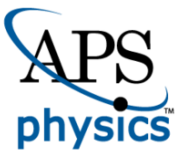

Network structure, topology, and dynamics in generalized models of synchronization

Kristina Lerman and Rumi Ghosh

Phys. Rev. E **86**, 026108 — Published 13 August 2012

DOI: 10.1103/PhysRevE.86.026108

Network Structure, Topology and Dynamics in Generalized Models of Synchronization

Kristina Lerman and Rumi Ghosh
USC Information Sciences Institute
4676 Admiralty Way, Marina del Rey, CA 90292

Network structure is a product of both its topology and interactions between its nodes. We explore this claim using the paradigm of distributed synchronization in a network of coupled oscillators. As the network evolves to a global steady state, nodes synchronize in stages, revealing network's underlying community structure. Traditional models of synchronization assume that interactions between nodes are mediated by a conservative process similar to diffusion. However, social and biological processes are often non-conservative. We propose a new model of synchronization in a network of oscillators coupled via non-conservative processes. We study dynamics of synchronization of a synthetic and real-world networks and show that the traditional and non-conservative models of synchronization reveal different structures within the same network.

PACS numbers: 05.45.Xt, 89.75.Hc, 89.75.-k, 89.65.Ef, 89.75.Fb, 02.10.Ud

I. INTRODUCTION

Community structure is an important characteristic of real-world complex networks, including biological networks composed of functional modules [1, 2], and social networks, which are often composed of groups of similar individuals [3]. Researchers developed a large variety of methods [3–6] that examine network topology, or connectivity between its nodes, to identify interesting structures. We claim, however, that community structure of complex networks is the product of both their topology *and dynamical processes* taking place on them. These processes, which are mediated by interactions between nodes, determine the phenomena taking place on the network, whether diffusion and other types of transport in biological networks, or epidemics and information spread in social networks. While it is generally acknowledged that network structure affects evolution of macroscopic phenomena [7–11], the impact of dynamics on our understanding of network structure is less appreciated.

We explore the connection between network structure, topology and dynamical processes by studying synchronization of coupled oscillators. Kuramoto [12] introduced a simple model of distributed synchronization that was adapted to networks of oscillators whose phases are coupled to their neighbors' phases [10, 13]. These systems demonstrate an interesting connection between dynamics and structure: as the network evolves to a global steady state, oscillators synchronize in stages, with synchronized groups revealing network's community structure [14, 15].

While the Kuramoto model is useful for studying physical systems whose elements are coupled via a diffusive process, it is not a good model of social and biological phenomena, for example, viral contagion. In a contagion process, rather than pick one neighbor to infect, a node will attempt to infect every neighbor. We introduce a new model of synchronization for such interactions and use it to study the structure of two example networks: a synthetic network with a hierarchical community structure and a benchmark social network. We show that the proposed model reveals a different view of community structure than that found by the traditional Kuramoto model, demonstrating the importance of accounting for interactions in the analysis of network structure.

II. GENERALIZED MODELS OF SYNCHRONIZATION

We consider a network of N identical oscillators, each interacting locally with its neighbors. We represent this network as an unweighted, undirected graph with an adjacency matrix A, such that $A[i,j] = 1$ if there exists an edge between nodes i and j; otherwise, $A[i,j] = 0$. In the linear version of the Kuramoto model, the phases of oscillators, specified by vector θ, evolve in time according to:

$$\frac{d\theta(t)}{dt} = \omega - K \cdot L\theta(t), \tag{1}$$

where the vector ω corresponds to natural frequencies of the nodes, K is a matrix of coupling constants of pairs of nodes, and $L = D - A$ is the Laplacian matrix of the graph, where the diagonal degree matrix D is defined as $D[i,i] = \sum_i A[i,j] = d_i$ and $D[i,j] = 0 \;\forall\; i \neq j$. When $\omega_i = 0$, network as a whole reaches a fully synchronized state in which $\theta_i(t) = \theta$, $\forall i$.

Oscillators in the linear Kuramoto model are coupled via a diffusive process whose dynamics is governed by the Laplacian matrix. Such a process cannot describe some real-world interactions, including social interactions that lead to the spread of a disease or information in social networks. Social interactions often have a non-conservative flavor: for example, a pathogen replicates as disease spreads. To distinguish dynamics arising from such interactions from those based on the Laplacian, we refer to the former as *non-conservative* and latter as *conservative*.

Consider a simple example of non-conservative interactions in which node i produces an amount $\alpha\theta_i$ of some substance, for example, a pathogen, but transmits an amount θ_i to each neighbor. Therefore, $(\alpha - d_i)\theta_i$ of

the amount of substance created by i is not transferred to any neighbor and is lost. Dynamics of a network of nodes coupled via such non-conservative interactions can be written as:

$$\frac{d\theta(t)}{dt} = \omega - K \cdot (\alpha I - A)\theta(t), \tag{2}$$

where I is the identity matrix and α is a constant. The model above introduces the *Replicator* operator $R = (\alpha I - A)$, the non-conservative counterpart of the Laplacian. In spite of non-conservation, under some conditions, specified below, the system reaches a steady state in which the variables $\theta_i(t)$ no longer change.

Both conservative and non-conservative models are special cases of the generalized linear synchronization model, which can be written in terms of the operator $\mathcal{L}(A)$ of the adjacency matrix A:

$$\frac{d\theta(t)}{dt} = \omega - K \cdot \mathcal{L}(A)\theta(t). \tag{3}$$

The operator $\mathcal{L}(A)$ captures the details of interactions and governs dynamics of synchronization. Different versions of $\mathcal{L}(A)$ are possible for both conservative and non-conservative interactions. If, for example, node i produces an amount θ_i and divides it evenly among its d_i neighbors, the conservative synchronization model in this case can be written in terms of the normalized Laplacian $\mathcal{L} = I - D^{-1}A$, rather than L.

A. Steady State

Solving Equation 3 above we get:

$$\theta(t) = e^{-K\cdot\mathcal{L}(A)t}\left(\theta_0 - \left(K \cdot \mathcal{L}(A)\right)^{-1}\omega\right) + \left(K \cdot \mathcal{L}(A)\right)^{-1}\omega, \tag{4}$$

where θ_0 is the initial phase of the oscillators at $t = 0$. Assuming that matrix $\mathcal{L}$ is diagonalizable, it can be written as an eigenvalue decomposition $\mathcal{L} = \sum_{i=1}^{N} \mathcal{X}[.,i]\lambda_i\mathcal{X}^{-1}[i,.]$, where the ith column of $\mathcal{X}$ is the ith eigenvector of $\mathcal{L}$ with eigenvalue λ_i. Then for $\omega_i = 0$ and $K_{ij} = c$, $\forall i, j$, Eq. 4 can be written as:

$$\theta(t) = \theta_0 e^{-c\mathcal{L}t} = \sum_{i=1}^{N} \mathcal{X}[.,i]e^{-c\lambda_i t}\mathcal{X}^{-1}[i,.]\theta_0 = \sum_{i=1}^{N} \mathcal{X}[.,i]e^{-c\lambda_i t}c_i \tag{5}$$

Here $c_i = \mathcal{X}^{-1}[i,.]\theta_0$ is a constant.

A non-trivial steady state $\theta(t \to \infty) \neq 0$ exists when $\lambda_1 = 0$. Under this condition, as $t \to \infty$, Eq. 5 reduces to $\theta^s = \mathcal{X}[.,1]c_1$, with constant c_1. The steady state solutions of different synchronization models are:

$\mathcal{L}(A) = D - A = L$**:** In this case $\mathcal{X}[.,1] = \bar{1}$ (vector of 1s). Hence, $\theta_i^s = \theta\ \forall\ i$, i.e., all oscillators have the same phase in the steady state.

$\mathcal{L}(A) = I - D^{-1}A$**:** In the steady state, $\theta_i^s \propto d[i]$, where $d[i]$ is the degree of node i.

$\mathcal{L}(A) = \alpha I - A = R$**:** For $\alpha = \lambda_{max}$, in the steady state θ^s is proportional to the eigenvector of A corresponding to its largest eigenvalue λ_{max}. For $\alpha > \lambda_{max}$ the steady state has a trivial solution $\theta_i^s \to 0,\ \forall i$.

$\mathcal{L}(A) = I + A$**:** This operator describes dynamics of a network of nodes coupled via a signaling process, studied by Hu et al. [16]. However, it can be shown that this system does not have a steady state.

B. Spectral Properties of $\mathcal{L}(A)$

The spectrum of $\mathcal{L}(A)$ gives information about topological and temporal scales of synchronization. Specifically, the temporal scale of synchronization, or the rate at which the system asymptotically relaxes to equilibrium, is dominated by the term in Eq. 5 corresponding to the smallest positive eigenvalue of $\mathcal{L}(A)$ (when the smallest eigenvalue is zero).

The dynamics of the conservative Kuramoto model is governed by the Laplacian L. Therefore, the synchronization time scale of the Kuramoto model is proportional to the inverse of smallest positive eigenvalue of L; gaps between its consecutive eigenvalues are related to the relative difference in synchronization time scales of different components [14, 15]; and the number of null eigenvalues of L gives the number of disconnected components.

In non-conservative synchronization, the synchronization time scale depends on the smallest positive eigenvalue of $R = \alpha I - A$ when $\alpha = \lambda_{max}$, the largest eigenvalue of A. Moreover, the spectrum of R gives insights into the community structure of the network. When a network has C disjoint communities, the adjacency matrix A has C eigenvalues that are significantly greater than the remaining $N - C$ eigenvalues [19]. Since the eigenvalue spectrum of $R = \lambda_{max}I - A$ is related to the spectrum of A, in a network with C disjoint communities, R has one null eigenvalue and $C - 1$ eigenvalues that are much closer to zero than the remaining $N - C$ eigenvalues.

The link between network structure and conservative dynamics has been studied in terms of the properties of random walks on graphs [5, 17], whose dynamics is described by the normalized Laplacian $\mathcal{L} = I - D^{-1}A$. This link is also captured by conductance of the network, which bounds the mixing time of a random walk on the network [18]. Conductance ϕ measures the quality of a partition of the network into two communities S and $\bar{S}$. It is defined as the ratio of the number of edges between nodes in S and nodes in $\bar{S}$ ($E(S,\bar{S})$) to the total number of edges within S ($vol(S)$): $\phi = \min_S E(S,\bar{S})/vol(S)$. According to Cheeger inequality ϕ is bounded by λ_2, the second largest eigenvalue of normalized Laplacian [4].

Therefore, when λ_2 is small, a good (i.e., low conductance) community division of the network exists. In this case, it will take a long time for a random walk to fully mix on the network [18]. This property of random walks has led to efficient algorithms to identify community structure of networks [5].

C. Synchronization and Network Structure

While spectral analysis can illuminate aspects of network structure, simulating dynamics of synchronization offers a more computationally efficient method to identify network structure. As demonstrated by Arenas et al. [14], nodes in the Kuramoto model synchronize in stages, with smaller components synchronizing before larger components, etc., until the entire network becomes synchronized. These stages reveal the hierarchical community structure of the network.

We quantify the degree of synchronization of nodes i and j at time t using similarity function:

$$s_{ij}(t) = cos\big(\theta_i(t) - \frac{\theta_i^s}{\theta_j^s}\theta_j(t)\big), \tag{6}$$

where θ_i^s is the phase of node i in the steady state. The rationale for this metric is that when nodes reach steady state, further interactions should not change their phases. In the conservative case, all phases are equal in the steady state; therefore, Eq. 6 reduces to the order parameter $s_{ij} = cos(\theta_i(t) - \theta_j(t))$ used in [14]. As nodes in the same community synchronize with each other, their similarity grows. Nodes are maximally similar ($s_{ij}(t) = 1$) in the steady state for both conservative and non-conservative synchronization models.

To find a network's community structure, we allow the network to evolve for some period of time according to the rules of a specific synchronization model. Next, we execute any of the standard clustering algorithms, such as the k-means algorithm or the agglomerative hierarchical clustering algorithm, to identify groups of similar nodes.

III. CASE STUDIES

We explore the differences between dynamics of conservative and non-conservative synchronization and the structures that emerge in two example networks: a synthetic network with a fixed hierarchical community structure and a benchmark social network. While the synthetic network does not have the statistical properties of naturally evolved real-world networks, we study it to demonstrate that running different processes on the same network leads to measurable differences in the discovered hierarchical structure.

A. Synthetic Network

The synthetic network, constructed following methods of [14], has $N = 256$ nodes evenly divided between four communities, with each community further sub-divided into four equal sub-communities. Each node randomly connects to $z_1 = 13$ nodes within its sub-community, $z_1 + z_2 = 17$ nodes within its community, and $z_{out} = 1$ nodes outside the community. Figure 1(a) shows the hinton diagram of the adjacency matrix of this network, with red entries indicating presence and blue absence of edges. Dense red blocks correspond to sub-communities, and sparse red blocks to second level communities. The spectra of L and R are shown in Figure 1(b). Each spectrum contains the eigenvalues of the operator, ranked in descending order, with the largest eigenvalue in the leading position. While we can already see differences in the spectra, they become more pronounced in real-world networks characterized by heterogeneous degree distribution.

We simulate synchronization in the synthetic network by letting nodes' phases evolve from some initial configuration. We ran 100 simulations of each synchronization model with initial values of θ_i drawn from a uniform random distribution $[-\pi, \pi]$, $\omega = 0$ and $\alpha = \lambda_{max}$ for non-conservative model. Figure 1(c) and (d) show the *similarity matrix* of the network after $t = 1500$ iterations of the two models. The matrix represents similarity, s_{ij}, of a pair of nodes, with color red corresponding to higher similarity values and blue to lower. The minimum similarity between any two nodes in the non-conservative system is 0.998, compared to 0.958 for the conservative system. The hierarchical community structure is visible in both similarity matrices.

To find the community structure, we execute a hierarchical agglomerative clustering algorithm on the similarity matrix at some time (here, after 1500 iterations). This procedure creates a dendrogram, which can be partitioned into four or 16 clusters. We use normalized mutual information MI to measure how well discovered clusters reproduce the actual communities [20]. When $MI = 1$, discovered clusters are the actual communities; while for $MI = 0$, they are independent of the actual communities. When we split each dendrogram into four clusters, we find $MI = 1.00$ in the conservative and $MI = 0.83$ in the non-conservative model. Splitting the dendrogram into 16 clusters leads to mutual information score of $MI = 0.66$ for the conservative model and $MI = 0.96$ for the non-conservative model. Non-conservative model appears to identify smaller structures faster and more accurately.

B. Karate Club Network

Next we study a real-world friendship network of members of a karate club [21], a widely studied social network benchmark shown in Fig. 2(a). During the course of the

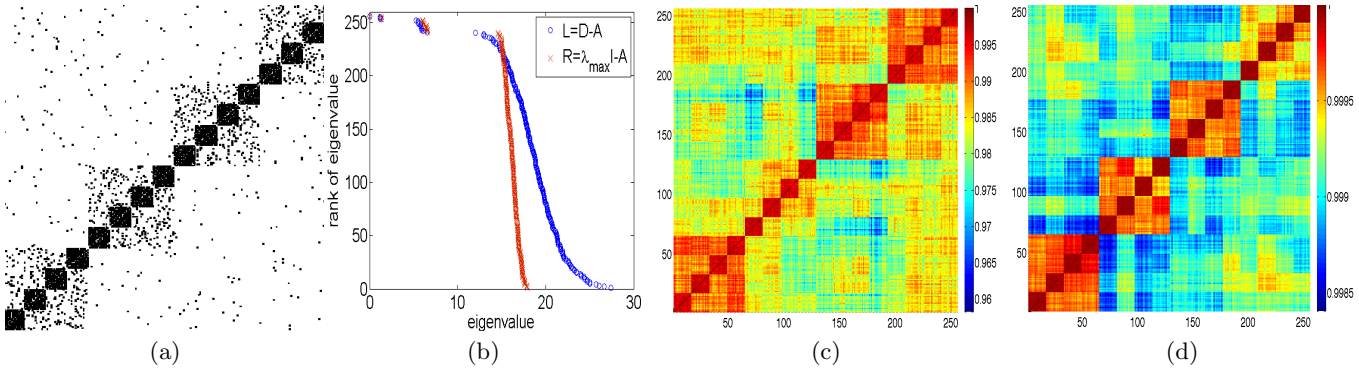

FIG. 1. Analysis of the synthetic network. (a) Hinton diagram of the adjacency matrix of the network. A point is red if an edge exists between nodes at that location; otherwise, it is blue. (b) Eigenvalue spectrum of the two conservative and non-conservative operators. Similarity matrix after $t = 1500$ iterations under the (c) conservative and (d) non-conservative synchronization models. Color indicates how similar two nodes are, with red corresponding to higher and blue to lower similarity.

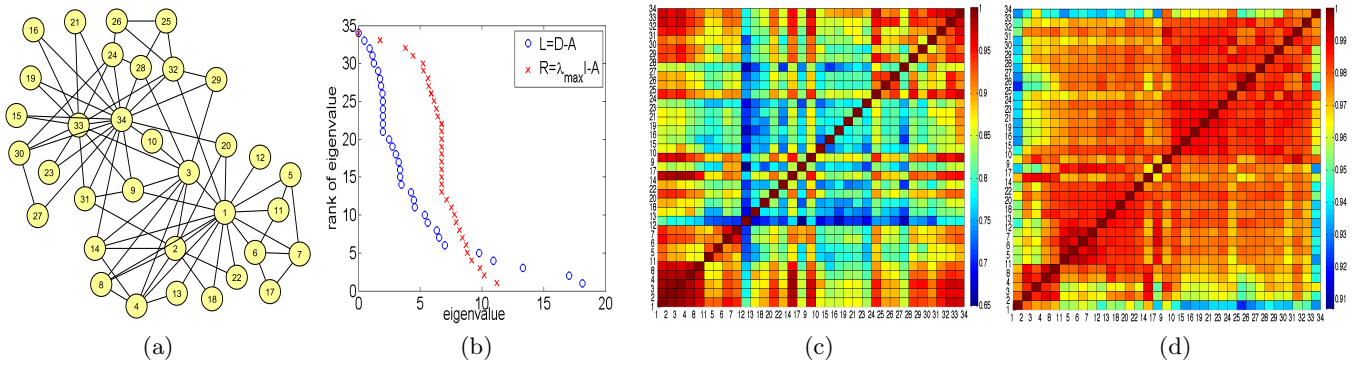

FIG. 2. Analysis of the karate club network. (a) Friendship graph of karate club members. (b) Comparison of eigenvalues of the Laplacian and Replicator operators. Similarity matrix after 1000 iterations of (c) the conservative and (d) non-conservative synchronization models. Color indicates similarity of node pairs, with red corresponding to higher and blue to lower similarity.

TABLE I. Normalized mutual information scores of the division of the karate club network into two groups under different synchronization models.

iterations	*10*	*1000*	*3000*	*3899*
conservative	0.046	0.046	0.046	0.733
non-conservative	0.046	0.301	1.000	1.000

study, a disagreement between the club's administrator and instructor resulted in the division of the club into two factions, represented by circles and squares, which we take as the actual communities for this network. There are greater differences between the spectra of L and R, shown in Fig. 2(b), than for the synthetic graph. The smallest positive eigenvalue of R is larger than that of L, implying that the non-conservative model synchronizes faster than the conservative model. Figure 2(c) and (d) shows similarity matrices after $t = 1000$ iterations of the two synchronization models. Minimum similarity in the conservative model is 0.65, while in the non-conservative model it is a much higher 0.91. After the same number of iterations, nodes are more synchronized in the non-conservative model.

We use hierarchical agglomerative algorithm to cluster nodes at different times based on synchronization similarity. Both models reveal rich hierarchical structure, though the two clusterings are very different. In the conservative model, high degree nodes (hubs) are deeper within the hierarchy, meaning they are more synchronized: 33 and 34 in one community and 1, 3 and 2 in the other community. Peripheral nodes (13, 18, 22) synchronize later, although nodes 15 and 10 never synchronize with their actual community and are mis-assigned. In the non-conservative model, peripheral nodes synchronize first, while the hubs synchronize last. Bridging nodes connected to both communities synchronize earlier in the conservative model than the non-conservative model and remain more synchronized. Table I reports MI scores of communities discovered by the two synchronization models at different times. Non-conservative model identifies communities faster than the conservative model, and the discovered communities are purer. Under both models nodes' community membership does not change after 3899 iterations, but similarity continues to grow until every node is equally similar to every other node.

IV. DISCUSSION

Our view of a network's structure depends not only on how the nodes are connected but also on how they interact. Interactions affect how we identify central nodes in the network [22, 23], and, as we show in this paper, also how we partition it into communities. We have demonstrated this by showing that in a network of coupled oscillators, different interactions lead different groups, or communities, of nodes to become synchronized over time. In practical terms this means that to find meaningful communities in real-world networks, community detection algorithms have to take into account the nature of interactions between nodes.

Note that our operational definition of a community, as a group of nodes that synchronize faster with each other than with other nodes, is different from the traditional definition of a community as more densely connected group of nodes. In fact, these two definitions are related, at least from the point of view of conservative processes, since the community structure found by conservative models is also one with low conductance. Conductance, however, may not be a good measure of group cohesiveness when nodes are interacting by broadcasting information or infecting each other with a pathogen.

Different interactions lead to distinct operators that govern dynamics of synchronization, each with its own spectral properties and characteristic topological and temporal scales. Investigating these relationships is the focus of future work.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 0915678, the MURI grant No. 1295-G-NA276 from the Air Force Office of Scientific Research, by the Air Force Research Laboratory (AFRL) under Contract Number FA8750-11-C-0127.

[1] E. Ravasz, A. L. Somera, D. A. Mongru, Z. N. Oltvai, and A. L. Barabási, Science **297**, 1551 (2002).
[2] A. W. Rives and T. Galitski, Proc. Natl. Acad. Sci. USA **100**, 1128 (2003).
[3] L. Freeman, pp. 39–97 (2003).
[4] F. R. K. Chung, *Spectral Graph Theory (CBMS Regional Conference Series in Mathematics, No. 92)* (American Mathematical Society, 1997).
[5] M. Rosvall and C. T. Bergstrom, Proc. Natl. Acad. Sci. USA **105**, 1118 (2008).
[6] S. Fortunato, Physics Reports **486**, 75 (2010).
[7] S. H. Strogatz, Nature **410**, 268 (2001).
[8] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).
[9] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, IEEE Symposium on Reliable Distributed Systems **0**, 25+ (2003).
[10] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. Hwang, Physics Reports **424**, 175 (2006).
[11] D. Centola, Science **329**, 1194 (2010).
[12] Y. Kuramoto, *Chemical Oscillations, Waves, and Turbulence* (Dover, Mineola, NY, 2003).
[13] T. Nishikawa, A. E. Motter, Y.-C. Lai, and F. C. Hoppensteadt, Phys. Rev. Lett. **91**, 014101 (2003).
[14] A. Arenas, A. Diaz-Guilera, and C. J. Pérez Vicente, Phys. Rev. Lett. **96**, 114102+ (2006).
[15] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, Physics Reports **469**, 93 (2008).
[16] Y. Hu, M. Li, P. Zhang, Y. Fan, and Z. Di, Phys. Rev. E **78**, 016115 (2008).
[17] R. Lambiotte, R. Sinatra, J. C. Delvenne, T. S. Evans, M. Barahona, and V. Latora (2010), 1012.1211.
[18] M. Jerrum and A. Sinclair, in *Proc. 20th ACM Symposium on Theory of Computing (STOC)* (ACM, New York, NY, USA, 1988), STOC '88, pp. 235–244.
[19] S. Chauhan, M. Girvan, and E. Ott, Phys. Rev. E **80**, 056114+ (2009).
[20] L. Danon, A. Daz-Guilera, J. Duch, and A. Arenas, J. Statistical Mechanics: Theory and Experiment **2005**, P09008 (2005).
[21] W. W. Zachary, J. Anthropological Research **33**, 452 (1977).
[22] S. Borgatti, Social Networks **27**, 55 (2005).
[23] R. Ghosh, K. Lerman, T. Surachawala, K. Voevodski, and S.-H. Teng (2011), 1102.4639.