon^{-1} I_n. \end{aligned}$$

Finally, for the boundedness assumption on $\|\frac{\partial\psi}{\partial X}(X)\|$, pick $\varepsilon > 0$ such that for all $X \in \mathbb{R}^n$

$$\varepsilon\left(\Pi B\frac{\partial\psi}{\partial X}(X)S(\varepsilon) + \left(B\frac{\partial\psi}{\partial X}(X)S(\varepsilon)\right)^\top \Pi\right) \leq I_n,$$

then (41) is negative semi-definite and this proves the claim of the lemma with $P = P(\varepsilon)$, $Q = Q(\varepsilon)$ and $R = R(\varepsilon)$. ■

References

[1] J. Fax and R. Murray, "Information flow and cooperative control of vehicle formations," *IEEE transactions on automatic control*, vol. 49, no. 9, pp. 1465–1476, 2004.

[2] Z. Qu, *Cooperative control of dynamical systems: applications to autonomous vehicles*. Springer Science & Business Media, 2009.

[3] L. Consolini, F. Morbidi, D. Prattichizzo, and M. Tosques, "Leader–follower formation control of nonholonomic mobile robots with input constraints," *Automatica*, vol. 44, no. 5, pp. 1343–1349, 2008.

[4] W. Yu, G. Chen, Z. Wang, and W. Yang, "Distributed consensus filtering in sensor networks," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 6, pp. 1568–1577, 2009.

[5] W. Yu, G. Chen, and M. Cao, "Distributed leader–follower flocking control for multi-agent dynamical systems with time-varying velocities," *Systems & Control Letters*, vol. 59, no. 9, pp. 543–552, 2010.

[6] Y. Tang, H. Gao, W. Zou, and J. Kurths, "Distributed synchronization in networks of agent systems with nonlinearities and random switchings," *IEEE Transactions on Cybernetics*, vol. 43, no. 1, pp. 358–370, 2012.

[7] X. Dong and G. Hu, "Time-varying formation control for general linear multi-agent systems with switching directed topologies," *Automatica*, vol. 73, pp. 47–55, 2016.

[8] Y. Hong, J. Hu, and L. Gao, "Tracking control for multi-agent consensus with an active leader and variable topology," *Automatica*, vol. 42, no. 7, pp. 1177–1182, 2006.

[9] Z. Li, G. Wen, Z. Duan, and W. Ren, "Designing fully distributed consensus protocols for linear multi-agent systems with directed graphs," *IEEE Transactions on Automatic Control*, vol. 60, no. 4, pp. 1152–1157, 2015.

[10] W. Ni and D. Cheng, "Leader-following consensus of multi-agent systems under fixed and switching topologies," *Systems & Control letters*, vol. 59, no. 3-4, pp. 209–217, 2010.

[11] R. Olfati-Saber, J. Fax, and R. Murray, "Consensus and cooperation in networked multi-agent systems," *Proceedings of the IEEE*, vol. 95, no. 1, pp. 215–233, 2007.

[12] W. Ren and R. Beard, "Consensus seeking in multiagent systems under dynamically changing interaction topologies," *IEEE Transactions on Automatic Control*, vol. 50, no. 5, pp. 655–661, 2005.

[13] X. Wang, S. Li, and P. Shi, "Distributed finite-time containment control for double-integrator multiagent systems," *IEEE Transactions on Cybernetics*, vol. 44, no. 9, pp. 1518–1528, 2013.

[14] H. Kim, H. Shim, and J. Seo, "Output consensus of heterogeneous uncertain linear multi-agent systems," *IEEE Transactions on Automatic Control*, vol. 56, no. 1, pp. 200–206, 2010.

[15] Y. Lv, Z. Li, Z. Duan, and J. Chen, "Distributed adaptive output feedback consensus protocols for linear systems on directed graphs with a leader of bounded input," *Automatica*, vol. 74, pp. 308–314, 2016.

[16] J. Wang, A. Lanzon, and I. Petersen, "Robust output feedback consensus for networked negative-imaginary systems," *IEEE Transactions on Automatic Control*, vol. 60, no. 9, pp. 2547–2552, 2015.

[17] T. Li and J.-F. Zhang, "Mean square average-consensus under measurement noises and fixed topologies: Necessary and sufficient conditions," *Automatica*, vol. 45, no. 8, pp. 1929–1936, 2009.

[18] T. Li and J. Zhang, "Consensus conditions of multiagent systems with time-varying topologies and stochastic communication noises," *IEEE Transactions on Automatic Control*, vol. 55(9), pp. 2043–2057, 2010.

[19] L. Cheng, Z. Hou, M. Tan, and X. Wang, "Necessary and sufficient conditions for consensus of double-integrator multi-agent systems with measurement noises," *IEEE Transactions on Automatic Control*, vol. 56(8), pp. 1958–1963, 2011.

[20] L. Cheng, Z. Hou, and M. Tan, "A mean square consensus protocol for linear multi-agent systems with communication noises and fixed topologies," *IEEE Transactions on Automatic Control*, vol. 59(1), pp. 261–267, 2014.

[21] L. Cheng, Y. Wang, W. Ren, Z. Hou, and M. Tan, "On convergence rate of leader-following consensus of linear multi-agent systems with communication noises," *IEEE Transactions on Automatic Control*, vol. 61(11), pp. 3586–3592, 2016.

[22] J. Hu and G. Feng, "Distributed tracking control of leader–follower multi-agent systems under noisy measurement," *Automatica*, vol. 46, no. 8, pp. 1382–1387, 2010.

[23] S. Battilotti and C. Califano, "Leader-following consensus for nonlinear agents with measurement feedback," *International Journal of Robust and Nonlinear Control*, vol. 29, no. 6, pp. 1694–1718, 2019.

[24] Z. Ding and Z. Li, "Distributed adaptive consensus control of nonlinear output-feedback systems on directed graphs," *Automatica*, vol. 72, pp. 46–52, 2016.

[25] Z. Li, W. Ren, X. Liu, and M. Fu, "Consensus of multi-agent systems with general linear and lipschitz nonlinear dynamics using distributed adaptive protocols," *IEEE Transactions on Automatic Control*, vol. 58, no. 7, pp. 1786–1791, 2012.

[26] W. Yu, G. Chen, M. Cao, and J. Kurths, "Second-order consensus for multiagent systems with directed topologies and nonlinear dynamics," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 40, no. 3, pp. 881–891, 2009.

[27] A. Cristofaro and M. Mattioni, "Hybrid consensus for multi-agent systems with time-driven jumps," *Nonlinear Analysis: Hybrid Systems*, vol. 43, p. 101113, 2021.

[28] M. Mattioni and S. Monaco, "Cluster partitioning of heterogeneous multi-agent systems," *Automatica*, vol. 138, p. 110136, 2022.

[29] M. Mattioni, "On multiconsensus of multi-agent systems under aperiodic and asynchronous sampling," *IEEE Control Systems Letters*, vol. 4, no. 4, pp. 839–844, 2020.

[30] M. Huang and J. Manton, "Coordination and consensus of networked agents with noisy measurements: stochastic algorithms and asymptotic behavior," *SIAM Journal on Control and Optimization*, vol. 48, no. 1, pp. 134–161, 2009.

[31] T. Li, F. Wu, and J.-F. Zhang, "Multi-agent consensus with relative-state-dependent measurement noises," *IEEE Transactions on Automatic Control*, vol. 59, no. 9, pp. 2463–2468, 2014.

[32] T. Li and J.-F. Zhang, "Consensus conditions of multi-agent systems with time-varying topologies and stochastic communication noises," *IEEE Transactions on Automatic Control*, vol. 55, no. 9, pp. 2043–2057, 2010.

[33] H. Chu, L. Gao, D. Yue, and C. Dou, "Consensus of lipschitz nonlinear multiagent systems with input delay via observer-based truncated prediction feedback," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 10, pp. 3784–3794, 2019.

[34] C. Wang, Z. Zuo, Z. Lin, and Z. Ding, "A truncated prediction approach to consensus control of lipschitz nonlinear multiagent systems with input delay," *IEEE Transactions on Control of Network Systems*, vol. 4, no. 4, pp. 716–724, 2016.

[35] G. Wen, Y. Yu, Z. Peng, and H. Wang, "Dynamical group consensus of heterogenous multi-agent systems with input time delays," *Neurocomputing*, vol. 175, pp. 278–286, 2016.

[36] Z. Zuo, C. Wang, and Z. Ding, "Robust consensus control of uncertain multi-agent systems with input delay: a model reduction method," *International Journal of Robust and Nonlinear Control*, vol. 27, no. 11, pp. 1874–1894, 2017.

[37] C. Deng, W.-W. Che, and Z.-G. Wu, "A dynamic periodic event-triggered approach to consensus of heterogeneous linear multiagent systems with time-varying communication delays," *IEEE Transactions on Cybernetics*, vol. 51, no. 4, pp. 1812–1821, 2021.

[38] C. Hadjicostis and T. Charalambous, "Average consensus in the presence of delays in directed graph topologies," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 763–768, 2013.

[39] J. Liu, X. Liu, W.-C. Xie, and H. Zhang, "Stochastic consensus seeking with communication delays," *Automatica*, vol. 47, no. 12, pp. 2689–2696, 2011.

[40] K. Liu, G. Xie, and L. Wang, "Containment control for second-order multi-agent systems with time-varying delays," *Systems & Control Letters*, vol. 67, pp. 24–31, 2014.

[41] S. Liu, T. Li, and L. Xie, "Distributed consensus for multiagent systems with communication delays and limited data rate," *SIAM Journal on Control and Optimization*, vol. 49, no. 6, pp. 2239–2262, 2011.

[42] U. Munz, A. Papachristodoulou, and F. Allgower, "Consensus in multi-agent systems with coupling delays and switching topology," *IEEE Transactions on Automatic Control*, vol. 56, no. 12, pp. 2976–2982, 2011.

[43] R. Olfati-Saber and R. Murray, "Consensus problems in networks of agents with switching topology and time-delays," *IEEE Transactions on Automatic Control*, vol. 49, no. 9, pp. 1520–1533, 2004.

[44] H. Ren and F. Deng, "Mean square consensus of leader-following multi-agent systems with measurement noises and time delays," *ISA transactions*, vol. 71, pp. 76–83, 2017.

[45] Y. Tang, H. Gao, W. Zhang, and J. Kurths, "Leader-following consensus of a class of stochastic delayed multi-agent systems with partial mixed impulses," *Automatica*, vol. 53, pp. 346–354, 2015.

[46] C. Wang, Z. Zuo, Z. Qi, and Z. Ding, "Predictor-based extended-state-observer design for consensus of MASs with delays and disturbances," *IEEE Transactions on Cybernetics*, vol. 49, no. 4, pp. 1259–1269, 2018.

[47] Z. Wang, H. Zhang, M. Fu, and H. Zhang, "Consensus for high-order multi-agent systems with communication delay," *Science China Information Sciences*, vol. 60, no. 9, pp. 1–12, 2017.

[48] Y. Zhang, R. Li, W. Zhao, and X. Huo, "Stochastic leader-following consensus of multi-agent systems with measurement noises and communication time-delays," *Neurocomputing*, vol. 282, pp. 136–145, 2018.

[49] B. Zhou and Z. Lin, "Consensus of high-order multi-agent systems with large input and communication delays," *Automatica*, vol. 50, no. 2, pp. 452–464, 2014.

[50] W. Zhu and D. Cheng, "Leader-following consensus of second-order agents with multiple time-varying delays," *Automatica*, vol. 46, no. 12, pp. 1994–1999, 2010.

[51] X. Zong, T. Li, and J.-F. Zhang, "Consensus conditions of continuous-time multi-agent systems with time-delays and measurement noises," *Automatica*, vol. 99, pp. 412–419, 2019.

[52] S. Battilotti, F. Cacace, C. Califano, and M. d'Angelo, "Stochastic leader-following for heterogeneous linear agents with communication delays," *IEEE Transactions on Automatic Control*, 2022.

[53] F. Cacace, M. d'Angelo, and L. Ricciardi Celsi, "Stochastic predictor-based leader-following control with input and communication delays," *International Journal of Control*, pp. 1–12, 2022.

[54] W. He, B. Zhang, Q.-L. Han, F. Qian, J. Kurths, and J. Cao, "Leader-following consensus of nonlinear multiagent systems with stochastic sampling," *IEEE transactions on cybernetics*, vol. 47, no. 2, pp. 327–338, 2016.

[55] Z. Li and Z. Duan, *Cooperative control of multi-agent systems: a consensus region approach*. CRC Press, 2017.

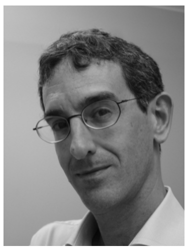
[56] F. Cacace, F. Conte, M. d'Angelo, A. Germani, and G. Palombo, "Filtering linear systems with large time-varying measurement delays," *Automatica*, vol. 136, p. 110084, 2022.

[57] F. Cacace, M. d'Angelo, and A. Germani, "LTV stochastic systems stabilization with large and variable input delay," *Automatica*, vol. 123, p. 109305, 2021.

[58] G. Ciccarella, M. Dalla Mora, and A. Germani, "A Luenberger-like observer for nonlinear systems," *International Journal of Control*, vol. 57, no. 3, pp. 537–556, 1993.

[59] W. Wonham, "On a matrix Riccati equation of stochastic control," *SIAM Journal on Control and Optimization*, vol. 6, no. 4, pp. 681–697, 1968.

Stefano Battilotti Stefano Battilotti was born in Rome, Italy, in 1962. He received his M.S. and Ph.D. degrees in electrical computer engineering from Sapienza University of Rome, respectively in 1987 and 1992. In 1993 he joined the Department of Computer, Control, and Management Engineering "Antonio Ruberti" at Sapienza University of Rome, where since 2005 he is a Professor of Automatic Control. His research interests are focused on observer and control design for nonlinear systems, including stochastic, delay and multi-agent systems.



Filippo Cacace graduated in electronic engineering at Politecnico di Milano in 1988, where he received the PhD degree in computer science in 1992. Since 2003 he has been working with Department of Engineering at Università Campus Bio-Medico di Roma, where he is an associate professor. His current research interests include nonlinear systems and observers, stochastic, distributed and delay systems, system identification, and applications to systems biology.

Claudia Califano received the Ph.D. degree in Systems Engineering in 1998 from the University of Rome La Sapienza. In 1999 she was awarded a CNR-NATO fellowship and a CNR grant. Between 1999 and 2000 she held a post doctoral position at the CNRS Laboratory Laboratoire des Signaux et Syst è mes-Gif sur Yvette-France. In 2009 and 2010 she has held a Maitre de Conference invited position at Ecole Centrale del Nantes, France. She is currently Associate Professor at the Department of Computer, Control, and Management Engineering "Antonio Ruberti" of the University of Rome La Sapienza. She is Senior Member of the IEEE Control System Society. Since 2012 she is a member of the IEEE CSS Technical Committee on Nonlinear Systems and Control and member of the Ifac Technical Committee Nonlinear Control Systems. Since 2014 she serves as an Associate Editor for IMA Journal of Mathematical Control and Information. Her main research interests concern theoretical aspect of nonlinear discrete time systems and continuous time delay systems.

Massimiliano d'Angelo was born in Italy in 1992. He received the M.Sc. degree and the Ph.D. degree (summa cum laude) in Control Engineering in 2016 and 2020, respectively, from the University of Rome La Sapienza where he is currently a PostDoc researcher. He has been a visiting student at Washington University in St. Louis in 2019 and a PostDoc at the University of Milano-Bicocca in 2020. His research interests include estimation and stochastic control theory, time-delay systems and systems biology.