Synchronization Interfaces and Overlapping Communities in Complex Networks

D. Li,[1] I. Leyva,[2] J. A. Almendral,[2] I. Sendiña-Nadal,[2] J. M. Buldú,[2] S. Havlin,[1] and S. Boccaletti[3,4]

[1]*Department of Physics, Minerva Center, Bar Ilan University, Ramat Gan 52900, Israel*
[2]*Complex Systems Group, Universidad Rey Juan Carlos, c/ Tulipán s/n, 28933 Móstoles, Madrid, Spain*
[3]*Embassy of Italy in Tel Aviv, 25 Hamered Street, 68125 Tel Aviv, Israel*
[4]*CNR–Istituto dei Sistemi Complessi, Via Madonna del Piano, 10, 50019 Sesto Fiorentino (Fi), Italy*

(Received 12 June 2008; published 14 October 2008)

We show that a complex network of phase oscillators may display interfaces between domains (clusters) of synchronized oscillations. The emergence and dynamics of these interfaces are studied for graphs composed of either dynamical domains (influenced by different forcing processes), or structural domains (modular networks). The obtained results allow us to give a functional definition of overlapping structures in modular networks, and suggest a practical method able to give information on overlapping clusters in both artificially constructed and real world modular networks.

DOI: 10.1103/PhysRevLett.101.168701 PACS numbers: 89.75.Hc, 05.45.Xt

The functioning of many natural (biological, neural, chemical) or artificial (technological) networks displays coordination of parallel tasks [1]. This phenomenon may be represented as the interplay between two simultaneous processes. The first (involving most of the network nodes) leads to the emergence of organized clusters (or moduli, or cohesive subgroups), where nodes in the same cluster adjust their dynamics into a common (synchronized) behavior to enhance the performance of a specific task. The second process (involving just a few nodes of the graph) is to form interfaces (or overlapping structures) between the moduli that are responsible for the coordination between the different tasks.

In this Letter, we report the first evidence that, under the presence of different functional (synchronized) clusters, interfaces appear and show a specific dynamical behavior that enables one to develop an algorithm for identifying their structure in a modular network. Indeed, the study of separate modular structures [2] and synchronization [3] in complex graphs has so far not unravelled the crucial point concerning the role of synchronization interfaces and their usefulness in detecting overlapping communities. We report results on networks consisting of two domains of interacting phase oscillators (each one synchronously evolving at a different frequency), where the nature of the two different frequency domains is the result of either a dynamical process (influenced by different forcing processes) or a structural design (modular network). Under these conditions, most of the oscillators will contribute to the synchronous behavior of the two clusters, whereas a few nodes will be in a *frustrated situation* due to contrasting inputs from the two clusters. Moreover, we propose an analytic treatment of an abstracted system that yields further insight. Based on our findings, we also develop an algorithm that is able to detect overlapping structures in both artificially constructed and real modular networks.

Let us start with the case of a generic random graph G of N coupled oscillators, whose original frequencies $\{\omega_i\}$ are randomly drawn from a uniform distribution in the interval 0.5 ± 0.25, subject simultaneously to an internal bidirectional coupling and an external pacemaking unidirectional forcing. The network dynamics is described by

$$\dot{\phi}_i = \begin{cases} \omega_i + \frac{d}{(k_i+k_{p_i})} \sum_{j=1}^{N} a_{ij} \sin(\phi_j - \phi_i) \\ + \frac{d_p k_{p_i}}{(k_i+k_{p_i})} \sin(\phi_{p_i} - \phi_i), \end{cases} \quad (1)$$

where dots denote temporal derivatives, k_i is the degree of the ith oscillator, ϕ_{p_i} is the instantaneous phase of a forcing oscillator having k_{p_i} unidirectional connections, d and d_p are coupling strengths, and the $\{a_{ij}\}$ are either 1 or 0 depending on whether or not a link exists between node i and node j.

In our simulations, we study a network G which consists of $N = 200$ phase oscillators arranged in an Erdös-Rényi configuration [4]. Initially, we set $d_p = 0$ and $k_{p_i} = 0 \forall i$, and we choose $d = 0.1$ so that G exhibits unsynchronized motion. Next, we arbitrarily divide the nodes into two groups: nodes from $i = 1, \ldots, 100$ (from $i = 101, \ldots, 200$) are assigned to the community A (B). We introduce two pacemakers of frequencies ω_{p_A} and ω_{p_B}, and connect the nodes in the first (second) group with the first (second) pacemaker. This implies in Eq. (1) that $\phi_{p_i} = \phi_{p_A} \equiv \omega_{p_A} t$ ($\phi_{p_i} = \phi_{p_B} \equiv \omega_{p_B} t$) for all the nodes in A (B). In order to assign its k_{p_i} links with the pacemaker to each node, we start at $t_0 = 0$ the evolution of Eq. (1) from random initial conditions in the unforced case ($k_{p_i} = 0 \forall i$), and add links between nodes of the two communities and the pacemakers at later times $t_l = t_0 + l\Delta t$. At precisely each time t_l, the pacemaker ϕ_{p_A} (ϕ_{p_B}) forms a connection with that node j in A (B) whose instantaneous phase at time t_l corresponds to the minimum $\min_j|\delta - \Delta\theta_j \mathrm{mod} 2\pi|$, with $\Delta\theta_j = \phi_j(t_l) - \phi_p(t_l)$, $\delta \in (0, 2\pi)$ setting a specific desired phase condition [5], and $\phi_p(t_l) = \phi_{p_A}(t_l)$ [$\phi_p(t_l) = \phi_{p_B}(t_l)$] for those nodes in A (B). In [5], it was demonstrated that, by operating this attachment over a given time interval, the resulting dynamics of any arbitrary network of oscillators can be entrained to any arbitrary frequency ω_p.

By selecting $\omega_{p_A} = 0.7$, $\omega_{p_B} = 0.3$, and $\sum_{i=1}^{N} k_{p_i} = 2000$ links to the pacemakers (1000 to entrain the nodes in A and the other 1000 to entrain those in B), and $d_p = 1$, the dynamics display two large communities of entrained oscillators. The situation is depicted in Fig. 1(a), where we report the instantaneous frequency of each oscillator in G as a function of time (averaged over a small window to smooth fluctuations). We observe that most of the nodes in A (B) have a constant frequency (that of the corresponding pacemaker), whereas the few nodes belonging to the synchronization interface exhibit a frequency that oscillates around the mean value of the frequency of the two communities $\bar{\omega} = (\omega_{p_A} + \omega_{p_B})/2$.

This switching mechanism is the result of the competition between two conflicting processes: the synchronization within G (controlled by parameter d) that would lead the whole network to exhibit a unique frequency, and the forcing of the two pacemakers (controlled by d_p) that tends to separate the nodes into two clusters of entrained oscillators. In order to quantify this competition and to describe the size of the synchronization interface, we fix $d_p = 1$ and gradually increase d. The results are shown in Figs. 1(a)–1(c) with (a) $d = 3.25$, (b) $d = 5.75$, and (c) $d = 9.75$. For intermediate coupling [Fig. 1(b)], it is observed that the interface attracts more and more members as the coupling increases. Because of the presence of the two forcing pacemakers, this interface is organized in an oscillating mode rather than in a constant frequency mode. As the coupling d is further increased, almost the entire system of oscillators eventually participates into this interface oscillating mode, as seen in Fig. 1(c). Finally, Fig. 1(d) gives evidence that in the low coupling regime ($d = 3.5$), the period T_O of the switching process is inversely proportional to $\omega_\Delta = (\omega_{p_A} - \omega_{p_B})/2$. Notice that similar switching dynamics were previously observed in

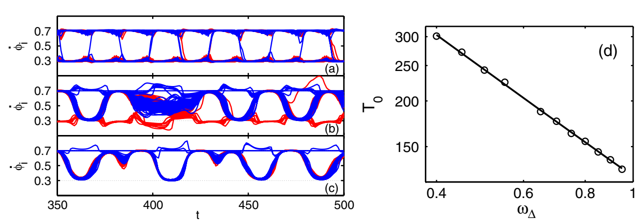

FIG. 1 (color online). (a–c) Instantaneous frequencies $\dot{\phi}_i(t)$ of each one of the 100 oscillators in A (light lines) [in B (dark lines)] vs time, obtained from simulation of Eq. (1) for (a) $d = 3.25$, (b) $d = 5.75$, and (c) $d = 9.75$. (d) Log-log plot of the period T_O of the switching process in the interface vs ω_Δ, for $d = 3.5$. The solid line represents a linear fit with $T_O \sim 120/\omega_\Delta$. Each point is the average of 5 independent realizations. Other parameters reported in the text.

the case of a chain of oscillators subject to two forcing frequencies applied to the two ends of the chain [6].

In order to give analytical insight to our findings, let us consider the simple case of three interacting phase oscillators described by $\dot{\phi}_1 = \omega_1 + K_1 \sin(\phi_3 - \phi_1)$, $\dot{\phi}_2 = \omega_2 + K_2 \sin(\phi_3 - \phi_2)$, $\dot{\phi}_3 = \omega_3 + K[\sin(\phi_1 - \phi_3) + \sin(\phi_2 - \phi_3)]$. Here, ϕ_3 is the phase of an oscillator (with natural frequency ω_3) receiving simultaneous coupling from two other oscillators at natural frequencies $\omega_1 \neq \omega_2 \neq \omega_3$; $K_1, K_2 \ll K$ are coupling constants. The oscillators 1 and 2 model the two functional (frequency) domains A and B, where the nature of the two different frequency domains results from either a dynamical process or a structural design. These two domains of synchronous oscillators simultaneously interact with the small group of nodes in the interface (modeled by the third oscillator), so that we can reasonably assume $K_1 = K_2 = 0$, and consequently $\phi_{1,2} = \omega_{1,2} t$.

The resulting equation $\dot{\phi}_3 = \omega_3 + 2K \sin(\frac{1}{2}(\omega_1 + \omega_2)t - \phi_3)\cos(\frac{1}{2}(\omega_1 - \omega_2)t)$ has an analytic solution

$$\dot{\theta}_3 = \frac{\tilde{A}\cos(\omega_\Delta t)\exp[\frac{2K}{\omega_\Delta}\sin(\omega_\Delta t)]}{1 + \tilde{B}\exp[\frac{4K}{\omega_\Delta}\sin(\omega_\Delta t)]}, \qquad (2)$$

for $\omega_3 = \bar{\omega}$, where $\theta_3 = \phi_3 - \bar{\omega}t$, $\bar{\omega} = \frac{1}{2}(\omega_1 + \omega_2)$, $\omega_\Delta = \frac{1}{2}(\omega_1 - \omega_2)$, and $\tilde{A}$, $\tilde{B}$ are suitable parameters. For intermediate coupling, the third node switches between the instantaneous frequencies of the other two, exactly as does the synchronization interface. Furthermore, in good agreement with the preceding discussion, Eq. (2) predicts that (once ω_3 is selected to be the mean frequency of the two forcing clusters) the instantaneous frequency $\dot{\phi}_3$ will oscillate around $\bar{\omega}$ with a period T_O *inversely proportional* to the difference in the frequencies of the two forcing clusters. Notice that as K increases (the strong coupling regime), $\dot{\theta}_3$ will approach zero, implying that the frequency of the third oscillator will be progressively damped to the mean frequency of the two other oscillators, even though there is no threshold for this damping phenomenon.

So far we have considered the behavior of interfaces as the result of the competition of dynamical domains. We now describe the competition of structural domains in modular graphs in the absence of the forcing [$d_p = 0$ and $k_{p_i} = 0 \forall i$ in Eq. (1)]. For this purpose, we construct the adjacency matrix of G by considering two large communities (A, B), each formed by 50 densely and randomly connected nodes (the average degree in the same community is 16), which overlap a small community O made of a complete graph of 5 nodes [3 random links to A (B)] that form symmetric connections to nodes in both A and B. Each of the 105 nodes of G is associated with a phase oscillator obeying Eq. (1), which is integrated for an initial distribution of frequencies such that nodes in A (B) [i.e., nodes from $i = 1$ to 50 (51 to 100)] have natural fre quencies uniformly distributed in the range 0.25 ± 0.25

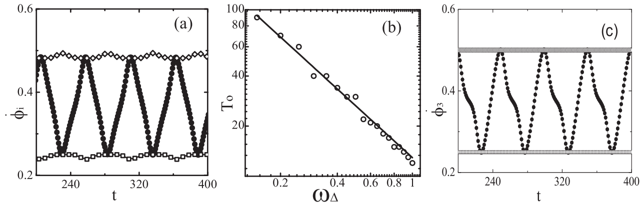

FIG. 2. (a) Instantaneous frequencies $\dot{\phi}_i(t)$ vs time from simulation of Eq. (1) with $d = 0.1$ (other parameters and stipulations are reported in the text). Squares, diamonds, and full circles represent, respectively, nodes belonging to A, B, and O. (b) Log-log plot of the switching period T_O of the oscillations in the frequency of the nodes in O vs the frequency difference ω_Δ. The solid line represents a linear fit with slope -1.002 ± 0.0069. (c) Solution of Eq. (2) for $\omega_1 = 0.5$, $\omega_2 = 0.25$ [representing the frequencies of the two main clusters in (a)], $K = 0.1$, $\tilde{A} = 0.019$, $\tilde{B} = 0.25$.

(0.5 ± 0.25), while nodes in O have frequencies uniformly distributed in the range 0.375 ± 0.05) (i.e., around the mean frequency of the two distributions).

Figure 2(a) shows that all oscillators in the communities A and B behave synchronously with almost constant frequency in time (closely approximating the mean of the original frequency distribution), while all oscillators in O constitute the synchronization interface and thus display an instantaneous frequency oscillating in time around the mean value of the two frequencies in the two clusters. Figure 2(b) shows, in agreement with the analytical prediction, that the period of the frequency oscillations of O scales inversely proportional to the frequency difference between the two communities A and B. Finally, Fig. 2(c) gives the numerical results of Eq. (2) for $\omega_1 = 0.5$, $\omega_2 = 0.25$ [representing the frequencies of the two main clusters in (a)]. A direct comparison with Fig. 2(a) allows one to appreciate how closely our numerical results agree with the analytical model. Furthermore, all other analytical predictions of Eq. (2) concerning the large coupling regime are confirmed in our simulations. Indeed, Fig. 3(a) and 3(b) shows that increasing the coupling strength d yields the synchronized frequency of the interface to lock almost always to the mean of the frequencies of the two communities, still showing persistent events of shooting to larger (or lower) frequencies.

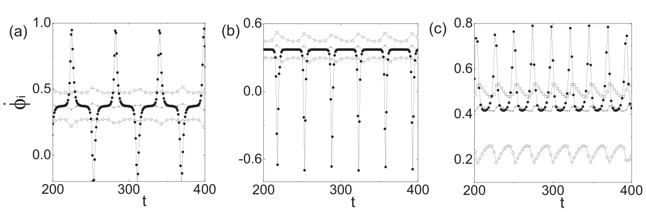

FIG. 3. Effects of coupling strength (a–b) and asymmetry (c). (a) $d = 0.5$; (b) $d = 0.95$. (c) $d = 0.1$, but 2 (5) links of the nodes in the interface go to nodes of cluster A (B). As in Fig. 2, squares, diamonds, and full circles refer, respectively, to nodes belonging to A, B, and O.

We have so far described the specific case in which the nodes of the interface had initially, by construction, the same number of links as the nodes of the two communities. It is therefore interesting to ask what happens when the nodes in O are asymmetrically connected to A and B. While the low coupling regime does not substantially differ from the symmetrical case, the high coupling regime [illustrated in Fig. 3(c)] exhibits frequency oscillations of the nodes in O that are biased toward the community in which the nodes have more connections.

The overall scenario reported above suggests a practical way to detect overlapping communities [Fig. 4(a)] in generic modular networks. It is important to remark that most of the definitions of network communities proposed so far lead essentially to a graph partition into components, such that a given node belongs to and only to one of the components of the partition [2]. The possibility, instead, that two components of a partition may have an overlapping set of nodes has been recently investigated by means of topological arguments [7]. The novelty of our approach consists in introducing a functional concept of overlapping structures that are defined in relationship to the dynamical response of the network as a whole. Namely, as far as synchronized behavior of phase oscillators is concerned, we define an overlapping structure as the set of nodes which, instead of following the constant frequency of one of the two domains, balance their instantaneous frequencies in between these two. Therefore, they cannot be considered as a functional part of any single domain. As shown below, this definition allows the detection of additional information including single overlapping nodes, as compared to previous studies.

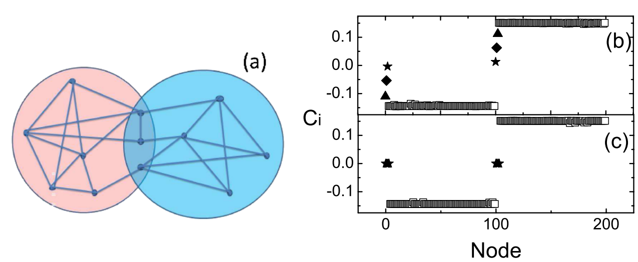

FIG. 4 (color online). (a) Graphical illustration of the constructed modular network, where the overlapping region of two circles (communities) represents the overlapping community. (b) C (see text for definition) vs node index in G. Overlapping nodes 1, 2, 3 (101, 102, 103), labeled, respectively, with triangle, diamond, and star, have [1,5], [2,4], [3,3] connections with nodes of [B, A] ([A, B]), while 4–100 (104–200) labeled with squares belong functionally to A (B) and $d = 0.15$. (c) Overlapping clusters each consisting of complete graphs of nodes 1–4 (101–104), labeled with stars, have symmetrical connections with A and B, while nodes 5–100 (105–200), labeled with squares, belong functionally to A (B); here $d = 0.1$.

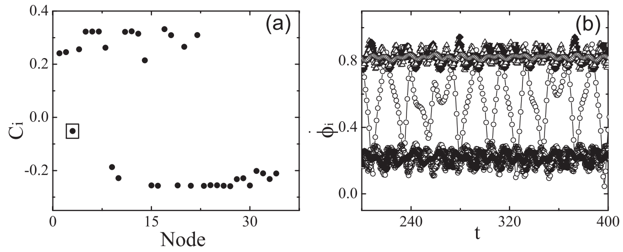

FIG. 5. (a) C (see text for definition) vs node index in the Zachary karate club network [8]. The values of C_i are visibly split into two main groups that identify exactly the split of the club that was originally observed in [8]. Node 3 (further labeled with a square) is ambiguously classified by different classical algorithms [2] and displays a value of C_i consistently different from that of any other node. (b) Instantaneous frequencies $\dot{\phi}_i(t)$ vs time from simulation of Eq. (1) with $d = 0.1$. The dynamics of node 3 (open circles) is the only one that reflects the typical oscillating behavior of the interfaces.

To illustrate this idea, we construct a network G made of two large moduli (A and B of 100 nodes each), where the majority of nodes form random connections (with average degree 15) with elements of the same community, while only very few nodes form links with nodes of both communities. Precisely, we denote by k_i^A (k_i^B) the total number of links these few nodes form with nodes in A (B), and define the ratio $D_i = k_i^A/k_i^B$ to evaluate the degree of overlapping of these nodes. For perfect overlap between the two clusters, $D_i = 1$. Under these conditions, Eq. (1) is simulated for $d_p = 0$, $k_{p_i} = 0 \forall i$, the set $\{\omega_i\}$ drawn from a Gaussian distribution with standard deviation 0.1 and mean value 0.3 (0.6) for nodes belonging to community A (B). To identify those overlapping nodes, we introduce the quantity $C_i = \mathrm{sgn}[\dot{\phi}_i(t) - \bar{\omega}]\min_t\{|\dot{\phi}_i(t) - \bar{\omega}|\}$, where $\bar{\omega}$ is the mean of the two averaged frequencies assigned to the two communities. This allows one to monitor how closely in time the dynamics of a node approaches $\bar{\omega}$. When the node is in the dynamics of the full synchronized interface, $C_i = 0$. Therefore, as D_i approaches 1, C_i is expected to approach 0. The results are shown in Fig. 4(b) and 4(c) for two different arrangements of the overlapping community: overlapping nodes [Fig. 4(b)] and overlapping clusters [Fig. 4(c)] which have symmetrical connections to two clusters. In both cases, two large synchronized clusters are identified very far from the overlapping synchronization, corresponding to those nodes performing distinct tasks and already classified by the structural partition. At the same time, the dynamical evolution manifests a group of nodes whose dynamics is significantly removed from the two main clusters (thus identifying the overlapping community). In Fig. 4(b), each overlapping node gives rise to a different value of C_i corresponding to its specific degree of overlap with D_i. The node with $D_i = 1$ yields $C_i = 0$. On the contrary, in Fig. 4(c), all the nodes inside the overlapping cluster are identified as a whole and have the same value of $C_i = 0$ due to the symmetrical connection.

Finally, we apply the algorithm to the network of friendship relationships between the members of a karate club, obtained from data collected by anthropologist Wayne Zachary [8]. The data consist of a graph having 34 nodes and 78 edges, that has been considered the comparative reference for most of the classical algorithms for the detection of separate modular structures [2]. We assign sets $\{\omega_i\}$ drawn from uniform distributions centered around 0.3 (0.7) with width 0.2 to nodes of the first (second) main community identified by classical studies [2]. Our algorithm shows not only that the values C_i allows one to describe perfectly the real split of the club [see Fig. 5(a)], but also that node 3 (which is ambiguously classified by different classical algorithms) constitutes the interface between the communities, and its associated instantaneous frequency [see Fig. 5(b)] displays the typical oscillating behavior of the interfaces that was introduced in our study. The fact that our algorithm gives such extra information on the structure of the Zachary club confirms that it can be of practical relevance for applications to other real networks to obtain novel insight on overlapping communities.

Work partly supported by EU Contract No. 043309 GABA, by the Spanish Ministry of S&T under Project No. FIS2006-08525, and by the URJC-CM under Projects No. URJC-CM-2006-CET-0643 and No. 2007-CET-1601. We also wish to thank the EU project DAPHNet, ONR, the Israel Science Foundation, Hadar Foundation, and the Center of Complexity Science for financial support.

[1] R. Albert and A. L. Barabási, Rev. Mod. Phys. **74**, 47 (2002); M. E. J. Newman, SIAM Rev. **45**, 167 (2003); S. Boccaletti *et al.*, Phys. Rep. **424**, 175 (2006).
[2] M. Girvan and M. E. J. Newman, Proc. Natl. Acad. Sci. U.S.A. **99**, 7821 (2002); F. Radicchi *et al.*, Proc. Natl. Acad. Sci. U.S.A. **101**, 2658 (2004); R. Guimerá and L. A. N. Amaral, Nature (London) **433**, 895 (2005).
[3] A. E. Motter, C. S. Zhou, and J. Kurths, Europhys. Lett. **69**, 334 (2005); Phys. Rev. E **71**, 016116 (2005); M. Chavez *et al.*, Phys. Rev. Lett. **94**, 218701 (2005); D.-U. Hwang *et al.*, Phys. Rev. Lett. **94**, 138701 (2005); S. Boccaletti *et al.*, Phys. Rev. E **74**, 016102 (2006).
[4] P. Erdös and A. Rényi, Publ. Math. Debrecen **6**, 290 (1959); P. Erdös and A. Rényi, Publ. of the Math. Inst. of the Hungarian Acad. of Sciences **5**, 17 (1960).
[5] I. Sendiña-Nadal *et al.*, PLoS ONE **3**, e2644 (2008).
[6] I. Leyva, E. Allaria, S. Boccaletti, and F. T. Arecchi, Phys. Rev. E **68**, 066209 (2003).
[7] G. Palla, I. Derényi, I. Farkas, and T. Vicsek, Nature (London) **435**, 814 (2005).
[8] W. W. Zachary, Journal of Anthropological Research **33**, 452 (1977).