

Distributed Decision Through Self-Synchronizing Sensor Networks in the Presence of Propagation Delays and Asymmetric Channels

Gesualdo Scutari, Sergio Barbarossa and Loreto Pescosolido

Dpt. INFOCOM, Univ. of Rome “La Sapienza”, Via Eudossiana 18, 00184 Rome, Italy

E-mail: {scutari, sergio, loreto}@infocom.uniroma1.it.*

Manuscript submitted on January 10, 2007. Finally accepted on September 15, 2007.

Abstract

In this paper we propose and analyze a distributed algorithm for achieving globally optimal decisions, either estimation or detection, through a self-synchronization mechanism among linearly coupled integrators initialized with local measurements. We model the interaction among the nodes as a directed graph with weights (possibly) dependent on the radio channels and we pose special attention to the effect of the propagation delay occurring in the exchange of data among sensors, as a function of the network geometry. We derive necessary and sufficient conditions for the proposed system to reach a consensus on globally optimal decision statistics. One of the major results proved in this work is that a consensus is reached with exponential convergence speed for any bounded delay condition if and only if the directed graph is quasi-strongly connected. We provide a closed form expression for the global consensus, showing that the effect of delays is, in general, the introduction of a bias in the final decision. Finally, we exploit our closed form expression to devise a double-step consensus mechanism able to provide an unbiased estimate with minimum extra complexity, without the need to know or estimate the channel parameters.

1 Introduction and Motivations

Endowing a sensor network with self-organizing capabilities is undoubtedly a useful goal to increase the resilience of the network against node failures (or simply switches to sleep mode) and avoid potentially dangerous congestion conditions around the sink nodes. Decentralizing decisions decreases also the vulnerability of the network against damages to the sink or control nodes. Distributed computation over a network and its application to statistical consensus theory has a long history (see, e.g., [1, 2]), including the pioneering work of Tsitsiklis, Bertsekas and Athans on asynchronous agreement problem for discrete-time distributed decision-making systems [3] and parallel computing [4, 5]. A sensor network may be seen indeed as a sort of distributed computer that has to evaluate a function of the

*This work has been partially funded by the WINSOC project, a Specific Targeted Research Project (Contract Number 0033914) co-funded by the INFSO DG of the European Commission within the RTD activities of the Thematic Priority Information Society Technologies, and by ARL/ERO Contract N62558-05-P-0458.

Part of this work was presented in the *IEEE Workshop on Signal Processing Advances in Wireless Communications, (SPAWC-2006)*, July 2006, for undirected graphs.

measurements gathered by each sensor, possibly without the need of a fusion center. This problem may be addressed by taking into account the vast literature on distributed consensus/agreement algorithms. These techniques have received great attention in the recent years within the literature on cooperative control and multiagent systems [6]–[22]. In particular, the conditions for achieving a consensus over a common specified value, like a linear combination of the observations, was solved for networked *continuous-time* dynamic systems by Olfati-Saber and Murray, under a variety of network topologies, also allowing for topology variations during the time necessary to achieve consensus [7, 8]. The *discrete-time* case was addressed by Tsitsiklis in [4] (see also [5]). Many recent works focused on the distributed computation of more general functions than the average of the initial measurements. These include average-max-min consensus [12], geometric mean consensus [13], and power mean consensus [14]. A study on the class of smooth functions that can be computed by distributed consensus algorithms was recently addressed in [15]. Application of consensus algorithms to data fusion problem and distributed filtering was proposed in [16, 17]. The study of the so called *alignment* problem, where all the agents eventually reach an agreement, but without specifying how the final value is related to the initial measurements, was carried out in [18]–[22]. Two recent excellent tutorials on distributed consensus and agreement techniques are [23] and [24].

Consensus may be also seen as a form of self-synchronization among coupled dynamical systems. In [25, 26], it was shown how to use the self-synchronization capabilities of a set of nonlinearly coupled dynamical systems to reach the *globally optimal* Best Linear Unbiased Estimator (BLUE), assuming symmetric communication links. As well known, the BLUE estimator coincides with the Maximum Likelihood (ML) estimate for the linear observation model with additive Gaussian noise. In particular, it was shown in [26] that reaching a consensus on the state derivative, rather than on the state itself (as in [7]–[24]), allows for better resilience against coupling noise. In [27], the authors provided a discrete-time decentralized consensus algorithm to compute the BLUE, based on linear coupling among the nodes, as a result of a distributed optimization incorporating the coupling noise.

The consensus protocols proposed in [7], [9]–[21] and [24]–[27] assume that the interactions among the nodes occur instantaneously, i.e., without any propagation delay. However, this assumption is not valid for large scale networks, where the distances among the nodes are large enough to introduce a nonnegligible communication delay. There are only a few recent works that study the consensus or the agreement problem, in the presence of propagation delays, namely [8], [28]–[32], [47] focusing on *continuous-time* systems, and [3, 4], [5, Ch. 7.3], [22], dealing with *discrete-time* protocols. More specifically, in [8, 28] the authors provide sufficient conditions for the convergence of a linear average consensus protocol, in the case of time-invariant *homogeneous* delays (i.e., equal delay for all the links) and assuming *symmetric* communication links. The most appealing feature of the dynamical system in [8] is the convergence of all the state variables to a *known, delay-independent*, function (equal to the average) of the observations. However, this desired property is paid in terms of convergence, since, in the presence of homogeneous delays, the system of [8] is able to reach a consensus if and only if the (common) delay is smaller than a given, topology-dependent, value. Moreover, the assumption of homogeneous delays and symmetric links is not appropriate for modeling the propagation delay in a typical wireless network, where the delays are proportional to the distance among the nodes and communication channels are typically *asymmetric*.

The protocol of [8] was generalized in [29] to time-invariant *inhomogeneous* delays (but symmetric channels) and in [30] to asymmetric channels. The dynamical systems studied in [29, 30] are guaranteed to reach an agreement, for any given set of finite propagation delays, provided that the network is *strongly connected*. Similar results, under weaker (sufficient) conditions on the (possibly time-varying) network topology, were obtained in [3, 4], [5, Ch.7.3] and [22] for the convergence of discrete-time asynchronous agreement algorithms. However, since the final agreement value is not a known function of the local measurements, agreement algorithms proposed in the cited papers are mostly appropriate for the alignment of mobile agents, but they cannot be immediately used to distributively compute prescribed functions of the sensors' measurements, like decision tests or global parameter estimates. In [47], the authors studied the convergence properties of the agreement algorithms proposed in [29]–[31] and provided a closed form expression of the achievable consensus in the case of strongly connected balanced digraphs, under the assumption that the initial conditions are nonnegative and the state trajectories remain in the nonnegative orthant of the state space. However, the final agreement value depends on delays, network topology and initial conditions of each node, so that the bias is practically unavoidable. In summary, in the presence of propagation delays, classical distributed protocols reaching consensus/agreement on the state [6]–[24], [27]–[31], [47] cannot be used to achieve prescribed functions of the sensors' measurements that are not biased by the channel parameters.

Ideally, we would like to have a totally decentralized system able to reach a global consensus on a final value which is a *known, delay-independent*, function of the local measurements (as in [8]), for *any* given set of *inhomogeneous* propagation delays (as in [29]) and *asymmetric* channels (as in [30]). In this paper, we propose a distributed dynamical system having all the above desired features. More specifically, we consider a set of linearly coupled first-order dynamical systems and we fully characterize its convergence properties, for a network with arbitrary topology (not necessarily strongly connected, as opposed to [29]–[31]) and (possibly) asymmetric communication channels. In particular, we consider an interaction model among the sensors that is directly related to the physical channel parameters. The network is modeled as a weighted directed graph, whose weights are directly related to the (flat-fading) channel coefficients between the nodes and to the transmit power. Furthermore, the geometry-dependent propagation delays between each pair of nodes, as well as possible time offsets among the nodes, are properly taken into account. The most appealing feature of the proposed system is that a consensus on a globally optimal decision statistic is achieved, for *any* (bounded) set of inhomogeneous delays and for *any* set of (asymmetric) communication channels, with the only requirement that the network be quasi-strongly connected and the channel coefficients be nonnegative.¹

In particular, our main contributions are the following: i) We provide the necessary and sufficient conditions ensuring local or global convergence, for *any* set of finite propagation delays and network topology; ii) We prove that the convergence is exponential, with convergence rate depending, in general, on the channel parameters and propagation delays; iii) We derive a closed form expression for the final consensus, as a function of the attenuation coefficients and propagation delays of each link; iv) We show how to get a final, *unbiased* function of the sensor's measurements, which coincides with the globally optimal decision statistics that it would have been computed by a fusion center having error-

¹This last requirement, if not immediately satisfied, requires some form of phase compensation at the receiver.

free access to all the nodes. The paper is organized as follows. Section 2 describes the proposed first-order linearly coupled dynamical system and shows how to design the system's parameters and the local processing so that the state derivative of each node converges, asymptotically, to the globally optimal decision statistics. Section 3 contains the main results of the paper, namely the necessary and sufficient conditions ensuring the global or local convergence of the proposed dynamical system, in the presence of inhomogeneous propagation delays and asymmetric channels. Finally, Section 4 contains numerical results validating our theoretical findings and draws some conclusions.

2 Reaching Globally Optimal Decisions Through Self-Synchronization

It was recently shown in [33] that, in many applications, an efficient sensor network design should incorporate some sort of in-network processing. In this paper, we show first a class of functions that can be computed with a totally distributed approach. Then, we illustrate the distributed mechanism able to achieve the globally optimal decision tests.

2.1 Forms of Consensus Achievable With a Decentralized Approach

If we denote by y_i , $i = 1, \dots, N$, the (scalar) measurement taken from node i , in a network composed of N nodes, we have shown in [25, 26] that it is possible to compute any function of the collected data expressible in the form

$$f(y_1, y_2, \dots, y_N) = h \left(\frac{\sum_{i=1}^N c_i g_i(y_i)}{\sum_{i=1}^N c_i} \right), \quad (1)$$

where $\{c_i\}$ are positive coefficients and $g_i(\cdot)$, $i = 1, \dots, N$, and $h(\cdot)$ are arbitrary (possibly nonlinear) real functions on \mathbb{R} , i.e., $g_i, h : \mathbb{R} \mapsto \mathbb{R}$, in a totally decentralized way, i.e., without the need of a sink node. In the vector observation case, the function may be generalized to the vector form

$$\mathbf{f}(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N) = \mathbf{h} \left(\left(\sum_{i=1}^N \mathbf{C}_i \right)^{-1} \left(\sum_{i=1}^N \mathbf{C}_i \mathbf{g}_i(\mathbf{y}_i) \right) \right), \quad (2)$$

where $\mathbf{y}_i = \{y_{i,k}\}_{k=1}^L$ is the vector containing the observations $\{y_{i,k}\}_{k=1}^L$ taken by sensor i , and $\mathbf{g}_i(\cdot)$ and $\mathbf{h}(\cdot)$ are arbitrary (possibly nonlinear) real functions on \mathbb{R}^L , i.e., $\mathbf{g}_i, \mathbf{h} : \mathbb{R}^L \mapsto \mathbb{R}$, and $\{\mathbf{C}_i\}$ are arbitrary square positive definite matrices.

Even though the class of functions expressible in the form (1) or (2) is not the most general one, nevertheless it includes many cases of practical interest, as shown in the following examples.

Example 1: ML or BLUE estimate. Let us consider the case where each sensor observes a vector in the form

$$\mathbf{y}_i = \mathbf{A}_i \boldsymbol{\xi} + \mathbf{w}_i, \quad i = 1, \dots, N, \quad (3)$$

where \mathbf{y}_i is the $M \times 1$ observation vector, $\boldsymbol{\xi}$ is the $L \times 1$ unknown common parameter vector, \mathbf{A}_i is the $M \times L$ mixing matrix of sensor i , and \mathbf{w}_i is the observation noise vector, modeled as a circularly symmetric Gaussian vector with zero mean and covariance matrix \mathbf{R}_i . We assume that the noise vectors affecting different sensors are statistically independent of each other, and that each matrix

\mathbf{A}_i is full column rank, which implies $M \geq L$. As well known, in this case the globally optimal ML estimate of $\boldsymbol{\xi}$ is [34]:

$$\hat{\boldsymbol{\xi}}_{ML} = \mathbf{f}(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N) = \left(\sum_{i=1}^N \mathbf{A}_i^T \mathbf{R}_i^{-1} \mathbf{A}_i \right)^{-1} \left(\sum_{i=1}^N \mathbf{A}_i^T \mathbf{R}_i^{-1} \mathbf{y}_i \right). \quad (4)$$

This expression is a special case of (2), with $\mathbf{C}_i = \mathbf{A}_i^T \mathbf{R}_i^{-1} \mathbf{A}_i$ and $\mathbf{g}_i(\mathbf{y}_i) = (\mathbf{A}_i^T \mathbf{R}_i^{-1} \mathbf{A}_i)^{-1} \mathbf{A}_i^T \mathbf{R}_i^{-1} \mathbf{y}_i$. If the noise pdf is unknown, (4) still represents a meaningful estimator, as it is the BLUE [34].

Example 2: Detection of a Gaussian process with unknown variance embedded in Gaussian noise with known variance. Let us consider now a detection problem. Let $y_i[k]$ denote the signal observed by sensor i , at time k . The detection problem can be cast as a binary hypothesis test, where the two hypotheses are

$$\begin{aligned} \mathcal{H}_0 : \quad & y_i[k] = w_i[k], \\ \mathcal{H}_1 : \quad & y_i[k] = s_i[k] + w_i[k], \end{aligned} \quad i = 1, \dots, N, \quad k = 1, \dots, K, \quad (5)$$

where $s_i[k]$ is the useful signal and $w_i[k]$ is the additive noise. Let us consider the case where the random sequences $s_i[k]$ and $w_i[k]$ are spatially uncorrelated and modeled as zero mean independent Gaussian random processes. A meaningful model consists in assuming that the noise variance is known, let us say equal to σ_w^2 , whereas the useful signal variance is not. Under these assumptions, the optimal detector consists in computing the generalized likelihood ratio test (GLRT) [34, 35]

$$T(\mathbf{y}) = f(\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N) = \frac{1}{K} \sum_{i=1}^N \sum_{k=1}^K y_i^2[k] \left(\frac{1}{\sigma_w^2} - \frac{1}{\hat{P}_i + \sigma_w^2} \right) - \sum_{i=1}^N \log \left(\frac{\hat{P}_i + \sigma_w^2}{\sigma_w^2} \right), \quad (6)$$

and comparing it with a threshold that depends on the desired false alarm rate. In (6), the term \hat{P}_i denotes the ML estimate of the signal power at node i , given by

$$\hat{P}_i = \left(\frac{1}{K} \sum_{k=1}^K y_i^2[k] - \sigma_w^2 \right)^+,$$

where $(x)^+ \triangleq \max(0, x)$. Also in this case, it is easy to check that (6) is a special case of (1), with $c_i = 1$ and $g_i(\mathbf{y}_i) = \frac{1}{K} \sum_{k=1}^K y_i^2[k] \left(\frac{1}{\sigma_w^2} - \frac{1}{\hat{P}_i + \sigma_w^2} \right) - \log \left(\frac{\hat{P}_i + \sigma_w^2}{\sigma_w^2} \right)$ [35].

These are only two examples, but the expressions (1) or (2) can be used to compute or approximate more general functions of the collected data, like, maximum and minimum [12], geometric mean [13], power mean [14] and so on, through appropriate choice of the parameters involved.

2.2 How to Achieve the Consensus in a Decentralized Way

The next question is how to achieve the aforementioned optimal decision statistics with a network having no fusion center. In [25, 26] we proposed an approach to solve this problem using a nonlinear interaction model among the nodes, based on an undirected graph and with no propagation delays in the exchange of information among the sensors. In this paper, we consider a linear interaction model, but we generalize the approach to a network where the propagation delays are taken into account and the network is described by a weighted *directed* graph (or *digraph*, for short), which is a fairly general

model to capture the non reciprocity of the communication links governing the interaction among the nodes.

The proposed sensor network is composed of N nodes, each equipped with four basic components: i) a *transducer* that senses the physical parameter of interest (e.g., temperature, concentration of contaminants, radiation, etc.); ii) a *local processing unit* that processes the measurement taken by the node; iii) a *dynamical system*, whose state is initialized with the local measurements and it evolves interactively with the states of nearby sensors; iv) a *radio interface* that transmits the state of the dynamical system and receives the state transmitted by the other nodes, thus ensuring the interaction among the nodes².

Scalar observations. In the scalar observation case, the dynamical system present in node i evolves according to the following linear functional differential equation

$$\begin{aligned} \dot{x}_i(t; \mathbf{y}) &= g_i(y_i) + \frac{K}{c_i} \sum_{j \in \mathcal{N}_i} a_{ij} (x_j(t - \tau_{ij}; \mathbf{y}) - x_i(t; \mathbf{y})), \quad t > 0, \\ x_i(t; \mathbf{y}) &= \tilde{\phi}_i(t), \quad t \in [-\tau, 0], \end{aligned} \quad i = 1, \dots, N, \quad (7)$$

where $x_i(t; \mathbf{y})$ is the state function associated to the i -th sensor that depends on the set of measurements $\mathbf{y} = \{y_i\}_{i=1}^N$; $g_i(y_i)$ is a function of the local observation y_i , whose form depends on the specific decision test; K is a positive coefficient measuring the global coupling strength; c_i is a positive coefficient that may be adjusted to achieve the desired consensus; $\tau_{ij} = T_{ij} + d_{ij}/c$ is a delay incorporating the propagation delay due to traveling the internode distance d_{ij} , at the speed of light c , plus a possible time offset T_{ij} between nodes i and j . The sensors are assumed to be fixed so that all the delays are constant. We also assume, realistically, that the maximum delay is bounded, with maximum value $\tau = \max_{ij} \tau_{ij}$. The coefficients a_{ij} are assumed to be nonnegative and, in general, dependent on transmit powers and channel parameters. For example, a_{ij} may represent the amplitude of the signal received from node i and transmitted from node j . In such a case, we have $a_{ij} = \sqrt{P_j |h_{ij}|^2 / d_{ij}^\eta}$, where P_j is the power of the signal transmitted from node j ; h_{ij} is a fading coefficient describing the channel between nodes j and i ; η is the path loss exponent. The nonnegativity of a_{ij} requires some form of channel compensation at the receiver side, like in the maximal ratio receiver, for example, if the channel coefficients $\{h_{ij}\}_{ij}$ are complex variables. Furthermore, we assume, realistically, that node i “hears” node j only if the power received from i exceeds a given threshold. In such a case, $a_{ij} \neq 0$, otherwise $a_{ij} = 0$. The set of nodes that sensor i hears is denoted by $\mathcal{N}_i = \{j = 1, \dots, N : a_{ij} \neq 0\}$. Observe that, in general, $a_{ij} \neq a_{ji}$, i.e., the channels are asymmetric. It is worth noticing that the state function of, let us say, node i depends, directly, only on the measurement y_i taken by the node itself and only indirectly on the measurements gathered by the other nodes. In other words, even though the state $x_i(t; \mathbf{y})$ gets to depend, eventually, on all the measurements, through the interaction with the other nodes, each node needs to know only its own measurement.

Because of the delays, the state evolution (7) for, let us say, $t > 0$, is uniquely defined provided

²The state value is transmitted by modulating an appropriate carrier. Because of space limitations, it is not possible to go into the details of this aspect in this paper. In parallel works, we have shown that a pulse position modulation of ultrawideband signals may be a valid candidate for implementing the interaction mechanism. Some preliminary remarks on the implementation of the proposed protocol can be found in [38].

that the initial state variables $x_i(t; \mathbf{y})$ are specified in the interval from $-\tau$ to 0. The initial conditions of (7) are assumed to be taken in the set of continuously differentiable and bounded functions $\tilde{\phi}_i(t)$ mapping the interval $[-\tau, 0]$ to \mathbb{R} (see Appendix B for more details).

Vector observations. If each sensor measures a set of, let us say, L physical parameters, the coupling mechanism (7) generalizes into the following expression³

$$\begin{aligned} \dot{\mathbf{x}}_i(t; \mathbf{y}) &= \mathbf{g}_i(\mathbf{y}_i) + K \mathbf{Q}_i^{-1} \sum_{j \in \mathcal{N}_i} a_{ij} (\mathbf{x}_j(t - \tau_{ij}; \mathbf{y}) - \mathbf{x}_i(t; \mathbf{y})), \quad t > 0, \\ \mathbf{x}_i(t; \mathbf{y}) &= \tilde{\phi}_i(t), \quad t \in [-\tau, 0], \end{aligned} \quad i = 1, \dots, N, \quad (8)$$

where $\mathbf{x}_i(t)$ is the L -size vector state of the i -th node; $\mathbf{g}_i(\mathbf{y}_i)$ is a vector function of the local observation $\mathbf{y}_i = \{y_{i,k}\}_{k=1}^L$, i.e. $\mathbf{g}_i: \mathbb{R}^L \mapsto \mathbb{R}^L$; and \mathbf{Q}_i is an $L \times L$ non-singular matrix that is a free parameter to be chosen according to the specific purpose of the sensor network.

As in (7), the initial conditions of (8) are assumed to be arbitrarily taken in the set of continuously differentiable and bounded (vectorial) functions $\tilde{\phi}_i(t)$ mapping the interval $[-\tau, 0]$ to \mathbb{R}^L , w.l.o.g..

2.3 Self-Synchronization

Differently from most papers dealing with average consensus problems [6]–[24] and [27]–[31], where the global consensus was intended to be the situation where all dynamical systems reach the same state value, we adopt here the alternative definition already introduced in our previous work [26]. We define the consensus (through network synchronization) with respect to the state *derivative*, rather than to the state, as follows.

Definition 1 *Given the dynamical system in (7) (or (8)), a solution $\{\mathbf{x}_i^*(t; \mathbf{y})\}$ of (7) (or (8)) is said to be a synchronized state of the system, if*

$$\dot{\mathbf{x}}_i^*(t; \mathbf{y}) = \boldsymbol{\omega}^*(\mathbf{y}), \quad \forall i = 1, 2, \dots, N. \quad (9)$$

The system (7) (or (8)) is said to globally synchronize if there exists a synchronized state as in (9), and all the state derivatives asymptotically converge to this common value, for any given set of initial conditions $\{\tilde{\phi}_i\}$, i.e.,

$$\lim_{t \rightarrow \infty} \|\dot{\mathbf{x}}_i(t; \mathbf{y}) - \boldsymbol{\omega}^*(\mathbf{y})\| = 0, \quad \forall \tilde{\phi}_i, \quad i = 1, 2, \dots, N, \quad (10)$$

where $\|\cdot\|$ denotes some vector norm and $\{\mathbf{x}_i(t; \mathbf{y})\}$ is a solution of (7) (or (8)). The synchronized state is said to be globally asymptotically stable if the system globally synchronizes, in the sense specified in (10). The system (7) (or (8)) is said to locally synchronize if there exist disjoint subsets of the nodes, called clusters, where the nodes in each cluster have state derivatives converging, asymptotically, to the same value, for any given set of initial conditions $\{\tilde{\phi}_i\}$.

Observe that, according to Definition 1, if there exists a globally asymptotically stable synchronized state, then it must necessarily be *unique* (in the derivative). In the case of local synchronization instead, the system may have multiple synchronized clusters, each of them with a different synchronized state.

³We assume that the coupling coefficients a_{ij} are the same for all estimated parameters. This assumption is justified by the fact that all the parameters are sent through the same physical channel.

As it will be shown in the next section, one of the reasons to introduce a novel definition of consensus, different from the classical one [6]–[24], [27]–[31], is that the convergence on the state derivative, rather than on the state, is not affected by the presence of propagation delays and there is a way to make the final consensus value to coincide with the globally optimal decision statistic. In the ensuing sections, we will provide necessary and sufficient conditions for the system in (7) (or (8)) to locally/globally synchronize according to Definition 1, along with the closed form expression of the synchronized state.

3 Necessary and Sufficient Conditions for Self-Synchronization

The problem we address now is to check if, in the presence of propagation delays and asymmetric communication links, the systems (7) and (8) can still be used to achieve globally optimal decision statistics in the form (1) and (2), in a totally distributed way. To derive our main results, we rely on some basic notions of digraph theory. To make the paper self-contained, in Appendix A.1 we recall the basic definitions of weak, quasi-strong and strong connectivity (WC, QSC, and SC, for short) of a digraph. In Appendix A.2, we recall the algebraic properties of the Laplacian matrix associated to a digraph and we derive the properties of its left eigenvectors, as they will play a fundamental role in computing the achievable forms of consensus. In Appendix B, we provide sufficient conditions for the marginal stability of linear delayed differential equations, as they will be instrumental to prove the main results of this paper.

The next theorem provides necessary and sufficient conditions for the proposed decentralized approach to achieve the desired consensus in the presence of propagation delays and asymmetric communication links.

Theorem 1 *Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ be the digraph associated to the network in (7), with Laplacian matrix \mathbf{L} . Let $\boldsymbol{\gamma} = [\gamma_1, \dots, \gamma_N]^T$ be a left eigenvector of \mathbf{L} corresponding to the zero eigenvalue, i.e., $\boldsymbol{\gamma}^T \mathbf{L} = \mathbf{0}_N^T$. Given the system in (7), assume that the following conditions are satisfied:*

- a1** *The coupling gain K and the coefficients $\{c_i\}$ are positive; the coefficients $\{a_{ij}\}$ are non-negative;*
- a2** *The propagation delays $\{\tau_{ij}\}$ are time-invariant and finite, i.e., $\tau_{ij} \leq \tau = \max_{i \neq j} \tau_{ij} < +\infty$, $\forall i \neq j$;*
- a3** *The initial conditions are taken in the set of continuously differentiable and bounded functions mapping the interval $[-\tau, 0]$ to \mathbb{R}^N .*

Then, system (7) globally synchronizes for any given set of propagation delays, if and only if the digraph \mathcal{G} is Quasi-Strongly Connected (QSC). The synchronized state is given by

$$\boldsymbol{\omega}^*(\mathbf{y}) = \frac{\sum_{i=1}^N \gamma_i c_i g_i(y_i)}{\sum_{i=1}^N \gamma_i c_i + K \sum_{i=1}^N \sum_{j \in \mathcal{N}_i} \gamma_i a_{ij} \tau_{ij}}, \quad (11)$$

where $\gamma_i > 0$ if and only if node i can reach all the other nodes of the digraph through a strong path, or $\gamma_i = 0$, otherwise.

The convergence is exponential, with rate arbitrarily close to $r \triangleq -\min_i \{|\operatorname{Re}\{s_i\}|\} : p(s_i) = 0$ and $s_i \neq 0\}$, where $p(s)$ is the characteristic function associated to system (7) (cf. Appendix B).

Proof. See Appendix C. ■

Theorem 1 has a very broad applicability, as it does not make any particular reference to the network topology. If, conversely, the topology has a specific structure, then we may have the following forms of consensus.⁴

Corollary 1 *Given system (7), assume that conditions **a1)**-**a3)** of Theorem 1 are satisfied. Then,*

1. *The system globally synchronizes and the synchronized state is given by*

$$\omega^*(\mathbf{y}) = g_r(y_r), \quad r \in \{1, 2, \dots, N\}, \quad (12)$$

if and only if the digraph \mathcal{G} contains only one spanning directed tree, with root node given by node r .

2. *The system globally synchronizes and the synchronized state is given by (11) with all γ_i 's positive if and only if the digraph \mathcal{G} is Strongly Connected (SC). The synchronized state becomes*

$$\omega^*(\mathbf{y}) = \frac{\sum_{i=1}^N c_i g_i(y_i)}{\sum_{i=1}^N c_i + K \sum_{i=1}^N \sum_{j \in \mathcal{N}_i} a_{ij} \tau_{ij}}, \quad (13)$$

if and only if, in addition, the digraph \mathcal{G} is balanced.

3. *The system locally synchronizes in K disjoint clusters $\mathcal{C}_1, \dots, \mathcal{C}_K \subseteq \{1, \dots, N\}$,⁵ with synchronized state derivatives for each cluster*

$$\dot{x}_q^*(t; \mathbf{y}_k) = \frac{\sum_{i \in \mathcal{C}_k} \gamma_i c_i g_i(y_i)}{\sum_{i \in \mathcal{C}_k} \gamma_i c_i + K \sum_{i \in \mathcal{C}_k} \sum_{j \in \mathcal{N}_i} \gamma_i a_{ij} \tau_{ij}}, \quad \forall q \in \mathcal{C}_k, \quad k = 1, \dots, K, \quad (14)$$

with $\mathbf{y}_k = \{y_i\}_{i \in \mathcal{C}_k}$, if and only if the digraph \mathcal{G} is weakly connected (WC) and contains a spanning directed forest with K Root Strongly Connected Components (RSCC).⁶

The proof of this corollary is a particular case of Theorem 1, except that we exploit the structure of the left eigenvector corresponding to the zero eigenvalue of the Laplacian matrix \mathbf{L} , as derived in Appendix A.2. The previous results can be extended to the vector case in (8), according to the following.⁷

Theorem 2 *Given the system (8), assume that conditions **a2)**-**a3)** of Theorem 1 are satisfied and that the matrices $\{\mathbf{Q}_i\}$ are positive definite. Then, the system synchronizes for any given set of propagation delays, if and only if the digraph \mathcal{G} is QSC. The synchronized state is given by*

$$\omega^*(\mathbf{y}) \triangleq \left(\sum_{i=1}^N \gamma_i \mathbf{Q}_i + \mathbf{I}_L \otimes \left(K \sum_{i=1}^N \sum_{j \in \mathcal{N}_i} \gamma_i a_{ij} \tau_{ij} \right) \right)^{-1} \left(\sum_{i=1}^N \gamma_i \mathbf{Q}_i \mathbf{g}_i(\mathbf{y}_i) \right), \quad (15)$$

⁴We focus, w.l.o.g., only on WC digraphs. In the case of non WC digraphs, Corollary 1 applies to each disjoint component of the digraph.

⁵In general, the clusters $\mathcal{C}_1, \dots, \mathcal{C}_K$ are not a partition of the set of nodes $\{1, \dots, N\}$.

⁶Please, see Appendix A.1 for the formal definition of (root) strongly connected component.

⁷The proof is omitted because of space limitations, but it follows the same guidelines as the proof for the scalar case, since the interaction among the nodes is the same.

where \otimes denotes the Kronecker product and $\gamma_i > 0$ if and only if node i can reach all the other nodes of the digraph by a strong path, or $\gamma_i = 0$, otherwise. The convergence is exponential, with rate arbitrarily close to $r \triangleq -\min_i\{|\operatorname{Re}\{s_i\}| : p(s_i) = 0 \text{ and } s_i \neq 0\}$, where $p(s)$ is the characteristic function associated to system (8).

3.1 Impact of propagation delays on convergence

The impact of delays in consensus-achieving algorithms has been analyzed in a series of works [5, Ch.7.3], [8], [22], [28]–[31]. Among these works, it is useful to distinguish between consensus algorithms, [8, 23], where the states of all the sensors converge to a prescribed function (typically the average) of the sensors’ initial values, and agreement algorithms, [5, Ch.7.3], [22], [28]–[31], [47], where the goal is to make all the states to converge to a common value, but without specifying how this value has to be related to the initial values. In our application, we can only rely on consensus algorithms, where the final consensus has to coincide with the globally optimal decision statistic.

The consensus algorithms analyzed in [8, 23] assume the same delay value for all the links, i.e., $\tau_{ij} = \tau$, and symmetric channels, i.e., $a_{ij} = a_{ji}$, for all $i \neq j$. Under these assumptions, the average consensus in [8, 23] is reached if and only if the common delay τ is upper bounded by [8]

$$\tau < \frac{\pi}{2} \frac{1}{\lambda_N}, \quad (16)$$

where λ_N denotes the maximum eigenvalue of the Laplacian associated to the undirected graph of the network (cf. Appendix A). Since, for any graph with N nodes, we have [39]

$$\frac{N}{N-1} d_{\max} \leq \lambda_N \leq 2 d_{\max}, \quad (17)$$

with d_{\max} denoting the maximum graph degree, increasing d_{\max} imposes a strong constraint on the maximum (common) tolerable delay. This implies, for example, that networks with hubs (i.e., nodes with very large degrees) that are commonly encountered in scale-free networks [36], are fragile against propagation delays, when using the consensus algorithms of [8, 23].

In our application, we were motivated to extend the approach of [8, 23] to the general case of inhomogeneous delays and asymmetric channel links. Nevertheless, in spite of the less restrictive assumptions, Theorem 1 shows that our proposed algorithm is more robust against propagation delays, since its convergence capability *is not affected by the delays*. Moreover, our approach is valid in the more general case of asymmetric communications links and the final value is not simply the average of the measurements, but a weighted average of functions of the measurements that can be made to coincide with the desired globally optimal decision statistic in the form (1) or (2), through a proper choice of the coefficients c_i .

An intuitive reason for explaining the main advantage of our approach is related to the use of an alternative definition of global consensus: As opposed to conventional methods requiring consensus on the state value, i.e. [6]–[24], [28]–[31], we require the convergence over the state derivative, i.e., we only require that the state trajectories converge towards parallel straight lines. The slope must be the same for all the trajectories and it has to coincide with the desired decision statistic. But the constant terms of each line may differ from sensor to sensor. This provides additional degrees of freedom that, eventually, make our approach more robust against propagation delays or link coefficients.

3.2 Effect of network topology on final consensus value

Theorem 1 generalizes all the previous (only sufficient) conditions known in the literature [8], [28]–[31] for the convergence of linear agreement/consensus protocols in the presence of propagation delays, since it provides a complete characterization of the synchronization capability of the system for any possible degree of connectivity in the network (not only for SC digraphs), as detailed next.

In general, the digraph modeling the interaction among the nodes may have one of the following structures: i) the digraph contains only one spanning directed tree, with a single root node, i.e., there exists only one node that can reach all the other nodes in the network through a strong directed path; ii) the digraph contains more than one spanning directed tree, i.e., there exist multiple nodes (possibly all the nodes), strongly connected to each other, that can reach all the other nodes by a strong directed path; iii) the digraph is weakly connected and contains a spanning forest, i.e., there exists no node that can reach every other node through a strong directed path; iv) the digraph is not even weakly connected. The last case is the least interesting, as it corresponds to a set of isolated subnetworks, that can be analyzed independently of each other, as one of the previous cases. In the first two cases, according to Theorem 1, system (7) achieves a *global* consensus, whereas in the third case the system forms clusters of consensus with, in general, different consensus values in each cluster, i.e., the system synchronizes only *locally*.

In other words, a global consensus is possible *if and only if* there exists at least one node (the root node of the spanning directed tree of the digraph) that can send its information, directly or indirectly, to all the other nodes. If no such a node exists, a global consensus cannot be reached. However, a *local* consensus is still achievable among all the nodes that are able to influence each other.

Interestingly, the closed form expressions of the synchronized state given by (11) confirms the above statements: The observation y_i of, let us say, sensor i affects the final consensus value if and only if such an information can reach all the other nodes by a strong directed path. As a by-product of this result, we have the following special cases (Corollary 1): If there is only one node that can reach all the others, then the final consensus depends only on the observation taken from that node; on the other extreme, the final consensus contains contributions from *all* the nodes *if and only if* the digraph is SC.

Moreover, if a node contributes to the final value, it does that through a weight that depends on its in/out-degree. The set of weights $\{\gamma_i\}$ in (11) can be interpreted as a measure of the “symmetry/asymmetry” of the communication links in the network: Some of these weights are equal to each other if and only if the subdigraph associated to the corresponding nodes is balanced (or undirected).

The synchronized state, as given in (11), suggests also an interesting interpretation of the consensus formation mechanism of system (7), based on the so called *condensation* digraph.⁸ From (11) in fact, one infers that all the nodes that are SC to each other (usually referred to as nodes of a strongly connected component (SCC)) produce the same effect on the final consensus as an *equivalent* single node that represents the consensus within that SCC. In fact, one may easily check if system in (7) locally or globally synchronizes and which nodes contribute on the consensus, simply reducing the original digraph to its equivalent *condensation* digraph and looking for the existence of a spanning

⁸Please, refer to Appendix A.1 for the definition of condensation digraph and the procedure for reducing a digraph into its condensation digraph.

directed tree in the condensation digraph. The only SCCs of the original digraph that will provide a contribution on the final consensus are the SCCs associated to the root nodes of the condensation digraph.

As an additional remark, the possibility to form clusters of consensus, rather than a global consensus, depends on the (channel) coefficients a_{ij} . If, e.g., these coefficients have the expression as given in Section 2.2, they may be altered by changing the transmit powers P_j . As a consequence, the nodes with the highest transmit power will be the most influential ones. If, for example, we want to set a certain parameter on each node, like, e.g., a decision threshold, we can use the same consensus mechanism used in this paper by assigning, for example, the desired value to, let us say, node i , and select the transmit powers so that node i is the only node that can reach every other node.

3.3 How to get unbiased estimates

The closed form expression of the synchronized state, as given in (11) (or in (15)), is valid for *any* given digraph (not only for undirected graphs as in [8]). Expression (11) shows a dependence of the final consensus on the network topology and propagation parameters, through the coefficients $\{a_{ij}\}$, $\{\gamma_i\}$ and the delays $\{\tau_{ij}\}$. This means that, even if the propagation delays do not affect the convergence of the proposed system, they introduce a bias on the final value, whose amount depends on both the delays and coefficients $\{a_{ij}\}$. The effect of propagation parameters and network topology on the final synchronized state is also contained in the eigenvector γ of the Laplacian \mathbf{L} . This implies that the final consensus resulting from (11) cannot be made to coincide with the desired decision statistics as given by (1), except that in the trivial case where all the delays are equal to zero and the digraph is balanced (and thus strongly connected). However, expression (11) suggests a method to get rid of any bias, as detailed next.

The bias due to the propagation delays can be removed using the following two-step algorithm.⁹ We let the system in (7) to evolve twice: The first time, the system evolves according to (7) and we denote by $\omega^*(\mathbf{y})$ the synchronized state; the second time, we set $g_i(y_i) = 1$ in (7), for all i , and the system is let to evolve again, denoting the final synchronized state by $\omega^*(\mathbf{1})$. From (11), if we take the ratio $\omega^*(\mathbf{y})/\omega^*(\mathbf{1})$, we get

$$\frac{\omega^*(\mathbf{y})}{\omega^*(\mathbf{1})} = \frac{\sum_{i=1}^N \gamma_i c_i g_i(y_i)}{\sum_{i=1}^N \gamma_i c_i}, \quad (18)$$

which coincides with the ideal value achievable in the absence of delays. Thus, this simple double-step algorithm allows us to remove the bias term depending on the delays and on the channel coefficients, without requiring the knowledge or estimate of neither set of parameters.

If the network is strongly connected and balanced, $\gamma_i = 1, \forall i$ and then the compensated consensus coincides with the desired value (1). If the network is unbalanced, the compensated consensus $\omega^*(\mathbf{y})/\omega^*(\mathbf{1})$ does not depend on the (channel) coefficients $\{a_{ij}\}$, but it is still biased, with a bias dependent on γ , i.e., on the network topology. Nevertheless, this residual bias can be eliminated in a decentralized way according to the following iterative algorithm. Let us denote by N_r the number of nodes in the RSCC of the digraph. At the beginning, every node sets $g_i(y_i) = 1$ and $c_i = 1$,

⁹We focus only on the scalar system in (7), because of space limitation.

$i = 1, \dots, N$ and the network is let to evolve. The final consensus value is denoted by $\omega^*(\mathbf{1})$. Then, the network is let to evolve N_r times, according to the following protocol. At step i , with $i = 1, \dots, N_r$, the nodes within the i -th SCC, set $g_i(y_i) = 1$, while all the other nodes set $g_k(y_k) = 0$ for all $k \neq i$. Let us denote by $\omega^*(\mathbf{e}_i)$ the final consensus value, where \mathbf{e}_i is the canonical vector having all zeros, except the i -th component, equal to one. Repeating this procedure for all the SCC's, at the end of the N_r steps, each node is now able to compute the ratio $\omega^*(\mathbf{e}_i)/\omega^*(\mathbf{1})$, which coincides with $\tilde{\gamma}_i := \gamma_i / \sum_k \gamma_k$. Thus, after $N_r + 1$ steps, every node knows its own (normalized) $\tilde{\gamma}_i$. This value is subsequently used to compensate for the network unbalance as follows. The compensation is achieved by simply setting, at each node $c_i = c_i / \tilde{\gamma}_i$. In fact, with this setting, the final ratio $\omega^*(\mathbf{y})/\omega^*(\mathbf{1})$ coincides with the desired unbiased expression given by (1), for all i such that $\gamma_i \neq 0$. If the digraph \mathcal{G} is SC, this procedure corresponds indeed to make the Laplacian matrix $\mathbf{L}(\mathcal{G})$ balanced, so that the left eigenvector of \mathbf{L} associated to the zero eigenvalue be proportional to the vector $\mathbf{1}_N$. This procedure only needs some kind of coordination among the nodes to make them working according to a described scheduling. It is important to remark that, since the eigenvector $\boldsymbol{\gamma}$ does not depend on the observations $\{y_i\}$, the proposed algorithm is required to be performed at the start-up phase of the network and repeated only if the network topology or the channels change through time. In summary, we can eliminate the effect of both delays and channel parameters on the final consensus value, thus achieving the optimal decision statistics as given in (1), with a totally decentralized algorithm, at the price of a slight increase of complexity and the need for some coordination among the nodes.¹⁰

3.4 Asymptotic convergence rate

Theorem 1 generalizes previous results on the convergence speed of classical linear consensus protocols (see, e.g. [8, 23]) to the case in which there are propagation delays. In spite of the presence of delays, the proposed system still converges to the consensus with *exponential* convergence rate, i.e., $\|\dot{\mathbf{x}}(t; \mathbf{y}) - \omega^*(\mathbf{y})\| \leq O(e^{rt})$, where the convergence factor $r < 0$ is defined in Theorem 1. Thus, the convergence speed is dominated by the slowest “mode” of the system. Moreover, as expected, the presence of delays affects the convergence speed, as the roots of characteristic equation $p(s) = 0$ associated to system (7) depend, in general, on both network topology and propagation delays. Unfortunately, a closed form expression of this dependence is not easy to derive, and the roots of $p(s) = 0$ need to be computed numerically.

In the special case of negligible delays instead, we can provide a bound of the convergence factor as a function of the channel parameters through the eigenvalues of the Laplacian matrix \mathbf{L} . Setting $\tau_{ij} = 0$ in (7), the characteristic equation associated to (7) becomes $p(s) = |s\mathbf{I} + \mathbf{L}| = 0$, whose solutions are just the eigenvalues of $-\mathbf{L}$. Under the assumptions of Theorem 1, it follows that

$$r = -\min_i \{\operatorname{Re}\{\lambda_i\} : \lambda_i \in \sigma\{\mathbf{L}\}, \text{ and } \lambda_i \neq 0\}, \quad (19)$$

where $\sigma\{\mathbf{L}\}$ denotes the spectrum of Laplacian matrix \mathbf{L} . The value of r is negative if and only if the digraph \mathcal{G} is QSC [45, Lemma 2]. Using (19), we can obtain bounds on the convergence rate as a

¹⁰In [47], the authors provided a closed form expression of the agreement obtainable with an SC balanced network achieving consensus on the state. However, that expression depends on the channel parameters and on the initial conditions in a way that the bias cannot be eliminated through a distributed procedure.

function of the network topology, as shown next. When the digraph is strongly connected we have¹¹

$$r \leq \kappa \triangleq -\lambda_2 \left(\frac{1}{2}(\mathbf{D}_\gamma \mathbf{L} + \mathbf{L}^T \mathbf{D}_\gamma) \right) < 0, \quad (20)$$

where $\mathbf{D}_\gamma \triangleq \text{diag}(\gamma_1, \dots, \gamma_N)$, γ is the left eigenvector of \mathbf{L} associated to the zero eigenvalue, normalized so that $\|\gamma\|_\infty = 1$, and $\lambda_2(\mathbf{A})$ denotes the second smallest eigenvalue of the symmetric matrix \mathbf{A} . It follows from (20) that system (7) reaches a consensus with rate at least κ . As special cases of (20) we have the following: i) In the case of balanced digraphs, $\mathbf{D}_\gamma = \mathbf{I}$ and thus $\kappa = -\lambda_2 \left(\frac{1}{2}(\mathbf{L} + \mathbf{L}^T) \right)$; ii) If the digraph is undirected (and connected) $\mathbf{D}_\gamma = \mathbf{I}$, $\mathbf{L} = \mathbf{L}^T$ and thus $\kappa = -\lambda_2(\mathbf{L})$, where $\lambda_2(\mathbf{L})$ is also known as the *algebraic connectivity* of the digraph [39]. For the case in which the digraph is QSC, some bounds of r can be found in [44, 46] using the generalization of the classical definition of algebraic connectivity. Moreover, interestingly, the convergence rate of the system, under the assumption of Theorem 1, can be related to the convergence rates of the SCCs of condensation digraph associated to the system. Because of lack of space, we suggest the interested reader to check [45] for more details.

In conclusion, according to the above results, we infer that the convergence rate of the proposed consensus algorithm, at least in the absence of propagation delays, is the same as that of the classical linear protocols achieving consensus on the state.

4 Numerical Results

In this section, we illustrate first some examples of consensus, for different network topologies. Then, we show an application of the proposed technique to an estimation problem, in the presence of random link coefficients. In both examples, the analog system (7) is implemented in discrete time, with sampling step size $T_s = 10^{-3}$.

Example 1: Different forms of consensus for different topologies

In Figure 1, we consider three topologies (top row), namely: (a) a SC digraph, (b) a QSC digraph with three SCCs, and (c) a WC (not QSC) digraph with a spanning forest composed by two trees. For each digraph, we also sketch its decomposition into SCCs (each one enclosed in a circle), corresponding to the nodes of the associated condensation digraph (whose root SCC is denoted by RSCC). In the bottom row of Figure 1, we plot the dynamical evolutions of the state derivatives of system (7) versus time, for the three network topologies, together with the theoretical asymptotic values predicted by (11) (dashed line with arrows). As proved by Theorem 1, the dynamical system in Figure 1a) achieves a global consensus, since the underlying digraph is SC. The network of Figure 1b), instead, is not SC, but the system is still able to globally synchronize, since there is a set of nodes, in the RSCC component, able to reach all the other nodes. The final consensus, in such a case, contains only the contributions of the nodes in the RSCC, since no other node belongs to the root of a spanning directed tree of the condensation digraph. Finally, the system in Figure 1c) cannot reach a global consensus since there is no node that can reach all the others, but it does admit two disjoint clusters, corresponding to the two RSCCs, namely RSCC₁ and RSCC₂. The middle lines of Figure 1c) refer to the nodes of the SCC component, not belonging to either RSCC₁ or RSCC₂, that are affected by the

¹¹The inequality in (20) follows from the fact that $\text{Re}\{\lambda_i\} \geq |\kappa|$ for all nonzero eigenvalues $\{\lambda_i\}$ of \mathbf{L} [45].

consensus achieved in the two RSCC components, but that cannot affect them. Observe that, in all the cases, the state derivatives of the (global or local) clusters converge to the values predicted by the closed form expression given in (11), (12) or (14), depending on the network topology.

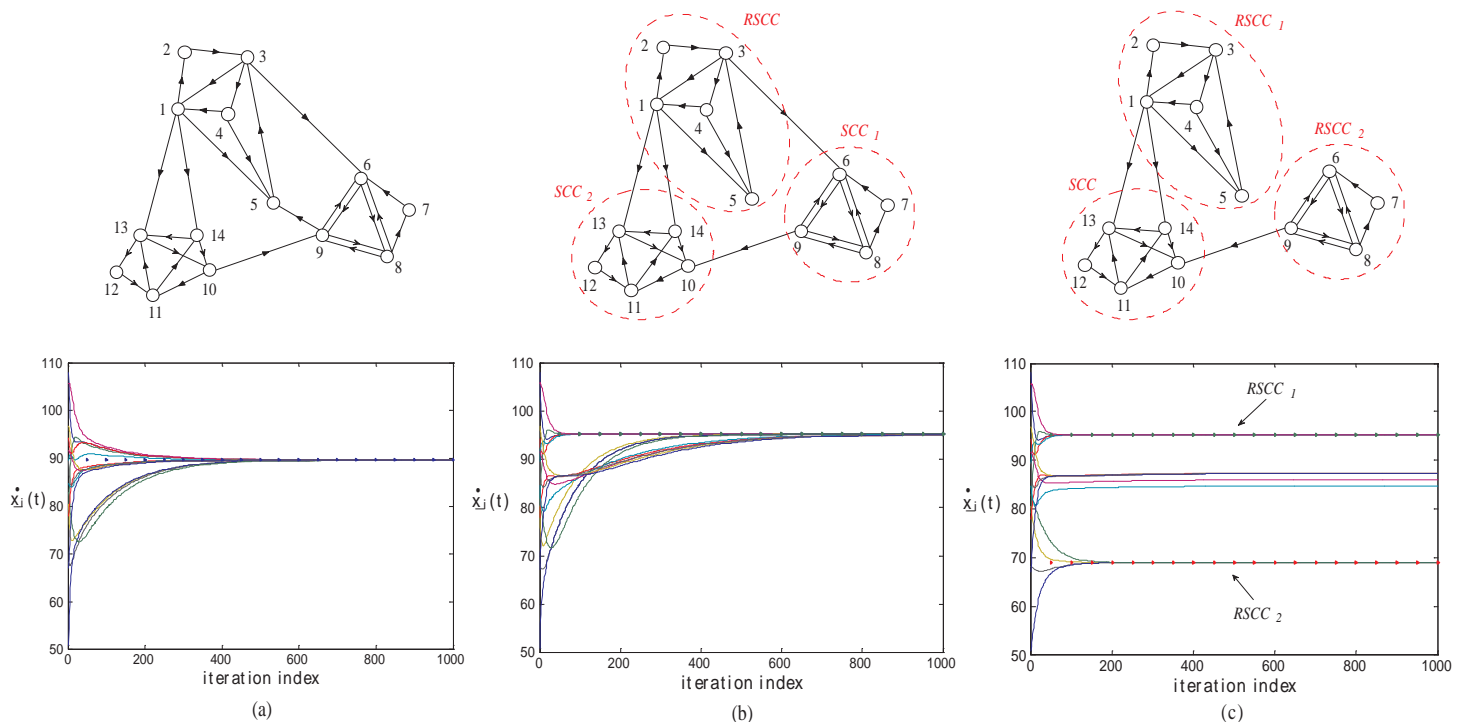


Figure 1: Consensus for three different network topologies: a) SC digraph; b) QSC digraph with three SCCs; c) WC digraph with a two trees forest; $T_s = 10^{-3}$, $\tau = 50T_s$, $K = 30$, $N=14$.

Example 2: Distributed optimal decisions through self-synchronization. The behaviors shown in the previous example refer to a given realization of the topology, with given link coefficients, and of the observations. In this example, we report a global parameter representing the variance obtained in the estimate of a scalar variable. Each sensor observes a variable $y_i = A_i\xi + w_i$, where w_i is additive zero mean Gaussian noise, with variance σ_i^2 . The goal is to estimate ξ . The estimate is performed through the interaction system (7), with functions $g_i(y_i) = y_i/A_i$ and coefficients $c_i = A_i^2/\sigma_i^2$, chosen in order to achieve the globally optimal ML estimate. The network is composed of 40 nodes, randomly spaced over a square of size D . The size of the square occupied by the network is chosen in order to have a maximum delay $\tau = 100T_s$. We set the threshold on the amplitude of the minimum useful signal to zero, so that, at least in principle, each node hears each other node. The corresponding digraph is then SC. To simulate a practical scenario, the channel coefficients a_{ij} are generated as i.i.d. Rayleigh random variables, to accommodate for channel fading. Each variable a_{ij} has a variance depending on the distance d_{ij} between nodes i and j , equal to¹² $\sigma_{ij}^2 = P_j/(1 + d_{ij}^2)$.

In Figure 2, we plot the estimated average state derivative (plus and minus the estimation standard deviation), as a function of the iteration index. Figure 2a) refers to the case in which there is only

¹²We use the attenuation factor $1/(1 + d_{ij}^2)$ instead of $1/d_{ij}^2$ to avoid the undesired event that, for $d_{ij} < 1$ the received power might be greater than the transmitted power.

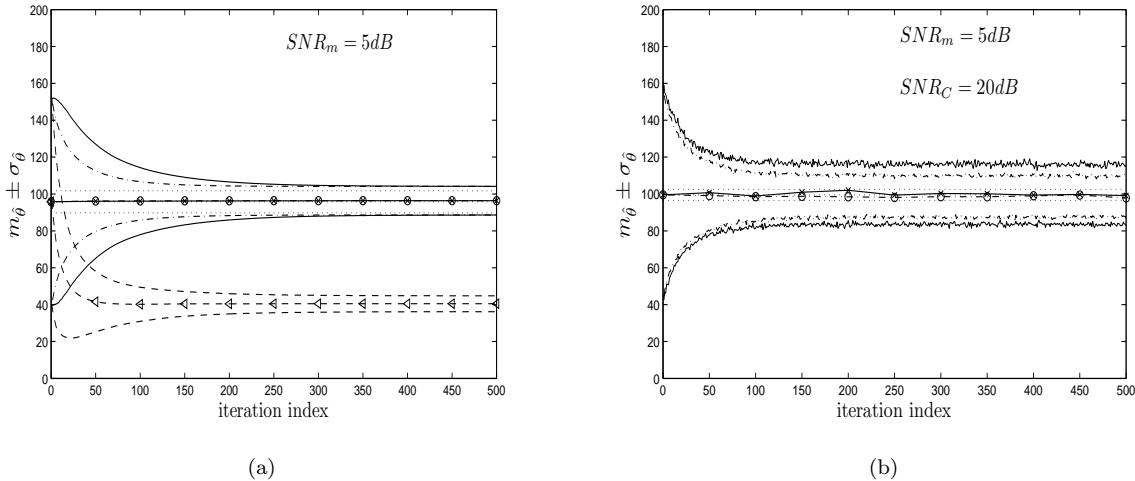


Figure 2: Estimated parameter vs. convergence time in the absence [subplot a)] and in the presence [subplot b)] of receiver noise: a) ML estimate by a centralized system (dotted lines); b) estimate by the proposed method, with no propagation delays, (dashed and dotted lines plus “x” for the average value); c) estimate by the proposed method, with propagation delays (dashed lines plus “<” for the average value); d) estimate by the two-step estimation method leading to (18) (solid lines plus “o” for the average value).

observation noise, but there is no noise in the exchange of information among the nodes. Conversely, Figure 2b) refers to the case where the node interaction is noisy, so that the state evolution of each sensor is described by the state equation $\dot{z}_i(t) = \dot{x}_i(t) + v_i(t)$, with $\dot{x}_i(t)$ given by (7), where $v_i(t)$ is white Gaussian noise; the SNR is $|\xi|^2/\sigma_w^2 = 20$ dB. The averages are taken across all the nodes, for 100 independent realizations of the network, where, in each realization we generated a new topology and a new set of channel coefficients and noise terms. The results refer to the following cases of interest: a) ML estimate achieved with a centralized system, with no communication errors between nodes and fusion center (dotted lines); b) estimate achieved with the proposed method, with no propagation delays, as a benchmark term (dashed and dotted lines plus \times marks for the average value); c) estimate achieved with the proposed method, in the presence of propagation delays (dashed lines plus triangles for the average value); d) estimate achieved with the two-step estimation method leading to (18) (solid lines plus circles for the average value). From Figure 2 we can see that, in the absence of delays, the (decentralized) iterative algorithm based on (7) behaves, asymptotically, as the (centralized) globally optimal ML estimator. In the presence of delays, we observe a clear bias (dashed lines), due to the large delay values, but with a final estimation variance still close to the ML estimator’s. Interestingly, the two-step procedure leading to (18) provides results very close to the optimal ML estimator, with no apparent bias, in spite of the large delays and the random channel fading coefficients. The only price paid with the two-step procedure, besides time, is a slight increase of the variance due to taking the ratio of two noisy consensus values, as evidenced in Fig. 2 b).

5 Concluding remarks

In conclusion, in this paper we have proposed a totally decentralized sensor network scheme capable to reach globally optimal decision tests through local exchange of information among the nodes, in

the presence of asymmetric communication channels and inhomogeneous time-invariant propagation delays. The method is particularly useful for applications where the goal of the network is to take decisions about a common event. Differently from the average consensus protocols available in the literature, our system globally synchronizes for *any* set of (finite) propagation delays *if and only if* the underlying digraph is QSC, with a final synchronized state that is a known function of the sensor measurements. In general, the synchronized state depends on the propagation parameters, such as delays and communication channels. Nevertheless, exploiting our closed form expression for the final consensus values, we have shown how to recover an unbiased estimate, for any set of delays and channel coefficients, without the need to knowing or estimating these coefficients. This desirable result is a distinctive property of the consensus achievable on the state derivative and cannot be obtained using classical consensus/agreement algorithms that reach a consensus on the state variables. If we couple the nice properties mentioned above with the properties reported in [26], where we showed that, in the absence of delays, the consensus protocol proposed in this paper and in [26] is also robust against coupling noise, we have, overall, a good candidate for a distributed sensor network.

As in many engineering problems, the advantages of our scheme come with their own prices. Three issues that deserve further investigations are the following: i) the states grow linearly with time; ii) the coefficients a_{ij} are nonnegative; and iii) a change of topology affects the convergence properties of the proposed scheme. The first issue has an impact on the choice of the radio interface responsible for the exchange of information between the nodes. To avoid the need for transmitting with a high dynamic range, the nodes must transmit a nonlinear, bounded function of the state value. One possibility, as proposed in [38], is to associate the state value to the phase of sinusoidal carrier or to the time shift of a pulse oscillator. The second point requires that the receiver be able to compensate for possible sign inversions. As far as the switching topology is concerned, it would be useful to devise methods aimed at increasing the resilience of our method against topology changes. But it is worth keeping in mind that the previous aspects are only the reverse of the medal of a method capable to achieve a globally optimal decision for any set of delays and for asymmetric channels.

6 Appendix

A Directed Graphs

The interaction among the sensors is properly described by a directed graph. For the reader's convenience, in this section, we briefly review the notation and basic results of graph theory that will be used throughout this paper. For the reader interested in a more in-depth study of this field, we recommend, for example, [39]–[41].

A.1 Basic Definitions

To take explicitly into account the possibility of unidirectional links among the network nodes, we represent the information topology among the nodes by their (weighted) directed graph.

Directed graph. Given N nodes, a (weighted) directed graph (or digraph) \mathcal{G} is defined as $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, where $\mathcal{V} \triangleq \{v_1, \dots, v_N\}$ is the set of nodes (or vertices) and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ is a set of edges (i.e., *ordered*

pairs of the nodes), with the convention that $e_{ij} \triangleq (v_i, v_j) \in \mathcal{E}$ (i.e., v_i and v_j are the head and the tail of the edge e_{ij} , respectively) means that the information flows from v_j to v_i . A digraph is weighted if a positive weight is associate to each edge, according to a proper map $\mathcal{W} : \mathcal{E} \rightarrow \mathbb{R}_+$, such that if $e_{ij} \triangleq (v_i, v_j) \in \mathcal{E}$, then $\mathcal{W}(e_{ij}) = a_{ij} > 0$, otherwise $a_{ij} = 0$. We focus in the following on weighted digraphs where the weights of loops (v_i, v_i) are zero, i.e., $a_{ii} = 0$ for all i . If $(v_i, v_j) \in \mathcal{E} \Leftrightarrow (v_j, v_i) \in \mathcal{E}$ (and $a_{ij} = a_{ji}, \forall i \neq j$), then the graph is said to be (weighted) *undirected*. For any node $v_i \in \mathcal{V}$, we define the information neighbor of v_i as

$$\mathcal{N}_i \triangleq \{j = 1, \dots, N : e_{ij} = (v_i, v_j) \in \mathcal{E}\}. \quad (21)$$

The set \mathcal{N}_i represents the set of indices of the nodes sending data to node i .

The in-degree and out-degree of node $v_i \in \mathcal{V}$ are, respectively, defined as:

$$\deg_{\text{in}}(v_i) \triangleq \sum_{j=1}^N a_{ij}, \quad \text{and} \quad \deg_{\text{out}}(v_i) \triangleq \sum_{j=1}^N a_{ji}. \quad (22)$$

Observe that for undirected graphs, $\deg_{\text{in}}(v_i) = \deg_{\text{out}}(v_i)$.

We may have the following class of digraphs.

Balanced digraph The node v_i of a digraph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ is said to be *balanced* if and only if its in-degree and out-degree coincide, i.e., $\deg_{\text{in}}(v_i) = \deg_{\text{out}}(v_i)$. A digraph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ is called *balanced* if and only if all its nodes are balanced, i.e.,

$$\sum_{j=1}^N a_{ij} = \sum_{j=1}^N a_{ji}, \quad \forall i = 1, \dots, N. \quad (23)$$

Path/cycle A *strong path* (or directed chain) in a digraph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ is a sequence of *distinct* nodes $v_0, \dots, v_q \in \mathcal{V}$ such that $(v_i, v_{i-1}) \in \mathcal{E}, \forall i = 1, \dots, q$. If $v_0 \equiv v_q$, the path is said to be *closed*. A *weak path* is a sequence of *distinct* nodes $v_0, \dots, v_q \in \mathcal{V}$ such that either $(v_{i-1}, v_i) \in \mathcal{E}$ or $(v_i, v_{i-1}) \in \mathcal{E}, \forall i = 1, \dots, q$. A (strong) *cycle* (or circuit) is a closed (strong) path.

Directed tree/forest A digraph with N nodes is a (rooted) *directed tree* if it has $N - 1$ edges and there exists a distinguished node, called the *root* node, which can reach *all* the other nodes by a (unique) strong path. Thus a directed tree cannot have cycles and every node, except the root, has one and only one incoming edge.¹³ A digraph is a (directed) *forest* if it consists of one or more directed trees. A subgraph $\mathcal{G}_s = \{\mathcal{V}_s, \mathcal{E}_s\}$ of a digraph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, with $\mathcal{V}_s \subseteq \mathcal{V}$ and $\mathcal{E}_s \subseteq \mathcal{E} \cap (\mathcal{V}_s \times \mathcal{V}_s)$, is a *spanning* directed tree (or a *spanning* directed forest), if it is a directed tree (or a directed forest) and it has the *same* node set as \mathcal{G} , i.e., $\mathcal{V}_s \equiv \mathcal{V}$. We say that a digraph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ contains a spanning tree (or a spanning forest) if there exists a subgraph of \mathcal{G} that is a spanning directed tree (or a spanning directed forest).

Connectivity. In a digraph there are many degrees of connectedness. In this paper we focus on the following. A digraph is *Strongly Connected* (SC) if, for *every* pair of nodes v_i and v_j , there exists a strong path from v_i to v_j and viceversa. A digraph is *Quasi-Strongly Connected* (QSC) if, for *every*

¹³Observe that some literature (e.g., [20, 21]) defines this concept using the opposite convention for the orientation of the edges.

pair of nodes v_i and v_j , there exists a node r that can reach both v_i and v_j by a *strong* path. A digraph is *weakly* connected (WC) if any pair of distinct nodes can be joined by a *weak* path. A digraph is *disconnected* if it is not weakly connected.

According to the above definitions, it is straightforward to see that strong connectivity implies quasi strong connectivity and that quasi strong connectivity implies weak connectivity, but the converse, in general, does not hold. For undirected graphs instead, the above notions of connectivity are equivalent: An undirected graph is connected if any two distinct nodes can be joined by a path. Moreover, it is easy to check that the quasi strong connectivity of a digraph is equivalent to the existence of at least one spanning directed tree in the graph (see, e.g., [42, p. 133]).

Condensation Digraph When a digraph \mathcal{G} is WC, it may still contain strongly connected subgraphs. A maximal subgraph of \mathcal{G} which is also SC is called *Strongly Connected Component* (SCC) of \mathcal{G} ¹⁴ [40]. Using this concept, any digraph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ can be partitioned into SCCs, let us say $\mathcal{G}_1 \triangleq \{\mathcal{V}_1, \mathcal{E}_1\}$, $\dots, \mathcal{G}_K \triangleq \{\mathcal{V}_K, \mathcal{E}_K\}$, where $\mathcal{V}_j \subseteq \mathcal{V}$ and $\mathcal{E}_j \subseteq \mathcal{E}$ denote the set of nodes and edges lying in the j -th SCC, respectively. Using this structure, one can reduce the original digraph \mathcal{G} to the so called *condensation* digraph $\mathcal{G}^* = \{\mathcal{V}^*, \mathcal{E}^*\}$, by substituting each node set \mathcal{V}_i of each SCC \mathcal{G}_i of \mathcal{G} with a distinct node $v_i^* \in \mathcal{V}^*$ of \mathcal{G}^* , and introducing an edge in \mathcal{G}^* from v_i^* to v_j^* if and only if there exists *some* edges from the i -th SCC \mathcal{G}_i and the j -th SCC \mathcal{G}_j [40, Ch. 3.2]. An SCC that is reduced to the root node of a directed tree of the condensation digraph is called *Root SCC* (RSCC). Observe that, by definition, the condensation digraph has no cycles [40, Lemma 3.2.3]. Building on this property, we have the following.

Lemma 1 *Let $\mathcal{G}^* = \{\mathcal{V}^*, \mathcal{E}^*\}$ be the condensation digraph of \mathcal{G} , composed by K nodes. Then, the nodes of \mathcal{G}^* can always be ordered as $v_1^*, \dots, v_K^* \in \mathcal{V}^*$, so that all the existing edges in \mathcal{G}^* are in the form*

$$(v_i^*, v_j^*) \in \mathcal{E}^*, \quad \text{with} \quad 1 \leq j < i \leq K, \quad (24)$$

where v_1^* has zero in-degree.

The ordering v_1^*, \dots, v_K^* satisfying (24) can be obtained by the following iterative procedure. Starting from v_1^* , remove v_1^* and all its out-coming edges from \mathcal{G}^* . Since the reduced digraph with $K - 1$ nodes has no (strong) cycles by construction, there must exist a node with zero in-degree in it. Let us denote such a node by v_2^* . Then, no edges in the form (v_2^*, v_j^*) , with $j > 2$, can exist in the reduced digraph (and thus in \mathcal{G}^*). This justifies (24) for $i = 2$ and $j = 1, 2$. The rest of (24), for $j > 2$, is obtained by repeating the same procedure on the remaining nodes.

The connectivity properties of a digraph are related to the structure of its *condensation* digraph, as given in the following Lemma (we omit the proof because of space limitations).

Lemma 2 *Let \mathcal{G}^* be the condensation digraph of \mathcal{G} . Then: i) \mathcal{G} is SC if and only if \mathcal{G}^* is composed by a single node; ii) \mathcal{G} is QSC if and only if \mathcal{G}^* contains a spanning directed tree; iii) if \mathcal{G} is WC, then \mathcal{G}^* contains either a spanning directed tree or a (weakly) connected directed forest.*

The concept of condensation digraph is useful to understand the network synchronization behavior, as shown in Section 3. A similar idea was already used in [10, 21] to study leadership problems in coordinated multi-agent systems.

¹⁴Maximal means that there is no larger SC subgraph containing the nodes of the considered component.

A.2 Spectral properties of a Digraph

We recall now some basic relationships between the connectivity properties of the digraph and the spectral properties of the matrices associated to the digraph, since they play a fundamental role in the stability analysis of the system proposed in this paper. In the following, we denote by $\mathbf{1}_N$ and $\mathbf{0}_N$ the N -length column vector of all ones and zeros, respectively.

Given a digraph $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$, we introduce the following matrices associated to \mathcal{G} :

- The $N \times N$ *adjacency matrix* \mathbf{A} is composed of entries $[\mathbf{A}]_{ij} \triangleq a_{ij}$, $i, j = 1, \dots, N$, equal to the weight associated with the edge e_{ij} , if $e_{ij} \in \mathcal{E}$, or equal to zero, otherwise;
- The *degree matrix* $\mathbf{\Delta}$ is a diagonal matrix with diagonal entries $[\mathbf{\Delta}]_{ii} \triangleq \deg_{\text{in}}(v_i)$;
- The (weighted) *Laplacian* \mathbf{L} is defined as

$$[\mathbf{L}]_{ij} \triangleq \begin{cases} \sum_{k \neq i=1}^N a_{ik}, & \text{if } j = i, \\ -a_{ij} & \text{if } j \neq i. \end{cases} \quad (25)$$

Using the adjacency matrix \mathbf{A} and the degree matrix $\mathbf{\Delta}$, the Laplacian can be rewritten in compact form as $\mathbf{L} \triangleq \mathbf{\Delta} - \mathbf{A}$.¹⁵

By definition, the Laplacian matrix \mathbf{L} in (25) has the following properties: i) it is a diagonally dominant matrix [43]; ii) it has zero row sums; and iii) it has nonnegative diagonal elements. From i)-iii), invoking Geršgorin's disk Theorem [43], we have that zero is an eigenvalue of \mathbf{L} corresponding to a right eigenvector in the $\text{Null}\{\mathbf{L}\} \supseteq \text{span}\{\mathbf{1}_N\}$, i.e.,

$$\mathbf{L} \mathbf{1}_N = \mathbf{0}_N, \quad (26)$$

while all the other eigenvalues have positive real part. This also means that $\text{rank}(\mathbf{L}) \leq N - 1$.

Moreover, from (23) and (26), it turns out that *balanced* digraphs can be equivalently characterized in terms of the Laplacian matrix \mathbf{L} : A digraph is balanced if and only if $\mathbf{1}_N$ is also a left eigenvector of \mathbf{L} associated with the zero eigenvalue, i.e.,

$$\mathbf{1}_N^T \mathbf{L} = \mathbf{0}_N^T, \quad (27)$$

or equivalently $\frac{1}{2}(\mathbf{L} + \mathbf{L}^T)$ is positive semidefinite.

The relationship between the connectivity properties of a digraph and the spectral properties of its Laplacian matrix are given by the following.

Lemma 3 *Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ be a digraph with Laplacian matrix \mathbf{L} . The multiplicity of the zero eigenvalue of \mathbf{L} is equal to the minimum number of directed trees contained in a spanning directed forest of \mathcal{G} .*

Corollary 2 *The zero eigenvalue of \mathbf{L} is simple if and only if \mathcal{G} contains a spanning directed tree (or, equivalently, it is QSC).*

¹⁵Observe that the definition of Laplacian matrix as given in (25) coincides with that used in the classical graph theory literature, except for the convention we adopted in the orientation of the edges. This leads to \mathbf{L} expressed in terms of the in-degrees matrix, rather than of the out-degrees matrix. Our choice is motivated by the physical interpretation we gave to the edges' weights, as detailed in Section 2.2.

Lemma 3 comes directly from Theorem 9 and Theorem 10 of [44]. Corollary 2 was independently proved in many papers, such as [9, Corollary 2], [20, Lemma 2]. Observe that, since the strong connectivity of the digraph implies QSC, the results provided in [8, 21] for SC digraphs, can be obtained as special case of Corollary 2. Specifically, we have the following.

Corollary 3 *Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ be a digraph with Laplacian matrix \mathbf{L} . If \mathcal{G} is SC, then \mathbf{L} has a simple zero eigenvalue and a positive left-eigenvector associated to the zero eigenvalue.*

According to Corollary 2, because of (26), the Laplacian of a QSC digraph has an isolated eigenvalue equal to zero, corresponding to a right eigenvector in the span $\{\mathbf{1}_N\}$. Observe that, for *undirected* graphs, Corollary 3 can be stated as follows: rank $(\mathbf{L}) = N - 1$ if and only if \mathcal{G} is connected [39]. For directed graphs, instead, the “only if” part does not hold.

We derive now the structure of the left-eigenvector γ of the Laplacian matrix \mathbf{L} associated to the zero eigenvalue, as a function of the network topology. This result is instrumental to prove the main theorem of this paper. We have the following.

Lemma 4 *Let $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ be a digraph with N nodes and Laplacian matrix \mathbf{L} . Assume that \mathcal{G} is QSC with K SCC's $\mathcal{G}_1 \triangleq \{\mathcal{V}_1, \mathcal{E}_1\}, \dots, \mathcal{G}_K \triangleq \{\mathcal{V}_K, \mathcal{E}_K\}$, with $\mathcal{V}_i \subseteq \mathcal{V}$, $\mathcal{E}_i \subseteq \mathcal{E}$, $|\mathcal{V}_i| = r_i$ and $\sum_i r_i = N$, numbered w.l.o.g. so that \mathcal{G}_1 coincides with the RSCC of \mathcal{G} . Then, the left-eigenvector $\gamma = [\gamma_1, \dots, \gamma_N]^T$ of \mathbf{L} associated to the zero eigenvalue has the following structure*

$$\gamma_i = \begin{cases} > 0, & \text{iff } v_i \in \mathcal{V}_1, \\ = 0, & \text{otherwise.} \end{cases} \quad (28)$$

If \mathcal{G}_1 is also balanced, then $\gamma_{r_1} \triangleq [\gamma_1, \dots, \gamma_{r_1}]^T \in \text{span}\{\mathbf{1}_{r_1}\}$, where $r_1 \triangleq |\mathcal{V}_1|$.

Proof. Since the digraph \mathcal{G} is QSC, the set \mathcal{V}_1 contains either all N or $0 < r_1 < N$ nodes of \mathcal{G} .

In the former case, $\mathcal{V}_1 \equiv \mathcal{V}$ and thus the digraph is SC, by definition. Hence, according to Corollary 3 (cf. Appendix A.2), the Laplacian matrix $\mathbf{L}(\mathcal{G})$ has a simple zero eigenvalue with left-eigenvector $\gamma > \mathbf{0}$. If, in addition, \mathcal{G} is balanced (and thus also SC), then $\gamma \in \text{span}\{\mathbf{1}_N\}$.

We consider now the latter case, i.e., $0 < r_1 < N$. According to Lemma 2, the condensation digraph $\mathcal{G}^* = \{\mathcal{V}^*, \mathcal{E}^*\}$ of \mathcal{G} contains a spanning directed tree composed by K nodes $v_1^*, \dots, v_K^* \in \mathcal{V}^*$ (associated to the K SCCs $\mathcal{G}_1, \dots, \mathcal{G}_K$), with root node v_1^* (associated to the SCC \mathcal{G}_1 of \mathcal{G}). From Lemma 1, we assume, w.l.o.g., that the nodes $v_1^*, \dots, v_K^* \in \mathcal{V}^*$ are ordered according to (24), so that the Laplacian matrix of \mathcal{G}^* be a lower triangular matrix.

Using the relationship between the original digraph \mathcal{G} and its condensation digraph \mathcal{G}^* (cf. Appendix A.1), the Laplacian matrix $\mathbf{L}(\mathcal{G})$ of \mathcal{G} can be written as a lower *block* triangular matrix, i.e.

$$\mathbf{L}(\mathcal{G}) = \begin{pmatrix} \mathbf{L}_1 & \mathbf{0} & \dots & \mathbf{0} \\ * & \mathbf{B}_2 & \ddots & \vdots \\ * & * & \ddots & \mathbf{0} \\ * & \dots & * & \mathbf{B}_K \end{pmatrix}, \quad (29)$$

with $\mathbf{L}_1 = \mathbf{L}(\mathcal{G}_1) \in \mathbb{R}^{r_1 \times r_1}$ denoting the $r_1 \times r_1$ Laplacian matrix associated to the root SCC \mathcal{G}_1 , and $\mathbf{B}_k = \mathbf{L}_k + \mathbf{D}_k \in \mathbb{R}^{r_k \times r_k}$, where $\mathbf{L}_k = \mathbf{L}(\mathcal{G}_k)$ is the $r_k \times r_k$ Laplacian matrix of the k -th SCC \mathcal{G}_k

of \mathcal{G} and \mathbf{D}_k is a nonnegative diagonal matrix whose i -th entry is equal to the sum of the weights associated to the edges outgoing from the nodes in $\mathcal{G}_1, \dots, \mathcal{G}_{k-1}$ and incoming in the i -th node in \mathcal{G}_k .¹⁶ Observe that, since \mathcal{G} is QSC, each \mathbf{D}_k has at least one positive entry; otherwise the SCC \mathcal{G}_k would be decoupled from the other SCCs.

Since each \mathcal{G}_k is SC by definition, we have $\text{rank}(\mathbf{L}_k) = r_k - 1$ (cf. Corollary 3) and $\text{Null}(\mathbf{L}_k) = \text{span}(\mathbf{1}_{r_k})$ (see (26)). Using these properties and the fact that $\mathbf{D}_k \mathbf{1}_{r_k} \neq \mathbf{0}_{r_k}$, we have that the rows (columns) of $\mathbf{L}_k + \mathbf{D}_k$ are linearly independent, or equivalently,

$$\text{rank}(\mathbf{B}_k) = r_k, \quad \forall k = 2, \dots, K. \quad (30)$$

Using (29) and (30), we derive now the structure of the left-eigenvector $\boldsymbol{\gamma}$ of $\mathbf{L}(\mathcal{G})$ in (29). Partitioning $\boldsymbol{\gamma} = [\boldsymbol{\gamma}_{r_1}^T, \boldsymbol{\gamma}_{r_2}^T, \dots, \boldsymbol{\gamma}_{r_K}^T]^T$ according to (29), with $\boldsymbol{\gamma}_{r_k} \in \mathbb{R}^{r_k}$, we have

$$\boldsymbol{\gamma}_{r_1}^T \mathbf{L}_1 = \mathbf{0}_{r_1}^T \quad \text{and} \quad \boldsymbol{\gamma}_{r_k}^T \mathbf{B}_k = \mathbf{0}_{r_k}^T, \quad k = 2, \dots, K, \quad (31)$$

which, using Corollary 3 and (30), provides

$$\boldsymbol{\gamma}_{r_1} > \mathbf{0}_{r_1} \quad \text{and} \quad \boldsymbol{\gamma}_{r_k} = \mathbf{0}_{r_k}, \quad k = 2, \dots, K, \quad (32)$$

respectively. If, in addition, \mathcal{G}_1 is balanced, then $\boldsymbol{\gamma}_{r_1} \in \text{span}\{\mathbf{1}_{r_1}\}$. ■

B Preliminary Results on Linear Functional Differential Equations

In this section, we introduce some basic definitions on linear functional differential equations and prove some intermediate results that will be extensively used in the proof of Theorem 1, as given in Appendix C.

To formally introduce the concept of functional differential equations, we need the following notations: let \mathbb{R}^N be an N -dimensional linear vector space over the real numbers with vector norm $\|\cdot\|$; denoting by τ the maximum time delay of the system, let $\mathcal{C} \triangleq \mathcal{C}([-\tau, 0], \mathbb{R}^N)$ [or $\mathcal{C}^1 \triangleq \mathcal{C}^1([-\tau, 0], \mathbb{R}^N)$] be the Banach space of continuous (or continuously differentiable) functions mapping the interval $[-\tau, 0]$ to \mathbb{R}^N with the topology of uniform convergence, i.e., for any $\boldsymbol{\phi} \in \mathcal{C}$ (or $\boldsymbol{\phi} \in \mathcal{C}^1$) the norm of $\boldsymbol{\phi}$ is defined as $\|\boldsymbol{\phi}\|_s = \sup_{-\tau \leq \vartheta \leq 0} \|\boldsymbol{\phi}(\vartheta)\|$.

Since in this paper we considered only linear coupling among the differential equations as given in (7), we focus on *Linear homogeneous* Delayed Differential Equations (LDDEs) with a finite number of heterogeneous non-commensurate delays¹⁷ [48]-[51]:

$$\dot{x}_i(t) = k_i \sum_{j \in \mathcal{N}_i} a_{ij} (x_j(t - \tau_{ij}) - x_i(t)), \quad i = 1, \dots, N, \quad t > t_0, \quad (33)$$

where k_i is any arbitrary positive constant. Equation (33) means that the derivative of the state variables \mathbf{x} at time t depends on t and $\mathbf{x}(\vartheta)$, for $\vartheta \in [t - \tau, t]$. Hence, to uniquely specify the evolution

¹⁶We assume, w.l.o.g., that in each SCC \mathcal{G}_k the nodes are numbered from 1 to r_k .

¹⁷Observe that the LDDE (33) falls in the more general class of LDDE's studied in [48]-[50]. In fact, it is straightforward to check that LDDE (33) can be rewritten in the canonical form of [48, Ch. 6, Eq. (6.3.2)], [50, Ch. 3, Eq. (3.1)].

of the state beyond time t_0 , it is required to specify the initial state variables $\mathbf{x}(t)$ in a time interval of length τ , from $t_0 - \tau$ to t_0 , i.e.,

$$x_i(t_0 + \vartheta) = \phi_i(\vartheta), \quad \vartheta \in [-\tau, 0], \quad i = 1, \dots, N. \quad (34)$$

Here after, the initial conditions $\phi = \{\phi_i\}_i$ are assumed to be taken in the set \mathcal{C}_β^1 of (continuously differentiable) functions that are bounded in the norm¹⁸ $|\phi|_s = \sup_{-\tau \leq \vartheta \leq 0} \|\phi(\vartheta)\|_\infty$, i.e.,

$$\mathcal{C}_\beta^1 \triangleq \left\{ \phi \in \mathcal{C}^1 : |\phi|_s = \sup_{-\tau \leq \vartheta \leq 0} \|\phi(\vartheta)\|_\infty \leq \beta < +\infty \right\}. \quad (35)$$

Given $\phi \in \mathcal{C}_\beta^1$ and t_0 , let $\mathbf{x}[t_0, \phi](t)$ denote the function \mathbf{x} at time t , with initial value ϕ at time t_0 , i.e., $\mathbf{x}(t_0 + \vartheta) = \phi(\vartheta)$, with $\vartheta \in [-\tau, 0]$. For the sake of notation, we define $\mathbf{x}[\phi](t) \triangleq \mathbf{x}[0, \phi](t)$. A function $\mathbf{x}[t_0, \phi](t)$ is said to be a solution to equation (33) on $[t_0 - \tau, t_0 + t_1)$ with initial value ϕ at time t_0 , if there exist $t_1 \geq 0$ such that: i) $\mathbf{x} \in \mathcal{C}([t_0 - \tau, t_0 + t_1], \mathbb{R}^n)$; ii) $\mathbf{x}[t_0, \phi](t)$ satisfies (33), $\forall t \in [t_0, t_0 + t_1]$; and iii) $\mathbf{x}(t_0 + \vartheta) = \phi(\vartheta)$, with $\vartheta \in [-\tau, 0]$. It follows from [50, Theorem 1.2] that such a solution to equation (33) exists and is unique. We focus now on two concepts related to the trajectories of (33), namely boundedness and stability.

Definition 2 *Given system (33), a solution $\mathbf{x}[t_0, \phi](t)$ is bounded if there exists a $\beta = \beta(t_0, \phi)$ such that $\|\mathbf{x}[t_0, \phi](t)\| < \beta(t_0, \phi)$ for $t \geq t_0 - \tau$. The solutions are uniformly bounded if, for any $\alpha > 0$, there exists a $\beta = \beta(\alpha) > 0$ such that for all $t_0 \in \mathbb{R}$, $\phi \in \mathcal{C}$ and $|\phi|_s < \alpha$ we have $\|\mathbf{x}[t_0, \phi](t)\| < \beta$ for all $t \geq t_0$.*

As far as the stability notion is concerned, it is not at all different from its counterpart for unretarded systems, except for the different assumptions on the initial conditions. The interested reader may refer, e.g., to [50, 51] for an in-depth treatment of this topic. We focus here on effective methods on proving the stability of LDDEs. Since system (33) is linear, the stability analysis can be carried out either in the time-domain [50, 51] or in the frequency-domain [48, 49], as for classical ordinary differential equations (see, e.g., [52]). In this paper we focus on the latter approach. The same conclusions can also be obtained using the time-domain analysis, based on Lyapunov-Krasovskii functional [53]. Before stating the major result, we need first the following intermediate definitions.

Let $\mathbb{C}_+ = \{s \in \mathbb{C} : \text{Re}\{s\} > 0\}$, $\mathbb{C}_- = \{s \in \mathbb{C} : \text{Re}\{s\} < 0\}$, and $\overline{\mathbb{C}}_+$ be the closure of \mathbb{C}_+ , i.e., $\overline{\mathbb{C}}_+ = \{s \in \mathbb{C} : \text{Re}\{s\} \geq 0\}$. Denoting by $\mathcal{H}^{n \times m}$ the set of $n \times m$ matrices whose entries are analytic¹⁹ and bounded functions in \mathbb{C}_+ , let us introduce the $N \times N$ diagonal degree matrix $\mathbf{\Delta} \geq \mathbf{0}_{N \times N}$ and the complex matrix $\mathbf{H}(s) \in \mathbb{C}^{N \times N}$, defined respectively as

$$\mathbf{\Delta} \triangleq \text{diag}(k_1 \text{deg}_{\text{in}}(v_1), \dots, k_N \text{deg}_{\text{in}}(v_N)) \quad \text{and} \quad [\mathbf{H}(s)]_{ij} \triangleq \begin{cases} 0, & \text{if } i = j, \\ k_i a_{ij} e^{-s\tau_{ij}}, & \text{if } i \neq j, \end{cases} \quad (36)$$

with $\text{deg}_{\text{in}}(v_i)$ given in (22) (see Appendix A). Observe that $\mathbf{H}(s) \in \mathcal{H}^{N \times N}$. We can now provide the main result of this section, stated in the following lemma.

¹⁸We used, without loss of generality, as vector norm $\|\cdot\|$ in \mathbb{R}^N the infinity norm $\|\cdot\|_\infty$, defined as $\|\mathbf{x}\|_\infty \triangleq \max_i |x_i|$.

¹⁹A complex function is said to be analytic (or holomorphic) on a region $\mathcal{D} \subseteq \mathbb{C}$ if it is complex differentiable at every point in \mathcal{D} , i.e., for any $z_0 \in \mathcal{D}$ the function satisfies the Cauchy-Riemann equations and has continuous first partial derivatives in the neighborhood of z_0 (see, e.g., [54, Theorem 11.2]).

Lemma 5 *Given system (33), assume that the following conditions are satisfied:*

- b1.** *The initial value functions $\phi \in \mathcal{C}_\beta^1$, and the solutions $\mathbf{x}[\phi](t)$ are bounded;*
- b2.** *The characteristic equation associated to (33)*

$$p(s) \triangleq \det(s\mathbf{I} + \mathbf{\Delta} - \mathbf{H}(s)) = 0, \quad (37)$$

with $\mathbf{\Delta}$ and $\mathbf{H}(s)$ defined in (36), has all roots $\{s_r\}_r \in \mathbb{C}_-$, with at most one simple root at $s = 0$.²⁰

Then, system (33) is marginally stable, i.e., $\forall \phi \in \mathcal{C}_\beta^1$ and $\text{Re}\{s_1\} < c < 0$, there exist t_1 and α , with $t_0 < t_1 < +\infty$ and $0 < \alpha < +\infty$, independent of ϕ , and a vector \mathbf{x}^∞ , with $\|\mathbf{x}^\infty\| < +\infty$, such that

$$\|\mathbf{x}[\phi](t) - \mathbf{x}^\infty\| \leq \alpha |\phi|_s e^{ct}, \quad \forall t > t_1. \quad (38)$$

Proof. Under assumption **b1**), according to [48, Theorem 6.5], the stability properties of system (33) are fully determined by the roots of the characteristic equation associated to (33), as detailed next.²¹ Denoting by $\mathcal{M} = \{\text{Re}\{s_r\} : p(s_r) = 0\}$ the set of real parts of the characteristic roots $\{s_r\}_r$, it follows from [48, Theorem 6.7] that, $\forall \phi \in \mathcal{C}_\beta^1$ and $c \notin \mathcal{M}$, there exist t_1 and α , with $0 < t_1 < +\infty$ and $0 < \alpha < +\infty$ independent on ϕ , such that

$$\left\| \mathbf{x}[\phi](t) - \lim_{l \rightarrow +\infty} \sum_{s_r \in C_l: \text{Re}\{s_r\} > c} \mathbf{p}_r(t) e^{s_r t} \right\| \leq \alpha |\phi|_s e^{ct}, \quad \forall t > t_1, \quad (39)$$

where each \mathbf{p}_r is a (vectorial) polynomial of degree less than the multiplicity of s_r , C_l denotes a contour in the complex plane of radius increasing with l and centered around $s = 0$ (see [48, Sec. 4.1] for more details on how such contours C_l need to be chosen), and the sum in (39) is taken over all characteristic roots s_r within the contour C_l and to the right of the line $\text{Re}\{s_r\} = c$. Observe that, since the number of such roots is finite²² [50, Theorem 1.5] (see also [48, Ch. 6.8 and Ch. 12]), the limit in (39) is always well-defined.

Using (39), we prove now that, under **b2**), system (33) is marginally stable. Invoking the property that, for any given $\gamma \in \mathbb{R}$, the number of characteristic roots s_r with $\text{Re}\{s_r\} > \gamma$ is finite, one can always choose the constant $c \notin \mathcal{M}$ in (39) so that $\text{Re}\{s_1\} < c < \text{Re}\{s_0\} = 0$, which leads to

$$\|\mathbf{x}[\phi](t) - \mathbf{p}_0\| \leq \alpha |\phi|_s e^{ct}, \quad \forall t > t_1, \quad (40)$$

where we have explicitly used the assumption that the (possible) root $s_0 = 0$ is simple and that there are no roots with positive real part. It follows from [48, Corollary 6.2] that this condition is also sufficient to guarantee that \mathbf{p}_0 is bounded, i.e., $\|\mathbf{p}_0\| < +\infty$. Setting in (40) $\mathbf{x}^\infty = \mathbf{p}_0$, we obtain (38), which completes the proof. ■

²⁰We assume, w.l.o.g., that the roots $\{s_r\}$ are arranged in nonincreasing order with respect to the real part, i.e., $0 = \text{Re}\{s_0\} > \text{Re}\{s_1\} \geq \text{Re}\{s_2\} \geq \dots$

²¹Observe that assumption **b1**) is only sufficient for the existence of the Laplace transform of the solutions to (33).

²²The reader who is familiar with Linear Retarded Functional Differential Equations (LRFDE) may observe that this result comes directly from the fact that the characteristic equation (37) does not have neutral roots [48, Ch. 12].

Remark 1. It follows from Lemma 5 that, under assumptions **b1)**-**b2)**, all the solutions to (33) asymptotically converge to the constant vector \mathbf{x}^∞ . Moreover, equation (38) provides an estimate of the convergence rate of the system: as for systems without delays, the convergence speed of (33) is exponential²³ with rate arbitrarily close to $\text{Re}\{s_1\}$ that, in general, depends on network topology and delays.

Remark 2. Lemma 5 generalizes results of [49, 50], where the authors provided alternative conditions for the *asymptotic* stability of a Linear Retarded Differential Equation (LRDE). Interestingly, Lemma 5 contains some of the conditions of [49, 50] as a special case: System (33) is asymptotically stable if all the characteristic roots of (37) have negative real part [see (38)]. This conclusion is the same as that for linear unretarded systems (see, e.g., [52]).

Moreover, Lemma 5 can be easily generalized to include the cases in which one is interested in “oscillatory” behaviors of system (33). It is straightforward to see that *bounded* oscillations arise if assumption **b2)** is replaced by the following condition: All the roots of characteristic equation (37) have negative real part and the roots with zero real part are simple.

C Proof of Theorem 1

In the following, for the sake of notation simplicity, we drop the dependence of the state function from the observation, as this dependence does not play any role in our proof.

C.1 Sufficiency

We prove that, under **a1)**-**a3)**, the quasi strong connectivity of digraph \mathcal{G} associated to the network in (7) is a sufficient condition for the system (7) to synchronize and that the synchronized state is given by (11). To this end, we organize the proof according to the following two steps.

We first show that, under **a1)**-**a2)** and the quasi-strong connectivity of \mathcal{G} , the set of LRFDEs (7) admits a solution in the form

$$x_i^*(t) = \alpha t + x_{i,0}^*, \quad i = 1, \dots, N, \quad (41)$$

if $\alpha = \omega^*$, where ω^* is defined in (11) and $\{x_{i,0}^*\}$ are constants that depend in general on the system parameters and the initial conditions. This guarantees the existence of the desired synchronized state (cf. Definition 1). Then, invoking results of Appendix B, we prove that, under **a1)**-**a3)** and the quasi-strong connectivity of \mathcal{G} , such a synchronized state is also globally asymptotically stable (according to Definition 1).

C.1.1 Existence of a synchronized state

Let us assume that conditions **a1)**-**a2)** are satisfied and that $\mathcal{G} = \{\mathcal{V}, \mathcal{E}\}$ associated to (7) is QSC. The synchronized state in the form (41) is a solution to (7) if and only if it satisfies equations (7) (cf. Appendix B), i.e., if and only if there exist α and $\{x_{i,0}^*\}$ such that the following system of linear equations is feasible:

$$\frac{c_i \Delta \omega_i(\alpha)}{K} + \sum_{j \in \mathcal{N}_i} a_{ij} (x_{j,0}^* - x_{i,0}^*) = 0, \quad \forall i = 1, \dots, N, \quad (42)$$

²³We say that $\mathbf{x}(t)$ converges exponentially toward \mathbf{x}^∞ with rate $r < 0$ if $\|\mathbf{x}(t) - \mathbf{x}^\infty\| \leq O(e^{rt})$.

where

$$\Delta\omega_i(\alpha) \triangleq g_i(y_i) - \alpha \left(1 + \frac{K}{c_i} \sum_{j \in \mathcal{N}_i} a_{ij} \tau_{ij} \right). \quad (43)$$

Introducing the weighted Laplacian $\mathbf{L} = \mathbf{L}(\mathcal{G})$ associated to digraph \mathcal{G} (cf. Section A), the system in (42) can be equivalently rewritten as

$$\mathbf{L}\mathbf{x}_0^* = \frac{1}{K}\mathbf{D}_c\Delta\boldsymbol{\omega}(\alpha), \quad (44)$$

where $\mathbf{x}_0^* \triangleq [x_{1,0}^*, \dots, x_{N,0}^*]^T$, $\mathbf{D}_c \triangleq \text{diag}(c_1, \dots, c_N)$, and $\Delta\boldsymbol{\omega}(\alpha) \triangleq [\Delta\omega_1(\alpha), \dots, \Delta\omega_N(\alpha)]^T$, with $\Delta\omega_i(\alpha)$ defined in (43). Observe that, under **a1)**-**a2)** and the quasi-strong connectivity of \mathcal{G} , the graph Laplacian \mathbf{L} has the following properties (cf. Corollary 2):

$$\text{rank}(\mathbf{L}) = N - 1, \quad \mathcal{N}(\mathbf{L}) = \text{span}\{\mathbf{1}_N\}, \quad \text{and} \quad \mathcal{N}(\mathbf{L}^T) = \text{span}\{\boldsymbol{\gamma}\}, \quad (45)$$

where $\mathcal{N}(\mathbf{L})$ denotes the (right) null-space of \mathbf{L} , and $\boldsymbol{\gamma}$ is a left eigenvector of \mathbf{L} corresponding to the (simple) zero eigenvalue of \mathbf{L} , i.e., $\boldsymbol{\gamma}^T\mathbf{L} = \mathbf{0}^T$.

Assume now that α is fixed. It follows from (45) that, for any *given* α , system (44) admits a solution if and only if $\mathbf{D}_c\Delta\boldsymbol{\omega}(\alpha) \in \text{span}\{\mathbf{L}\}$. Because of (45), we have

$$\mathbf{D}_c\Delta\boldsymbol{\omega}(\alpha) \in \text{span}\{\mathbf{L}\} \Leftrightarrow \boldsymbol{\gamma}^T\mathbf{D}_c\Delta\boldsymbol{\omega}(\alpha) = 0. \quad (46)$$

It is easy to check that the value of α that satisfies (46) is $\alpha = \omega^*$, with ω^* defined in (11). Hence, if $\alpha = \omega^*$, the synchronized state in the form (41) is a solution to (7), for any given set of $\{\tau_{ij}\}$, $\{g_i\}$, $\{c_i\}$, $\{a_{ij}\}$ and $K \neq 0$. The structure of the left eigenvector $\boldsymbol{\gamma}$ associated to the zero eigenvalue of \mathbf{L} as given in (11) comes directly from Lemma 4.

Setting $\alpha = \omega^*$, system (44) admits ∞^1 solutions, given by

$$\mathbf{x}_0^* = \frac{1}{K}\mathbf{L}^\sharp\mathbf{D}_c\Delta\boldsymbol{\omega}(\omega^*) + \text{span}\{\mathbf{1}_N\} \triangleq \bar{\mathbf{x}}_0 + \text{span}\{\mathbf{1}_N\}, \quad (47)$$

where

$$\bar{\mathbf{x}}_0 \triangleq \frac{1}{K}\mathbf{L}^\sharp\mathbf{D}_c\Delta\boldsymbol{\omega}(\omega^*), \quad (48)$$

$\Delta\omega_i(\omega^*)$ is obtained by (43) setting $\alpha = \omega^*$ and \mathbf{L}_A^\sharp is the generalized inverse of the Laplacian \mathbf{L} [56].

C.1.2 Global Asymptotic Stability of the Synchronized State

To prove the global asymptotic stability of the synchronized state of system (7), whose existence has been proved in Appendix C.1.1, we use the following intermediate result (see Appendix B for the definitions used in the lemma).

Lemma 6 ([55, Theorem 2.2]) *Let $\mathbf{H}(s) \in \mathcal{H}^{N \times N}$ and $\rho(\mathbf{H}(s))$ denote the spectral radius of $\mathbf{H}(s)$. Then, $\rho(\mathbf{H}(s))$ is a subharmonic²⁴ bounded (above) function on $\overline{\mathbb{C}}_+$.*

²⁴See, e.g., [54, Ch. 12], [55], for the definition of subharmonic function.

We first rewrite system (7) in a more convenient form, as detailed next. Consider the following change of variables

$$\Psi_i(t) \triangleq x_i(t) - (\omega^* t + \bar{x}_{i,0}), \quad i = 1, \dots, N, \quad (49)$$

where ω^* and $\{\bar{x}_{i,0}\}$ are defined in (11) and (48), respectively, so that the original system (7) can be equivalently rewritten in terms of the new variables $\{\Psi_i(t)\}_i$ as

$$\dot{\Psi}_i(t) = \Delta\omega_i(\omega^*) + \frac{K}{c_i} \sum_{j \in \mathcal{N}_i} a_{ij} (\Psi_j(t - \tau_{ij}) - \Psi_i(t) + \bar{x}_{j,0} - \bar{x}_{i,0}), \quad i = 1, \dots, N, \quad t \geq 0, \quad (50)$$

$$\Psi_i(\vartheta) = \phi_i(\vartheta) \triangleq \tilde{\phi}_i(\vartheta) - \omega^* \vartheta - \bar{x}_{i,0}, \quad \vartheta \in [-\tau, 0],$$

where $\tilde{\phi}$ are the initial value functions of the original system (7). Using (48) [see also (42), with $\alpha = \omega^*$], system (50) becomes

$$\dot{\Psi}_i(t) = k_i \sum_{j \in \mathcal{N}_i} a_{ij} (\Psi_j(t - \tau_{ij}) - \Psi_i(t)), \quad i = 1, \dots, N, \quad t \geq 0, \quad (51)$$

$$\Psi_i(\vartheta) = \phi_i(\vartheta), \quad \vartheta \in [-\tau, 0],$$

where, for the sake of convenience, we defined $k_i \triangleq K/c_i > 0$, for $i = 1, \dots, N$.

It follows from (51) that the synchronized state of system (7), as given in (41), is globally asymptotically stable (according to Definition 1) if system in (51) is marginally stable (cf. Appendix B). To prove the marginal stability of system (51), it is sufficient to show that system (51) satisfies Lemma 5. To this end, we organize the rest of the proof in the following two steps.

Step 1. We show that, under **a1)**, **a3)**, all the solutions $\Psi[\phi](t)$ to system (51) are uniformly bounded, as required by assumption **b1)** of Lemma 5;

Step 2. We prove that, under **a1)-a3)** and the quasi-strong connectivity of the digraph, the characteristic equation associated to system (51) has all the roots in \mathbb{C}_- and a simple root in $s = 0$, which satisfies assumption **b2)** of Lemma 5.

Step 1. Given any arbitrary $\beta < +\infty$, assumptions **a1)**, **a3)** are sufficient to guarantee that all the trajectories of (51) are uniformly bounded (cf. Definition 2), as shown next. Since $\phi \in \mathcal{C}_\beta^1$, we have (see (35))

$$|\Psi_i(\vartheta)| \leq \beta, \quad \forall i = 1, \dots, N, \quad \vartheta \in [-\tau, 0]. \quad (52)$$

Condition (52) is sufficient for $\{\Psi_i[\phi](t)\}$ to be uniformly bounded for all $t > 0$. In fact, assume that $\{\Psi_i[\phi](t)\}$ are not bounded. Then, according to Definition 2 of Appendix B, there must exist some $\bar{t} > 0$ and a set $\mathcal{J} \subseteq \{1, \dots, N\}$ such that²⁵

$$\|\Psi(t)\|_\infty \leq \beta, \quad \forall t < \bar{t}, \quad (53)$$

and

$$|\Psi_j(\bar{t})| = \beta \quad \text{and} \quad \dot{\Psi}_j(\bar{t}) = \begin{cases} < 0, & \text{if } \Psi_j(\bar{t}) = -\beta, \\ > 0, & \text{if } \Psi_j(\bar{t}) = \beta, \end{cases} \quad j \in \mathcal{J}. \quad (54)$$

This would imply that, for some $t > \bar{t}$, $|\Psi_j(\bar{t})| > \beta$.

Given (53)-(54) and $j \in \mathcal{J}$, we have two possibilities, namely:

²⁵For the sake of notation, we omit in the following the dependence of $\Psi[\phi](t)$ on ϕ .

i) $\Psi_j(\bar{t}) = \beta$ and $\dot{\Psi}_j(\bar{t}) > 0$. Since $|\Psi_i(\bar{t} - \tau_{ji})| < \beta, \forall i = 1, \dots, N$ (see (53)) and $k_j > 0$, from (51) we have

$$\dot{\Psi}_j(\bar{t}) = k_j \sum_{i \in \mathcal{N}_j} a_{ji} (\Psi_i(\bar{t} - \tau_{ji}) - \Psi_j(\bar{t})) = k_j \sum_{i \in \mathcal{N}_j} a_{ji} (\Psi_i(\bar{t} - \tau_{ji}) - \beta) \leq 0, \quad (55)$$

where in the last inequality, we used (52). Expression (55) contradicts the assumption $\dot{\Psi}_j(\bar{t}) > 0$.

ii) $\Psi_j(\bar{t}) = -\beta$ and $\dot{\Psi}_j(\bar{t}) < 0$. However, since

$$\dot{\Psi}_j(\bar{t}) = k_j \sum_{i \in \mathcal{N}_j} a_{ji} (\Psi_i(\bar{t} - \tau_{ji}) - \Psi_j(\bar{t})) = k_j \sum_{i \in \mathcal{N}_j} a_{ji} (\Psi_i(\bar{t} - \tau_{ji}) + \beta) \geq 0, \quad (56)$$

a contradiction with $\dot{\Psi}_j(\bar{t}) < 0$ results.

Thus, any solution $\Psi[\phi](t)$ to (51) is bounded and remains in \mathcal{C}_β^1 for any $t > 0$.

Step 2. We study now the characteristic equation (37) associated to system (51), assuming that **a1)-a3)** are satisfied and that digraph \mathcal{G} is QSC. First of all, observe that, since $\mathbf{\Delta} - \mathbf{H}(0) = K\mathbf{D}_c\mathbf{L}$, we have

$$p(0) = \det(\mathbf{\Delta} - \mathbf{H}(0)) = K \det(\mathbf{D}_c) \det(\mathbf{L}) = 0, \quad (57)$$

where $\mathbf{\Delta}$ and $\mathbf{H}(s)$ are defined in (36), and the last equality in (57) is due to the properties of Laplacian matrix $\mathbf{L} = \mathbf{L}(\mathcal{G})$ of digraph \mathcal{G} (see (26)). It follows from (57) that $p(s)$ has a root in $s = 0$, corresponding to the zero eigenvalue of the Laplacian \mathbf{L} (recall that $K \det(\mathbf{D}_c) \neq 0$). Since the digraph is assumed to be QSC, according to Corollary 2, such a root is simple.

Thus, to complete the proof, we need to show that $p(s)$ does not have any solution in $\overline{\mathcal{C}}_+ \setminus \{0\}$, i.e.,

$$\det(s\mathbf{I} + \mathbf{\Delta} - \mathbf{H}(s)) \neq 0, \quad \forall s \in \overline{\mathcal{C}}_+ \setminus \{0\}. \quad (58)$$

Since $s\mathbf{I} + \mathbf{\Delta}$ is nonsingular in $\overline{\mathcal{C}}_+ \setminus \{0\}$ [recall that, under **a1)**, $\mathbf{\Delta} \geq \mathbf{0}$, with at least one positive diagonal entry], (58) is equivalent to

$$\det(\mathbf{I} - (s\mathbf{I} + \mathbf{\Delta})^{-1} \mathbf{H}(s)) \neq 0, \quad \forall s \in \overline{\mathcal{C}}_+ \setminus \{0\}, \quad (59)$$

which leads to the following sufficient condition for (58):

$$\rho(s) \triangleq \rho((s\mathbf{I} + \mathbf{\Delta})^{-1} \mathbf{H}(s)) < 1, \quad \forall s \in \overline{\mathcal{C}}_+ \setminus \{0\}. \quad (60)$$

Since $(s\mathbf{I} + \mathbf{\Delta})^{-1} \in \mathcal{H}^{N \times N}$ and $\mathbf{H}(s) \in \mathcal{H}^{N \times N}$ (cf. Appendix B), it follows from Lemma 6 that the spectral radius $\rho(s)$ in (60) is a subharmonic function on $\overline{\mathcal{C}}_+$. As a direct consequence, we have, among all, that $\rho(s)$ is a continuous bounded function on $\overline{\mathcal{C}}_+^{26}$ and satisfies the *maximum modulus principle* (see, e.g., [54, Ch. 12]): $\rho(s)$ achieves its *global* maximum only on the boundary of $\overline{\mathcal{C}}_+$.²⁷ Since $\rho(s)$ is strictly proper in $\overline{\mathcal{C}}_+$, i.e., $\rho(s) \rightarrow 0$ as $|s| \rightarrow +\infty$ while keeping $s \in \overline{\mathcal{C}}_+$, it follows that

$$\sup_{s \in \mathbb{C}_+} \rho(s) < \sup_{s \in \overline{\mathcal{C}}_+} \rho(s) \leq \sup_{\omega \in \mathbb{R}} \rho(j\omega). \quad (61)$$

²⁶Observe that $\rho(s)$ is well-defined in $s = 0$, and $\rho(0) = 1$.

²⁷According to the maximum modulus theorem [54, Theorem 10.24], the only possibility for $\rho(s)$ to reach its global maximum also on the interior of $\overline{\mathcal{C}}_+$ is that $\rho(s)$ be constant over all $\overline{\mathcal{C}}_+$, which is not the case.

Using (61), we infer that condition (60) is satisfied if

$$\rho(j\omega) = \rho\left((j\omega\mathbf{I} + \mathbf{\Delta})^{-1} \mathbf{H}(j\omega)\right) < 1, \quad \forall \omega \in \mathbb{R} \setminus \{0\}. \quad (62)$$

For any matrix norm $\|\cdot\|$, since $\rho(\mathbf{A}) \leq \|\mathbf{A}\| \forall \mathbf{A} \in \mathbb{C}^{N \times M}$ [43, Theorem 5.6.9], using the *maximum row sum matrix norm* $\|\cdot\|_\infty$ defined as [43, Definition 5.6.5]

$$\|\mathbf{A}\|_\infty \triangleq \max_{r=1, \dots, N} \sum_{q=1}^M \left| [\mathbf{A}]_{rq} \right|, \quad (63)$$

we have

$$\begin{aligned} \rho(j\omega) &\leq \left\| (j\omega\mathbf{I} + \mathbf{\Delta})^{-1} \mathbf{H}(j\omega) \right\|_\infty = \max_{r=1, \dots, N} \sum_{q \neq r} \left| \frac{k_r a_{rq}}{j\omega + k_r \deg_{\text{in}}(v_r)} e^{-j\omega\tau_{rq}} \right| \\ &= \max_{r=1, \dots, N} \left| \frac{k_r \deg_{\text{in}}(v_r)}{j\omega + k_r \deg_{\text{in}}(v_r)} \right| \leq 1, \end{aligned} \quad (64)$$

where in the last inequality the equality is reached if and only if $\omega = 0$. It follows from (64) that $\rho(j\omega) < 1$ for all $\omega \neq 0$, which guarantees that condition (60) is satisfied.

This proves that assumption **b2)** of Lemma 5 holds true. Hence, all the trajectories $\Psi[\phi](t) \rightarrow \Psi^\infty$ as $t \rightarrow +\infty$, with exponential rate arbitrarily close to $r \triangleq \{\min_i \text{Re}\{s_i\} : p(s_i) = 0 \text{ and } s_i \neq 0\}$, where $p(s)$ is defined in (37) and Ψ^∞ satisfies the linear system of equations $\mathbf{L}\Psi^\infty = \mathbf{0}$, whose solution is $\Psi^\infty \in \text{span}\{\mathbf{1}_N\}$ (Corollary 2). In other words, system (51) exponentially reaches the consensus on the state. \blacksquare

C.2 Necessity

We prove the necessity of the condition by showing that, if the digraph \mathcal{G} of (7) is not QSC, different clusters of nodes synchronize on different values. This local synchronization is in contrast with the definition of (global) synchronization, as given in Definition 1. Hence, if the overall system has to synchronize, the digraph associated to the system must be QSC.

Assume that the digraph \mathcal{G} associated to (7) is not QSC, but WC with K SCCs and, let us say, $r \leq K$ (distinct) RSCCs. Then, according to Lemma 2, the condensation digraph $\mathcal{G}^* = \{\mathcal{V}^*, \mathcal{E}^*\}$ contains a spanning directed forest with r distinct roots (associated to the r RSCCs of the K SCCs of \mathcal{G}). Ordering the nodes $v_1^*, \dots, v_K^* \in \mathcal{V}^*$ according to Lemma 1, and exploring the relationship between \mathcal{G}^* and \mathcal{G} (cf. Appendix C.1.1), one can write the Laplacian matrix $\mathbf{L} = \mathbf{L}(\mathcal{G})$ as an r -reducible matrix [43], i.e.,

$$\mathbf{L}(\mathcal{G}) = \begin{pmatrix} \mathbf{L}_1 & \mathbf{0} & \cdots & \cdots & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & \ddots & \cdots & \cdots & \vdots \\ \mathbf{0} & \mathbf{0} & \mathbf{L}_r & \ddots & \cdots & \vdots \\ * & * & * & \mathbf{B}_{r+1} & \ddots & \vdots \\ * & * & * & * & \ddots & \mathbf{0} \\ * & * & * & * & * & \mathbf{B}_K \end{pmatrix}, \quad (65)$$

where the first r diagonal blocks are the Laplacian matrices of the r RSCCs of \mathcal{G} , and the $\{\mathbf{B}_k\}$, with $k > r$, are the nonsingular matrices associated to the remaining SCCs. Each of these matrices can be

written as the linear combination of the Laplacian matrix of the corresponding SCC and a nonnegative diagonal matrix with at least one positive diagonal entry. The structure of \mathbf{L} given in (65) shows that the RSCCs associated to the first r diagonal blocks are totally decoupled from each other. Hence, (at least) the state derivatives of the nodes in each of these r RSCCs reach a common value (since the corresponding subdigraphs are SC by construction) that, in general, is different for any of the SCCs. This is sufficient for the overall system not to reach a global synchronization.

References

- [1] E. Eisenberg, D. Gale, "Consensus of Subjective Probabilities: The Pari-mutuel Method," *The Annals of Mathematical Statistics*, pp. 165–168, Mar. 1959.
- [2] M. H. DeGroot, "Reaching a Consensus," *Journal of the American Statistical Association*, pp. 118–121, Mar. 1974.
- [3] J. N. Tsitsiklis, D. P. Bertsekas, M. Athans, "Distributed Asynchronous Deterministic and Stochastic Gradient Optimization Algorithms," *IEEE Trans. on Automatic Control*, pp. 803–812, Sep. 1986.
- [4] J. N. Tsitsiklis, "Problems in Decentralized Decisions Making and Computation," *Ph.D. Dissertation*, MIT, Cambridge, MA, Nov. 1982. Available at <http://web.mit.edu/jnt/www/Papers/PhD-84-jnt.pdf>.
- [5] D. P Bertsekas and J.N. Tsitsiklis, *Parallel and Distributed Computation: Numerical Methods*, Athena Scientific, 1989.
- [6] V. Borkar, P. Varaiya, "Asymptotic Agreement in Distributed Estimation," *IEEE Trans. on Automatic Control*, pp. 650–655, 1982.
- [7] R. Olfati-Saber and R. M. Murray, "Consensus Protocols for Networks of Dynamic Agents," in Proc. of the *2003 American Control Conference*, June 2003.
- [8] R. Olfati-Saber, and R.M. Murray, "Consensus Problems in Networks of Agents with Switching Topology and Time-Delays," *IEEE Trans. on Automatic Control*, vol. 49, pp. 1520-1533, Sep., 2004.
- [9] W. Ren, R. W. Beard, and T. W. McLain, "Coordination Variables and Consensus Building in Multiple Vehicle Systems," in Proc. of the *Block Island Workshop on Cooperative Control, Springer-Verlag Series*, vol. 309, p. 171-188, 2005.
- [10] C. W. Wu, "Agreement and Consensus Problems in Groups of Autonomous Agents with Linear Dynamics," in Proc. of *IEEE International Symposium on Circuits and Systems (ISCAS 2005)*, pp. 292 - 295, 23-26 May 2005.
- [11] L. Xiao, S. Boyd, and S. Lall, "A scheme for Robust Distributed Sensor Fusion based on Average Consensus," in Proc. of the *Fourth International Symposium on Information Processing in Sensor Networks, IPSN 2005*, 15 April 2005, pp. 63–70.
- [12] J. Cortés, "Finite-time Convergent Gradient Flows with Applications to Network Consensus," *Automatica*, vol. 42, no. 11, pp. 1993–2000, November 2006.
- [13] R. Olfati-Saber, E. Franco, E. Frazzoli, and J. S. Shamma, "Belief Consensus and Distributed Hypothesis Testing in Sensor Networks," *Lecture Notes in Control and Information Sciences*, vol 331, pp. 169 - 182, July 2006.
- [14] D. Bauso, L. Giarre, and R. Pesenti, "Nonlinear Protocols for Optimal Distributed Consensus in Networks of Dynamic Agents," *Systems & Control Letters*, vol 55, no. 11, pp. 918-928, July 2006.
- [15] J. Cortés, "Distributed Algorithms for Reaching Consensus on General Functions," submitted to *Automatica*, October 2006. Available at http://www.soe.ucsc.edu/~jcortes/research/data/2006_Co-auto.pdf.
- [16] R. Olfati-Saber and J. Shamma, "Consensus Filters for Sensor Networks and Distributed Sensor Fusion," in Proc. of the *Joint 44th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC'05)*, Seville, Spain, December 2005.

- [17] R. Olfati-Saber, "Distributed Kalman Filter with Embedded Consensus Filters," in Proc. of the *Joint 44th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC'05)*, Seville, Spain, December 2005.
- [18] A. Jadbabaie, J. Lin, A. S. Morse, "Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbor Rules," *IEEE Trans. on Automatic Control*, pp. 988–1001, June 2003.
- [19] L. Moreau, "Stability of Multiagent Systems with Time-dependent Communication Links," *IEEE Transactions on Automatic Control*, vol. 50, no. 2, pp. 169–182, February 2005.
- [20] Z. Lin, B. Francis, and M. Maggiore, "Necessary and Sufficient Graphical Conditions for Formation Control of Unicycles," in *IEEE Trans. on Automatic Control*, pp. 121–127, 2005.
- [21] Z. Jin and R. M. Murray, "Consensus Controllability for Coordinated Multiple Vehicle Control, in Proc. of *Sixth Int. Conf. on Cooperative Control and Optimization*, Feb. 1–3, 2006, Gainesville, Florida.
- [22] V. D. Blondel, J. M. Hendrickx, A. Olshevsky, and J. N. Tsitsiklis, "Convergence in Multiagent Coordination, Consensus, and Flocking," in Proc. of the *Joint 44th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC'05)*, Seville, Spain, December 2005.
- [23] R. Olfati-Saber, J. A. Fax, R. M. Murray, "Consensus and Cooperation in Networked Multi-agent Systems," in *Proc. of the IEEE*, vol. 95, no. 1, pp. 215–233, Jan. 2007.
- [24] W. Ren, R. W. Beard, and E. M. Atkins, "Information Consensus in Multivehicle Cooperative Control: Collective Group Behavior Through Local Interaction," *IEEE Control Systems Magazine*, vol. 27, no. 2, pp. 71–82, April 2007.
- [25] S. Barbarossa, "Self-organizing Sensor Networks with Information Propagation based on Mutual Coupling of Dynamic Systems," in *Proc. of IWWAN 2005*, London, UK, 2005.
- [26] S. Barbarossa, and G. Scutari, "Decentralized Maximum Likelihood Estimation for Sensor Networks Composed of Self-synchronizing Locally Coupled Oscillators," to appear on *IEEE Transactions on Signal Processing*, July 2007. Available at <http://arxiv.org/abs/cs.DC/0612042>.
- [27] I. Schizas, A. Ribeiro, and G. B. Giannakis "Distributed Estimation with Ad Hoc Wireless Sensor Networks," in Proc. of *European Signal. Process. Conf. (EUSIPCO 06)*, Florence, Italy, September 4–8, 2006.
- [28] M.G. Earl, and S.H. Strogatz, "Synchronization in Oscillator Networks with Delayed Coupling: A Stability Criterion," *Physical Rev. E*, Vol. 67, pp. 1–4, 2003.
- [29] A. Papachristodoulou and A. Jadbabaie, "Synchronization in Oscillator Networks: Switching Topologies and Presence of Nonhomogeneous delays," in Proc of the *43th IEEE Conference on Decision and Control and European Control Conference (ECC-CDC 2005)*, Seville, Spain, December 2005.
- [30] D.J. Lee and M.W. Spong, "Agreement With Non-uniform Information Delays," in Proc. of the *American Control Conference (ACC 2006)*, Minneapolis, MN, June 14–16, 2006.
- [31] L. Wang and F. Xiao, "A New Approach to Consensus Problem for Discrete-Time Multiagent Systems with Time-Delays," in Proc. of the *American Control Conference (ACC 2006)*, Minneapolis, MN, June 14–16, 2006.
- [32] G. Scutari, S. Barbarossa, and L. Pescosolido, "Optimal Decentralized Estimation Through Self-synchronizing Networks in the Presence of Propagation Delays," in Proc. of *IEEE Workshop on Signal Processing Advances in Wireless Communications, (SPAWC-2006)*, Cannes, France, July 2006.
- [33] A. Giridhar, and P.R. Kumar, "Computing and Communications Functions," *IEEE Trans. on Information Theory*, pp. 388–404, March 2005.
- [34] S. M. Kay, *Fundamentals of Statistical Signal Processing*, Prentice Hall Signal Proc. Series, Upper Saddle River, NJ, 1998.
- [35] L. Pescosolido, S. Barbarossa and G. Scutari, "Decentralized Detection and Localization Through Sensor Networks Designed As a Population of Self-Synchronizing Oscillators," in Proc. of the *IEEE International Conference on Acoustics, Speech and Signal Processing, (ICASSP 2006)*, Vol. 4, pp. 981–984, Toulouse, May 15–19, 2006.
- [36] A.-L. Barabasi and R. Albert, "Emergence of Scaling in Random Networks," *Science*, vol. 286, pp. 509–512, 1999.

- [37] S. Barbarossa, G. Scutari, A. Swami, "Achieving Consensus in Self-organizing Wireless Sensor Networks: The impact of Network Topology on Energy Consumption," in Proc. of the *IEEE International Conference on Acoustics, Speech and Signal Processing, (ICASSP 2007)*, Honolulu, Hawaii, April 15-20, 2007.
- [38] S. Barbarossa and G. Scutari, "Bio-inspired Sensor Network Design: Distributed Decision Through Self-synchronization," *IEEE Signal Processing Magazine*, vol. 24, no. 3, May pp. 2635, 2007.
- [39] M. Fiedler, "Algebraic Connectivity of Graphs", *Journal of Czechoslovak Math.*, vol. 23, pp. 298–305, 1973.
- [40] R. A. Brualdi and H. J. Ryser, *Combinatorial Matrix Theory*, Cambridge University Press, 1991.
- [41] C. Godsil, and G. Royle, *Algebraic Graph Theory*, Springer, New York, 2001.
- [42] C. Berge and A. Ghouila-Houri, *Programming, Games and Transportation Networks*, John Wiley & Sons, New York, 1965.
- [43] R. Horn, and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, 1985.
- [44] C. W. Wu, "On Bounds of Extremal Eigenvalues of Irreducible and m-reducible Matrices," *Linear Algebra and its Applications*, vol. 402, pp. 29-45, June 2005.
- [45] C. W. Wu, "On Rayleigh–Ritz Ratios of a Generalized Laplacian Matrix of Directed Graphs," *Linear Algebra and its Applications*, vol. 402, pp. 207227, February 2005.
- [46] C. W. Wu, "Synchronization in an Array of Chaotic Systems Coupled via a Directed Graph," in Proc. of the *IEEE Int. Symposium on Circuits and System (ISCASS 2005)*, Kobe, Japan, May 23-26, 2005.
- [47] V. Chellaboina, W. M. Haddad, Q. Hui, and J. Ramakrishnan, "On System State Equipartitioning and Semistability in Network Dynamical Systems with Arbitrary Time-Delays," in Proc of the *44th IEEE Conference on Decision and Control (CDC 2006)*, San Diego, CA, USA, December 13-15, 2006.
- [48] R. Bellman and K.L. Cooke, *Differential-Difference Equations*, New York Academic Press, 1963.
- [49] G. Stepan, *Retarded Dynamical Systems: Stability and Characteristic Functions*, Pitman Research Notes in Mathematics Series, Longman Scientific & Technical, 1989.
- [50] K. Gu, V.L. Kharitonov, J. Chen, *Stability of Tyme-Delay Systems*, Control Engineering Series, Birkhauser, 2002.
- [51] J. K. Hale, S.M.V. Lunel, *Introduction to Functional Differential Equations*, Springer-Verlag, 1991.
- [52] T. Kailath, *Linear Systems*, Prentice Hall, 1979.
- [53] J. R. Haddock and J. Terjeki, "Liapunov-Razumikhin Functions and an Invariance Principle for Functional Differential Equations," *Jour. of Differential Equations*, vol. 48, pp. 95-122, 1983.
- [54] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, International Student Ed., 1970.
- [55] S. Boyd and S. A. Desoer, "Subharmonic Functions and Performance Bounds on Linear Time-Invariant Feedback Systems," in *IMA Jour. of Mathematical Control & Information*, Vol. 2, pp. 153-170, 1985.
- [56] S.L. Campbell and C.D. Meyer, *Generalized Inverses of Linear Transformations*, Dover Publications, 1991.