A Theoretical Study of Different Leader Roles in Networks

Wei Wang
Nonlinear Systems Laboratory, MIT
Cambridge, Massachusetts, 02139
Email: wangwei@mit.edu

Jean-Jacques E. Slotine
Nonlinear Systems Laboratory, MIT
Cambridge, Massachusetts, 02139
Email: jjs@mit.edu

Abstract— We study synchronization conditions for distributed dynamic networks with different types of leaders. The role of a "power" leader specifying a desired global state trajectory through local interactions has long been recognized and modeled. This paper introduces the complementary notion of a "knowledge" leader holding information on the target dynamics, which is propagated to the entire network through local adaptation mechanisms. Different types of leaders can co-exist in the same network. For instance, in a network of locally connected oscillators, the power leader may set the global phase while the knowledge leader may set the global frequency and the global amplitude. Knowledge-based leader-followers networks have many analogs in biology, e.g., in evolutionary processes and disease propagation.

I. INTRODUCTION

Recent results in the theoretical study of synchronization and group cooperation [10], [13], [22], [5], [24], [30], [32], [28], [33] have greatly helped understand distributed networks in the natural world and emulate them in artificial systems. In these networks, each element only gets local information from a set of neighbors but the whole system exhibits a collective behavior. Examples of such networked systems pervade nature at every scale, including for instance neural networks, pacemaker cells, flashing fireflies, chirping crickets, and the aggregate motions of bird flocks, fish schools, animal herds and bee swarms. For diffusion-coupled networks with arbitrary size and general structure, explicit conditions on the coupling strengths can be derived for synchronization to occur, based on network connectivity and uncoupled element dynamics [28], [33].

In a network composed of peers, the phase of the collective behavior is hard to predict, since it depends on the initial conditions of all the coupled elements. To let the whole network converge to a specific trajectory, a "leader" can be added [10], [13]. Here the leader is an element whose dynamics is independent from and thus followed by all the others. Such leader-followers network occurs in natural aggregate motions, with the leader specifying "where to go." We shall refer to this kind of leader as the *power leader*. A synchronization condition for a dynamic network with a power leader was derived in [28], [33] and will be briefly reviewed here (section 2).

The main goal of this paper is to introduce a different type of leader, which we shall refer to as a *knowledge leader* (section 3). In this case, the network members' dynamics can all be different. The knowledge leader is the one whose dynamics properties are fixed (or changes comparatively slowly), with the followers obtaining dynamic knowledge from the leader through local adaptation mechanisms. In this sense, a knowledge leader can be understood as the one who indicates "how to go." Such knowledge leaders may exist in many natural processes. For instance, in evolutionary biology [21], [23], the adaptive model we describe could represent genotype-phenotype mappings. Similar mechanisms occur in infectious-disease dynamics [16]. Knowledge leaders may also exist in animal aggregate motion as a junior or injured member with limited capacities. Using Lyapunov analysis, we shall derive conditions of synchronization and also for dynamics convergence for networks with knowledge leaders.

Both types of leaders may co-exist (section 4), and be located anywhere in the network. In a circuit of electronic oscillators, the power leader may be a local clock setting global phase while the knowledge leader sets global frequency or amplitude, for instance. Both types of leaders can be virtual (as e.g. in [13] in the case of power leaders) and may be used for instance to coordinate behaviors in groups of robots of different types.

II. POWER LEADER

Consider the dynamics of a coupled network containing one power leader and n power followers

$$\dot{\mathbf{x}}_0 = \mathbf{f}(\mathbf{x}_0, t) \tag{1}$$

$$\dot{\mathbf{x}}_i = \mathbf{f}(\mathbf{x}_i, t) + \sum_{j \in \mathcal{N}_i} \mathbf{K}_{ji} (\mathbf{x}_j - \mathbf{x}_i) + \gamma_i \mathbf{K}_{0i} (\mathbf{x}_0 - \mathbf{x}_i)$$

Here vector $\mathbf{x}_0 \in \mathbb{R}^m$ is the state of the leader whose dynamics is independent, and $\mathbf{x}_i$ the state of the ith follower, $i = 1, \ldots, n$. The vector function $\mathbf{f}$ represents the uncoupled dynamics, which is assumed to be identical for each element. For notational simplicity, the coupling forces are set to be diffusive, where all coupling gains are symmetric positive definite, and the couplings between the followers are bi-directional with $\mathbf{K}_{ji} = \mathbf{K}_{ij}$ if both $i, j \neq 0$. $\mathcal{N}_i$ denotes the set of peer-neighbors of element i, which for instance could be defined as the set of the followers within a certain distance around element i. γ_i is equal to either 0 or 1, representing the connection from the leader to the followers. In our model, the

network connectivity can be very general. Thus $\mathcal{N}_i$ and γ_i can be defined arbitrarily. An example is illustrated in Figure 1(a).
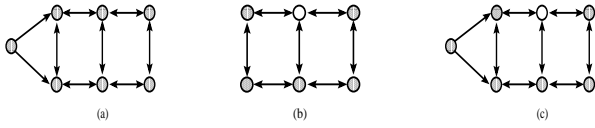

Fig. 1. The graphs illustrate networked systems with (a). a power leader (the left-most node); (b). a knowledge leader (the hollow node); (c). both leaders. The arrows indicate the directions of the couplings.

Since the dynamics of $\mathbf{x}_0$ is independent, it can be considered as an external input to the rest of the network. Therefore it can be seen as an undirected graph with n nodes containing only the followers. We further assume that it has τ inner links.

In the theorem given below, λ represents an eigenvalue and subscript s the symmetric part of a matrix; notation $\mathbf{I}^n_{\gamma_i \mathbf{K}_{0i}}$ denotes an $n \times n$ block diagonal matrix with the ith diagonal entry as $\gamma_i \mathbf{K}_{0i}$; $\mathbf{L_K}$ is the weighted Laplacian matrix [7] and

$$\mathbf{L_K} \;=\; \mathbf{D}\, \mathbf{I}^\tau_{\mathbf{K}_{ij}}\, \mathbf{D}^T$$

where the $n \times \tau$ block matrix $\mathbf{D}$ is a generalized incidence matrix by replacing each number 1 or -1 in the incidence matrix [7] with identity matrix $\mathbf{I} \in \mathbb{R}^{m \times m}$ or $-\mathbf{I}$. Note that the incidence matrix here is obtained by assigning an arbitrary orientation to the undirected graph which contains all the followers. $\mathbf{I}^\tau_{\mathbf{K}_{ij}}$ is a $\tau \times \tau$ block diagonal matrix with the kth diagonal entry $\mathbf{K}_{ij}$ corresponding to the weight of the kth link which connects the nodes i and j.

Theorem 1: The states of all the followers will converge exponentially to the state of the leader if

$$\lambda_{min}(\mathbf{L_K} + \mathbf{I}^n_{\gamma_i \mathbf{K}_{0i}}) > \max_{i=1}^{n} \lambda_{max}(\frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}_i, t))_s \quad \text{uniformly.}$$

The proof of Theorem 1 (Theorem 7 in [33]) is based on Contraction Theory [14], [15], the details of which can be found in [28], [33].

Remarks

• Theorem 1 can be extended to study networks with unidirectional couplings between the followers, with positive semi-definite couplings, or with switching structures [28], [33]. Moreover, the leader does not have to be single, but it can be a group of leading elements, for instance, a synchronized leaders' group. In addition, in some cases it may receive feedback from the followers as well. This is the case in synchronization propagation, where the node density is unevenly distributed through the network. Since the synchronization rate depends on network connectivity, a high-density region will synchronize very quickly despite disturbances from other parts of the network. The inputs from these leaders then facilitate synchronization in low-density regions, where the elements may not be able to synchronize by themselves. A simple simulation was given in [28], and [34] observed a similar phenomenon by setting different interior connection weights inside different subgroups. Note that the leaders group here is

very similar to the concept of *core group* in infectious disease dynamics [16], which is a group of the most active individuals. A small change in the core group will make a big difference in whether or not an epidemic can occur in the whole population.

• Synchronization can be made to propagate from the center outward in a more active way, for instance, through diffusion of a chemical produced by leaders or high-level elements and having the ability to expand the communication channels it passes through, i.e., to increase the gains through diffusion. Such a mechanism represents a hierarchical combination with gain dynamics. By extending the state, the analysis tools provided here can apply more generally to combinations where the gain dynamics are coupled to each other (with arbitrary connectivity) and to the $\mathbf{x}_i$.

• Different leaders $\mathbf{x}_0^j$ of arbitrary dynamics can define different *primitives* which can be combined. Contraction of the follower dynamics ($i = 1, \ldots, n$)

$$\dot{\mathbf{x}}_i = \mathbf{f}(\mathbf{x}_i, t) + \sum_{j \in \mathcal{N}_i} \mathbf{K}_{ji}\,(\mathbf{x}_j - \mathbf{x}_i) + \sum_j \alpha_j(t)\, \gamma_i^j\, \mathbf{K}_{0i}^j\,(\mathbf{x}_0^j - \mathbf{x}_i)$$

is preserved if $\sum_j \alpha_j(t) \geq 1, \; \forall t \geq 0$.

• Besides its dubious moral implications, Theorem 1 also means that it is easy to detract a group from its nominal behavior by introducing a "covert" element, with possible applications to group control games, ethology, and animal and plant mimicry.

• Having a leader in a moving formation may yield other advantages, such as energy saving in aerodynamics [2], [26].

• In the central nervous system, the recently recognized pervasiveness of electrical synapses facilitates such bilateral, diffusion-like mechanisms.

III. KNOWLEDGE LEADER

A knowledge-based leader-followers network is composed of elements with non-identical but adaptive dynamics. A knowledge leader may be located in any position inside a network as we illustrated in Figure 1(b). Its dynamics is fixed or slowly changing, while those of the followers are learned from the leader through adaptation. If we consider the power leader as the one which tells the rest of the network "where to go," a knowledge leader indicates "how to go." Synchronization or group agreement can still be achieved in such a network with only local interactions.

Consider a coupled network containing n elements without a power leader

$$\dot{\mathbf{x}}_i = \mathbf{f}(\mathbf{x}_i, \mathbf{a}_i, t) + \sum_{j \in \mathcal{N}_i} \mathbf{K}_{ji}(\mathbf{x}_j - \mathbf{x}_i) \quad i = 1, \ldots, n \quad (2)$$

where the connectivity can be general. Assume that the uncoupled dynamics $\mathbf{f}(\mathbf{x}_i, \mathbf{a}_i, t)$ is continuous, smooth, and identical to each element except the value of a parameter set $\mathbf{a}_i$, which is different for knowledge followers but has a fixed value $\mathbf{a}$ for all the knowledge leaders. Denote Ω as the set of the followers, whose adaptation laws are based on local interactions

$$\dot{\mathbf{a}}_i \;=\; \mathbf{P}_i \mathbf{W}^T(\mathbf{x}_i, t) \sum_{j \in \mathcal{N}_i} \mathbf{K}_{ji}\,(\mathbf{x}_j - \mathbf{x}_i) \quad \forall\, i \in \Omega \quad (3)$$

where $\mathbf{P}_i > 0$ is constant and symmetric, and $\mathbf{W}(\mathbf{x}_i, t)$ is defined as

$$\mathbf{f}(\mathbf{x}_i, \mathbf{a}_i, t) = \mathbf{f}(\mathbf{x}_i, \mathbf{a}, t) + \mathbf{W}(\mathbf{x}_i, t)\tilde{\mathbf{a}}_i$$

with estimation error $\tilde{\mathbf{a}}_i = \mathbf{a}_i - \mathbf{a}$.

To prove convergence, we first define $\mathbf{L}_{\mathbf{K}\Lambda}$ as a symmetric matrix with

$$\mathbf{L}_{\mathbf{K}\Lambda} = \mathbf{D} \, (\mathbf{I}^\tau_{\mathbf{K}_{ij}} \mathbf{I}^\tau_{\Lambda_{ij}})_s \, \mathbf{D}^T = \mathbf{D} \, \mathbf{I}^\tau_{(\mathbf{K}\Lambda)_{ijs}} \, \mathbf{D}^T \quad (4)$$

Here $\mathbf{I}^\tau_{\Lambda_{ij}}$ is a $\tau \times \tau$ block diagonal matrix with the kth diagonal entry

$$\Lambda_{ij} = \int_0^1 \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}_j + \chi(\mathbf{x}_i - \mathbf{x}_j), \mathbf{a}, t) \, d\chi$$

corresponding to the kth link which has been assigned an orientation by the incidence matrix $\mathbf{D}$. $\mathbf{I}^\tau_{(\mathbf{K}\Lambda)_{ijs}}$ is defined in a similar manner with $(\mathbf{K}\Lambda)_{ijs}$ the symmetric part of $\mathbf{K}_{ij}\Lambda_{ij}$.

We then define a Lyapunov-like function

$$V = \frac{1}{2} \left(\mathbf{x}^T \mathbf{L}_{\mathbf{K}} \mathbf{x} + \sum_{i \in \Omega} \tilde{\mathbf{a}}_i^T \mathbf{P}_i^{-1} \tilde{\mathbf{a}}_i \right)$$

where $\mathbf{x}^T = [\mathbf{x}_1^T, \mathbf{x}_2^T, \ldots, \mathbf{x}_n^T]$, and

$$\dot{\mathbf{x}} = \begin{bmatrix} \mathbf{f}(\mathbf{x}_1, \mathbf{a}_1, t) \\ \ldots \\ \mathbf{f}(\mathbf{x}_n, \mathbf{a}_n, t) \end{bmatrix} - \mathbf{L}_{\mathbf{K}} \mathbf{x}$$

$$= \begin{bmatrix} \mathbf{f}(\mathbf{x}_1, \mathbf{a}, t) \\ \ldots \\ \mathbf{f}(\mathbf{x}_n, \mathbf{a}, t) \end{bmatrix} - \mathbf{L}_{\mathbf{K}} \mathbf{x} + \begin{bmatrix} \mathbf{W}(\mathbf{x}_1, t)(\mathbf{a}_1 - \mathbf{a}) \\ \ldots \\ \mathbf{W}(\mathbf{x}_n, t)(\mathbf{a}_n - \mathbf{a}) \end{bmatrix}$$

so that

$$\dot{V} = \mathbf{x}^T \mathbf{L}_{\mathbf{K}} \dot{\mathbf{x}} + \sum_{i \in \Omega} \tilde{\mathbf{a}}_i^T \mathbf{P}_i^{-1} \dot{\mathbf{a}}_i$$

$$= \mathbf{x}^T \mathbf{L}_{\mathbf{K}} \left(\begin{bmatrix} \mathbf{f}(\mathbf{x}_1, \mathbf{a}, t) \\ \ldots \\ \mathbf{f}(\mathbf{x}_n, \mathbf{a}, t) \end{bmatrix} - \mathbf{L}_{\mathbf{K}} \mathbf{x} \right)$$

$$= \mathbf{x}^T \left(\mathbf{L}_{\mathbf{K}\Lambda} - \mathbf{L}_{\mathbf{K}}^2 \right) \mathbf{x}$$

Note that $\mathbf{a}_i - \mathbf{a} = 0$ if element i is a knowledge leader.

To complete the proof, we use the following lemma, which is derived in Appendix 1.

Lemma 1: Giving any $\mathbf{x}^T = [\mathbf{x}_1^T, \mathbf{x}_2^T, \ldots, \mathbf{x}_n^T]$ where $\mathbf{x}_i \in \mathbb{R}^m$, if

$$\frac{\lambda_{m+1}^2(\mathbf{L}_{\mathbf{K}})}{\lambda_n(\mathbf{L})} > \max_k \lambda_{max}(\mathbf{K}\Lambda)_{ijs} \quad (5)$$

$\mathbf{x}^T \left(\mathbf{L}_{\mathbf{K}\Lambda} - \mathbf{L}_{\mathbf{K}}^2 \right) \mathbf{x} \leq 0$ and the equality is true if and only if $\mathbf{x}_1 = \mathbf{x}_2 = \cdots = \mathbf{x}_n$.

Note that for condition (5) to be true, we need a connected network (otherwise $\lambda_{m+1}(\mathbf{L}_{\mathbf{K}}) = 0$), an upper bounded $\lambda_{max}(\mathbf{K}\Lambda)_{ijs}$, and strong enough coupling strengths (to make the inequality to be true). For an example, if $m = 1$ and all the coupling gains are identical with value κ, condition (5) turns to be

$$\kappa > \frac{\lambda_n(\mathbf{L})}{\lambda_2^2(\mathbf{L})} \, \max \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}_i, \mathbf{a}, t)$$

Theorem 2: For a knowledge-based leader-followers network, the states of all the elements will converge together asymptotically if condition (5) is verified and all the states are bounded. Furthermore, $\forall \, i \in \Omega$, $\mathbf{a}_i$ will converge to $\mathbf{a}$ if

$$\exists \, \alpha > 0, T > 0, \; \forall t \geq 0 \quad \int_t^{t+T} \mathbf{W}^T(\mathbf{x}_i, r) \mathbf{W}(\mathbf{x}_i, r) dr \geq \alpha \mathbf{I}$$
$$(6)$$

Proof: Condition (5) means V is non-increasing. Assuming all the functions are smoothly differentiable, the boundedness of $\ddot{V}$ can be concluded if all the states are bounded. According to Barbalat's lemma [27], $\dot{V}$ will then tend to 0 asymptotically, implying that all the states $\mathbf{x}_i$ converge together. Hence, $\mathbf{W}(\mathbf{x}_i, t)\tilde{\mathbf{a}}_i$ will tend to zero. This in turn, using standard persistency of excitation arguments [20], [27], leads to the convergence of the followers' parameters under condition (6). $\square$

Theorem 2 implies that new elements can be added into the network without prior knowledge of the individual dynamic parameters, and that elements in an existing network have the ability to recover dynamic information if temporarily lost. Similar knowledge-based leader-followers mechanism may exist in many natural processes. In evolutionary biology, knowledge leaders are essential to keep the evolution processes un-invasible or evolutionary stable [21], [23]. In reproduction, for instance, the leaders could be senior members. The knowledge-based mechanism may also describe evolutionary mutation or disease infection [16], where the leaders are mutants or invaders. Knowledge-based leader-following may also occur in animal aggregate motions or human social activities. In a bird flock, for instance, the knowledge leader can be a junior or injured member whose moving capacity is limited, and which is protected by others through dynamic adaptation.

Note that in evolutionary biology, the adaptive model we describe may loosely represent a genotype-phenotype mapping, where adaptation occurring in genotypic space is based on the interactions of behavioral phenotypes.

Example 3.1:: Consider six neurons of FitzHugh-Nagumo type (a well known spiking neuron model [6], [18], [19]), connected as in Figure 1(b)

$$\begin{cases} \dot{v}_i = v_i(\alpha_i - v_i)(v_i - 1) - w_i + I_i + \sum_{j \in \mathcal{N}_i} k_{ij}(v_j - v_i) \\ \dot{w}_i = \beta_i v_i - \gamma_i w_i \qquad i = 1, \ldots, 6 \end{cases}$$

Let the parameter set $\mathbf{a}_i = [\alpha_i, I_i, \gamma_i, \beta_i]^T$ be fixed for the knowledge leader, and let those of the other neurons change according to the adaptation law (3). Simulation results are plotted in Figure 2. $\square$

Remarks

• Leaders holding different parameters may exist in the same network, a situation somewhat analogous to a human society containing experts in different fields. For instance, consider (2) again. Assume the dynamics $\mathbf{f}$ contains l parameter sets $\mathbf{a}^1, \mathbf{a}^2, \ldots, \mathbf{a}^l$ with

$$\mathbf{f}(\mathbf{x}_i, \mathbf{a}_i^1, \ldots, \mathbf{a}_i^l, t) = \mathbf{f}(\mathbf{x}_i, \mathbf{a}^1, \ldots, \mathbf{a}^l, t) + \sum_{k=1}^l \mathbf{W}_k(\mathbf{x}_i, t)\tilde{\mathbf{a}}_i^k$$
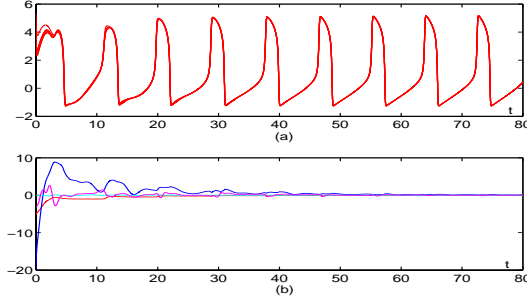
Fig. 2. Simulation results of Example 3.1. With initial conditions chosen arbitrarily, the plots show that (a).states v_i ($i = 1, \ldots, 6$) synchronize in the time space; (b).estimation error set $\tilde{\mathbf{a}}_i$ of any of the knowledge followers vanish in the time space.

Denoting by $\Omega^1, \Omega^2, \ldots, \Omega^l$ the followers sets corresponding to different parameters, the adaptation laws are, for $k = 1, 2, \ldots, l$,

$$\dot{\mathbf{a}}_i^k = \mathbf{P}_i^k \mathbf{W}_k^T(\mathbf{x}_i, t) \sum_{j \in \mathcal{N}_i} \mathbf{K}_{ji} (\mathbf{x}_j - \mathbf{x}_i) \quad \forall\, i \in \Omega^k$$

States and parameters will converge under the same conditions as those given in Theorem 2.

• To further improve the convergence rate, the adaptation law (3) may be refined as

$$\hat{\mathbf{a}}_i = \mathbf{a}_i + \mathbf{Q}_i \mathbf{W}^T(\mathbf{x}_i, t) \sum_{j \in \mathcal{N}_i} \mathbf{K}_{ji} (\mathbf{x}_j - \mathbf{x}_i)$$

where $\mathbf{Q}_i > 0$ is constant and symmetric, $\mathbf{a}_i$ is defined by (3), and $\hat{\mathbf{a}}_i$ is used in real dynamics, i.e., $\mathbf{f}_i = \mathbf{f}(\mathbf{x}_i, \hat{\mathbf{a}}_i, t)$. Note that in the theoretical analysis we should use a modified Lyapunov-like function

$$V = \frac{1}{2} \left(\mathbf{x}^T \mathbf{L_K} \mathbf{x} + \sum_{i \in \Omega} \tilde{\mathbf{a}}_i^T \mathbf{P}_i^{-1} \tilde{\mathbf{a}}_i \right) + \sum_{i \in \Omega} \int_0^t \mathbf{z}_i^T \mathbf{Q}_i \mathbf{z}_i dt$$

where $\tilde{\mathbf{a}}_i = \mathbf{a}_i - \mathbf{a}$ and $\mathbf{z}_i = \mathbf{W}^T(\mathbf{x}_i, t) \sum_{j \in \mathcal{N}_i} \mathbf{K}_{ji} (\mathbf{x}_j - \mathbf{x}_i)$.

• The number of leaders in a knowledge-based network can be arbitrary. At the limit all elements could be adaptive, i.e., there is no leader at all, in which case the system may converge to any odd parameter set depending on initial conditions. While all states will still converge together, the desired individual behaviors (such as oscillations) may not be preserved.

• Synchronization conditions derived in Theorem 2 are very similar to those in [28], [33] for coupled networks without any leader or adaptation. Note that if condition (5) is true, $\forall$ neighbored i, j, $\mathbf{x}_i - \mathbf{x}_j$ are bounded. Thus the boundedness of the states are simply determined by $\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, \mathbf{a}, t) + \mathbf{u}$ where the input $\mathbf{u}$ is bounded [12].

• Condition (6) is true if the stable system behaviors are sufficiently rich or persistently exciting [27]. This is the case, for instance, when the individual elements are oscillators, where the possibilities that any component of $\mathbf{x}_i$ converges to zero can be excluded by dynamic analysis showing that zero is an unstable state.

• Both power leaders and knowledge leaders could be virtual, which is common in animal aggregate motions. For instance, a landmark may be used as a virtual power leader. Similarly, when hunting, an escaping prey could specify both the where and the how of the movement.

• In a neural context, parameter changes in low order models may lead to entirely different qualitative behaviors [9], [25]. Thus, modifying the parameters of a single element, the knowledge leader, can select radically different qualitative properties for entire networks.

IV. PACIFIC COEXISTENCE

Different types of leaders can co-exist in the same network. As illustrated in Figure 1(c), a power leader guiding the direction may use state measurements from its neighbors to adapt its parameters to the values of the knowledge leaders. In addition, a power leader could be also a knowledge leader.

Consider the power-based leader-followers network (1) again, assuming the dynamics $\mathbf{f}$ contains a parameter set $\mathbf{a}$. There are knowledge leaders holding the fixed value $\mathbf{a}$ and knowledge followers using adaptation to learn. If $0 \in \Omega$, the set of the knowledge followers, we have

$$\dot{\mathbf{a}}_0 = \mathbf{P}_0 \mathbf{W}^T(\mathbf{x}_0, t) \sum_{i=1}^n \gamma_i \mathbf{K}_{0i} (\mathbf{x}_i - \mathbf{x}_0)$$

while if $i \in \Omega$ with $i = 1, \ldots, n$,

$$\dot{\mathbf{a}}_i = \mathbf{P}_i \mathbf{W}^T(\mathbf{x}_i, t) \Big(\sum_{j \in \mathcal{N}_i} \mathbf{K}_{ji} (\mathbf{x}_j - \mathbf{x}_i) + \gamma_i \mathbf{K}_{0i} (\mathbf{x}_0 - \mathbf{x}_i) \Big)$$

To prove state convergence, we first define several (extended) Laplacian matrices for a power-based network structure:

• $\mathbf{L_K}$, the weighted Laplacian of the followers network.

• $\vec{\mathbf{L}}_{\mathbf{K}}$, the weighted Laplacian of the whole network, which is non-symmetric since we have uni-directional links between the leader and the followers. Thus,

$$\vec{\mathbf{L}}_{\mathbf{K}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ -\mathbf{b} & \mathbf{C} \end{bmatrix} \quad \text{where} \quad \mathbf{b} = \begin{bmatrix} \vdots \\ \gamma_i \mathbf{K}_{0i} \\ \vdots \end{bmatrix}, \quad \mathbf{C} = \mathbf{L_K} + \mathbf{I}_{\gamma_i \mathbf{K}_{0i}}^n$$

$\mathbf{C}$ is positive definite if the whole network is connected.

• $\bar{\mathbf{L}}_{\mathbf{K}}$, the weighted Laplacian of the whole network which we consider as an undirected graph. Thus,

$$\bar{\mathbf{L}}_{\mathbf{K}} = \vec{\mathbf{L}}_{\mathbf{K}}^T + \begin{bmatrix} \sum_{i=1}^n \gamma_i \mathbf{K}_{0i} & \mathbf{0} \\ -\mathbf{b} & \mathbf{0} \end{bmatrix}$$

Define the Lyapunov-like function

$$V = \frac{1}{2} \left(\mathbf{x}^T \bar{\mathbf{L}}_{\mathbf{K}} \mathbf{x} + \sum_{i \in \Omega} \tilde{\mathbf{a}}_i^T \mathbf{P}_i^{-1} \tilde{\mathbf{a}}_i \right)$$

We can show that

$$\dot{V} = \mathbf{x}^T \bar{\mathbf{L}}_{\mathbf{K}} \left(\begin{bmatrix} \mathbf{f}(\mathbf{x}_1, \mathbf{a}, t) \\ \dots \\ \mathbf{f}(\mathbf{x}_n, \mathbf{a}, t) \end{bmatrix} - \vec{\mathbf{L}}_{\mathbf{K}} \mathbf{x} \right)$$
$$= \mathbf{x}^T \left(\bar{\mathbf{L}}_{\mathbf{K}\Lambda} - \vec{\mathbf{L}}_{\mathbf{K}}^T \vec{\mathbf{L}}_{\mathbf{K}} \right) \mathbf{x}$$

where $\bar{\mathbf{L}}_{\mathbf{K}\Lambda}$ is defined similarly to (4), except that here the incidence matrix is based on the whole network. See Appendix 2 for the condition for $\bar{\mathbf{L}}_{\mathbf{K}\Lambda} - \vec{\mathbf{L}}_{\mathbf{K}}^T \vec{\mathbf{L}}_{\mathbf{K}}$ to be negative semi-definite. Following the same proofs as those in Sections III, this then implies that all the states $\mathbf{x}_i$, $i = 0, 1, \dots, n$ will converge together asymptotically. Parameter convergence conditions are also the same.

Appendix 1: Proof of Lemma **1**

For notational simplicity, we show the derivations for the case $m = 1$.

Notice that 0 is always one of the eigenvalues of $\mathbf{L}_{\mathbf{K}\Lambda} - \mathbf{L}_{\mathbf{K}}^2$, with one corresponding eigenvector $\mathbf{v} = [1, 1, \dots, 1]^T$. According to Weyl's Theorem [8],

$$\lambda_{n-k+1}(\mathbf{L}_{\mathbf{K}\Lambda} - \mathbf{L}_{\mathbf{K}}^2) \leq \lambda_n(\mathbf{L}_{\mathbf{K}\Lambda}) - \lambda_k(\mathbf{L}_{\mathbf{K}}^2)$$

where $k = 1, 2, \dots, n$, and the eigenvalues λ_i are arranged in increasing order for $i = 1, 2, \dots, n$. This implies that, $\forall k > 1$, $\lambda_{n-k+1}(\mathbf{L}_{\mathbf{K}\Lambda} - \mathbf{L}_{\mathbf{K}}^2) < 0$ if

$$\lambda_n(\mathbf{L}_{\mathbf{K}\Lambda}) < \lambda_2(\mathbf{L}_{\mathbf{K}}^2) \tag{7}$$

Therefore, $\lambda_n(\mathbf{L}_{\mathbf{K}\Lambda} - \mathbf{L}_{\mathbf{K}}^2) = 0$, i.e., $\mathbf{L}_{\mathbf{K}\Lambda} - \mathbf{L}_{\mathbf{K}}^2$ is negative semi-definite.

Denote $\max_k \lambda_{max}(\mathbf{K}\Lambda)_{ijs} = \bar{\lambda}$. If $\bar{\lambda} \leq 0$, we have $\lambda_n(\mathbf{L}_{\mathbf{K}\Lambda}) \leq 0$ and both the conditions (7) and (5) are always true; if $\bar{\lambda} > 0$,

$$\lambda_n(\mathbf{L}_{\mathbf{K}\Lambda}) \leq \bar{\lambda} \lambda_n(\mathbf{L})$$

where $\mathbf{L}$ is the graph Laplacian matrix. Considering the fact that $\lambda_2(\mathbf{L}_{\mathbf{K}}^2) = \lambda_2^2(\mathbf{L}_{\mathbf{K}})$, condition (5) is sufficient to guarantee (7).

The eigenvectors of real symmetric matrix form an orthogonal basis of the state space. Without loss of generality, consider such an orthogonal eigenvector set, $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n\}$, of $\mathbf{L}_{\mathbf{K}\Lambda} - \mathbf{L}_{\mathbf{K}}^2$, where $\mathbf{v}_n = [1, 1, \dots, 1]^T$ is the only zero eigenvector. For any $\mathbf{x}$, we have

$$\mathbf{x} = \sum_{i=1}^{n} k_i \mathbf{v}_i \quad \text{and} \quad \mathbf{x}^T \left(\mathbf{L}_{\mathbf{K}\Lambda} - \mathbf{L}_{\mathbf{K}}^2 \right) \mathbf{x} = \sum_{i=1}^{n-1} \lambda_i k_i^2 \mathbf{v}_i^T \mathbf{v}_i$$

Since the eigenvalue $\lambda_i < 0$ $\forall i < n$, $\mathbf{x}^T \left(\mathbf{L}_{\mathbf{K}\Lambda} - \mathbf{L}_{\mathbf{K}}^2 \right) \mathbf{x} = 0$ if and only if $\mathbf{x} = k_n \mathbf{v}_n$, that is, $\mathbf{x}_1 = \mathbf{x}_2 = \dots = \mathbf{x}_n$.

In the case $m > 1$, we can follow the same proof except that the zero eigenvalue now has multiplicity m, and the corresponding eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_m\}$ are linear combinations of the orthogonal set $[\mathbf{I}, \mathbf{I}, \dots, \mathbf{I}]^T$ where $\mathbf{I} \in \mathbb{R}^{m \times m}$ is the identity matrix.

Appendix 2: Network with Both Types of Leaders

Similarly to the proof above, $\bar{\mathbf{L}}_{\mathbf{K}\Lambda} - \vec{\mathbf{L}}_{\mathbf{K}}^T \vec{\mathbf{L}}_{\mathbf{K}}$ is negative semi-definite if

$$\lambda_{n+1}(\bar{\mathbf{L}}_{\mathbf{K}\Lambda}) < \lambda_2(\vec{\mathbf{L}}_{\mathbf{K}}^T \vec{\mathbf{L}}_{\mathbf{K}})$$

and its only eigendirection for the zero eigenvalue is thus $\mathbf{v} = [1, 1, \dots, 1]^T$. Since

$$\vec{\mathbf{L}}_{\mathbf{K}}^T \vec{\mathbf{L}}_{\mathbf{K}} = \begin{bmatrix} \mathbf{b}^T \mathbf{b} & -\mathbf{b}^T \mathbf{C} \\ -\mathbf{C}\mathbf{b} & \mathbf{C}^2 \end{bmatrix}$$

we have

$$\lambda_2(\vec{\mathbf{L}}_{\mathbf{K}}^T \vec{\mathbf{L}}_{\mathbf{K}}) \geq \lambda_1(\mathbf{C}^2) = \lambda_1^2(\mathbf{C})$$

according to the Interlacing Eigenvalues Theorem for bordered matrices [8]. Thus a sufficient condition to guarantee negative semi-definite is

$$\lambda_1^2(\mathbf{C}) > \lambda_{n+1}(\bar{\mathbf{L}}_{\mathbf{K}\Lambda}) \tag{8}$$

This condition is similar to the one we derived in Theorem 1 for synchronization of pure power-based leader-followers network. Assuming all the coupling strengths are identical with value κ, condition (8) becomes

$$\kappa > \frac{\lambda_{n+1}(\bar{\mathbf{L}})}{\lambda_1^2(\mathbf{L} + \mathbf{I}_{\gamma_i}^n)} \max \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{x}_i, \mathbf{a}, t)$$

REFERENCES

[1] Cortes, J., and Bullo, F. (2003) Coordination and Geometric Optimization via Distributed Dynamical Systems, *SIAM Journal of Control and Optimization*

[2] Cutts, C., and Speakman, J. (1994) Energy Savings in Formation flight of Pink-Footed Geese, *J. Exp. Biol.*, **189**:251-261

[3] D'Andrea, R., and Dullerud, G.E. (2003) Distributed Control Design for Spatially Interconnected Systems, *IEEE Transactions on Automatic Control*, **48**(9):1478–1495

[4] Fiedler, M. (1973) Algebraic Connectivity of Graphs, *Czechoslovak Mathematical Journal*, **23**(98): 298-305

[5] Fierro, R., Song, P., Das, A., and Kumar, V. (2002) Cooperative Control of Robot Formations, in *Cooperative Control and Optimization: Series on Applied Optimization*, Kluwer Academic Press, 79-93

[6] FitzHugh, R.A. (1961) Impulses and Physiological States in Theoretical Models of Nerve Membrane, *Biophys. J.*, 1:445-466

[7] Godsil, C., and Royle, G. (2001) Algebraic Graph Theory, *Springer*

[8] Horn, R.A., and Johnson, C.R. (1985) Matrix Analysis, *Cambridge University Press*

[9] Izhikevich. E. (2003) Simple Model of Spiking Neuron, *IEEE Trans. on Neural Networks*

[10] Jadbabaie, A., Lin, J., and Morse, A.S. (2003) Coordination of Groups of Mobile Autonomous Agents Using Nearest Neighbor Rules, *IEEE Transactions on Automatic Control*, **48**:988-1001

[11] Jadbabaie, A., Motee, N., and Barahona, M. (2004) On the Stability of the Kuramoto Model of Coupled Nonlinear Oscillators, *Proceedings of the American Control Conference*

[12] Khalil H.K. (1996) Nonlinear Systems, *Prentice-Hall*

[13] Leonard, N.E., and Fiorelli, E. (2001) Virtual Leaders, Artificial Potentials and Coordinated Control of Groups, *40th IEEE Conference on Decision and Control*

[14] Lohmiller, W., and Slotine, J.J.E. (1998) On Contraction Analysis for Nonlinear Systems, *Automatica*, **34**(6)

[15] Lohmiller, W. (1999) Contraction Analysis of Nonlinear Systems, *Ph.D. Thesis, Department of Mechanical Engineering, MIT*

[16] May, R.M., Gupta, S., and McLean, A.R. (2001) Infectious Disease Dynamics: What Characterizes a Successful Invader? *Phil. Trans. R. Soc. Lond. B*, **356**:901-910

[17] Mohar, B. (1991) Eigenvalues, Diameter, and Mean Distance in Graphs, *Graphs and Combinatorics* **7**:53-64

[18] Murray, J.D. (1993) Mathematical Biology, *Springer-Verlag*

[19] Nagumo, J., Arimoto, S., and Yoshizawa, S. (1962) An Active Pulse Transmission Line Simulating Nerve Axon, *Proc. Inst. Radio Engineers*, 50:2061-2070

[20] Narendra, K., and Annaswamy, A. (1989) Stable Adaptive Systems, *Prentice-Hall*

[21] Nowak, M.A., and Sigmund, K. (2004) Evolutionary Dynamics of Biological Games, *Science*, **303**:793-799

[22] Olfati-Saber, R., and Murray, R.M. (2003) Consensus Protocols for Networks of Dynamic Agents, *American Control Conference, Denver, Colorado*

[23] Page, K.M., and Nowak, M.A. (2002) Unifying Evolutionary Dynamics, *J. theor. Biol.*, **219**:93-98

[24] Pikovsky, A., Rosenblum, M., and Kurths, J. (2003) Synchronization: A Universal Concept in Nonlinear Sciences, *Cambridge University Press*

[25] Schnitzler, A., and Gross, J. (2005) Normal and Pathological Oscillatory Communication in the Brain. *Nat. Rev. Neurosci.*

[26] Seiler, P., Pant, A., and Hedrick J.K. (2003) A Systems Interpretation for Observations of Bird V-formations, *Journal of Theoretical Biology*, **221**:279-287

[27] Slotine, J.J.E., and Li, W. (1991) Applied Nonlinear Control, *Prentice Hall*

[28] Slotine, J.J.E., and Wang, W. (2003) A Study of Synchronization and Group Cooperation Using Partial Contraction Theory, *Block Island Workshop on Cooperative Control*, Kumar V. Editor, Springer-Verlag

[29] Stan, G-B., and Sepulchre, R. (December, 2003) Dissipativity Characterization of a Class of Sscillators and Networks of Oscillators, *Proceedings of the 42nd IEEE Conference on Decision and Control, Hawai*

[30] Strogatz, S. (2003) Sync: The Emerging Science of Spontaneous Order, *New York: Hyperion*

[31] Tanner, H., Jadbabaie, A., and Pappas, G.J., (December 2003) Stable Flocking of Mobile Agents, Part I: Fixed Topology IEEE, *Conference on Decision and Control, Maui, HI*

[32] Vicsek, T. (2002) The Bigger Picture, *Nature*, **418**:131

[33] Wang, W., and Slotine, J.J.E. (2003) On Partial Contraction Analysis for Coupled Nonlinear Oscillators, submitted to *Biological Cybernetics*

[34] Yen, S.C., Menschik, E.D., and Finkel, L.H. (1999) Perceptual Grouping in Striate Cortical Networks Mediated by Synchronization and Desynchronization, *Neurocomputing*, 26-27(1-3):609-616