Complex Networks: Information Theory and Collective Dynamics

Jean P. Martins, Danilo V. Vargas

{jean,dvargas}@icmc.usp.br

University of São Paulo
Institute of Mathematics and Computer Sciences,
São Carlos, 13566-590, SP, Brazil

Abstract

In this report we briefly discuss about two aspects related to the understanding of complex networks' structure and dynamics. First of all, is shown how we can use an Information theory approach to describe networks' characteristics such as heterogeneity and assortativeness; in addition, some network models like "Random Graphs" and "Lattice Networks" are classified following that approach. After that, the collective dynamics will be presented with a focus on synchronization of identical, non-identical and chaotic oscillators.

Contents

1 Fundamentals

In this report we briefly discuss about two aspects related to complex networks' structure and dynamics. First of all, it is shown how we can use an Information theory approach to describe networks' characteristics as heterogeneity and assortativeness; in addition, some network models like "Random Graphs" and "Lattice Networks" are classified following that approach. Also, we are going to discuss the collective dynamics with a focus on synchronization of identical, non-identical and chaotic oscillators together with some background in chaotic systems.

Therefore, some important fundamentals are summarized in this section. We start with a short description about complex networks, after that the concepts of Information Theory used in Section 2 are described followed by a preview about the collective dynamics contents discussed in Section 3.

1.1 Complex Networks

Complex networks describe a wide variety of systems, from the chemicals reactions within cells to social networks and the Internet. The complex networks perspective aim an general explanation for several questions about the structure and the dynamics of these networks: "What makes common structures emerge in so different contexts?", "How self-organization emerges from uncoordinated interactions?".

These and other questions have been studied by a statistical mechanics approach, so far the most successfully option since that complex networks became a well studied topic. In this section we review the main concepts used along the text, ranging from the degree probability distribution,node degree correlations until preferential attachment on evolving networks.

Random graphs were first studied by the hungarian Paul Erdös and Alfred Rényi, according to their models, every pair of nodes has a probability p of being connected, leading to a graph with approximately $pN(N-1)/2$ edges distributed randomly [1]. Although that was the beginning of studies in complex networks, today is known real networks deviates from a random graph, consequently, to understand the topology of the interactions between nodes and components in a network is important topic.

An important aspect concerning complex network is the *degree distribution*. Not every node has the same number of edges (node degree), in this case, the diversity of node's degrees can be characterized by a distribution function P_k, which gives the probability that a random selected node has exactly k edges. One of the most important discoveries of this field was that most real large networks don't have a Poisson distribution degree distribution, what means that they diverge from a random graph, it is rather common in these cases degree distributions following a power-law such $P_k \sim k^{-\gamma}$, these networks are called scale-free.

The scale-free networks contain a low number of high-degree nodes, and a high number of low-degree nodes. This characteristic is consequence of the dynamical process that creates scale-free networks, which states that along the time in real networks, high-degree nodes tend to receive more new connections than low-degree ones.

1.2 Information Theory

The Information Theory [2] was initially developed as a subfield of communication theory to answer questions like "What is the ultimate data compression?" and "What is the ultimate transmission rate of communication?", however, along the years it has made fundamental contributions in much more diverse areas such statistical physics, computer science, statistical inference and probability and statistics [3]. Most of the questions that emerged by that time were explained by concepts related to entropy and mutual information, in this section we briefly discuss these and another concepts required for better understand the subject of Section 2.

First of all, we define entropy as a measure of the average uncertainty in a random variable X with a probability mass function of $p(x)$. In other words, the entropy is the average number of bits (assuming logarithms to base 2) required to describe a random variable.

Definition 1 (Entropy [3]) *Let X be a discrete random variable with alphabet ξ and probability mass function $p(X = x)$, with $x \in \xi$. The entropy $H(X)$ is defined by:*

$$H(X) = -\sum_{x \in \xi} p(x) \log p(x) \tag{1}$$

It is interesting to observe that $H(X) = 1$ (its maximum value) when $p(x)$ equal $1/2$, i.e. maximum uncertainty. Also noteworthy, $H(X) = 0$ when $p(x) = 0$ or $p(x) = 1$, what makes sense because in this cases the variable is not random at all, and so there is not uncertainty.

As a result of entropy's definition we can define mutual information as a measure of the amount of information that one random variable contains about another random variable. In other words, mutual information is the reduction of uncertainty of one random variable due the information about another random variable.

Definition 2 (Mutual Information [3]) *Consider two random variables X and Y with a joint probability mass function $p(X = x, Y = y)$, and marginal probability mass functions $p(X = x)$ and $p(Y = y)$. The mutual information $I(X, Y)$ is defined by:*

$$I(X, Y) = \sum_{x \in \xi} \sum_{y \in \xi} p(x)p(y) \log \frac{p(x, y)}{p(x)p(y)} \tag{2}$$

Information theory is vast area, despite of that, Entropy and Mutual Information are sufficient for the context discussed at Section 2, hence we stop here, a more profound discussion about the topic can be found in [3] and [4].

1.3 Collective Dynamics

In the real world, everything is dynamic. And the interaction of multiple dynamic units is very complex. To give a measure of complexity, the orbit of two masses interacting through gravity is known to be analytically solvable. In other words, it is possible to predict the behavior of those masses by solving the differential equations.

If we add a third body to the previous system, however, the system is not analytical solvable anymore, but chaotic. In fact, the orbit is so complex that even if the system is simplified to be a problem with two masses much bigger than the third mass. This simplified system is still not analytically solvable and presents chaotic dynamics caused by homoclinic crossings.

Therefore, if we wish to understand the behavior of collective network dynamics a different view should be used. The understanding of chaos is very important, specially because when many objects are interacting the system is almost certainly chaotic. Instead of trying to explicitely predict systems, which can be most of the times unattainable in collective dynamics. The approach have to be more qualitative, that is, using characteristics of the system such as its Lyapunov exponents.

The literature of collective dynamics is vast as well as the background to understand it. In this essay, it will be presented briefly the most important background aspects to understand it and an overview of the area, specially in synchronous systems.

2 Information Theory for Complex Networks

Several correlation measures have been defined in order to characterize the structure of complex networks. Here we show how mutual information and entropy can be used to define some characteristics of networks on static graphs. This chapter describes and expands the analytical analysis described on the paper [5] for some simple examples. The original paper purposes that real networks are clustered in a well-defined domain of the entropy-noise space.

First of all we need to define some entropy-based measures that could be used to quantify the heterogeneity of a network by looking at the degree distribution (P_k). The degree distribution gives us the probability of a randomly selected node to have degree k. Furthermore, several statistical properties such as: average degree, clustering and diameter were defined to try characterize the properties of networks, most of them can be obtained from the adjacency matrix. In this text, we will focus three important characteristics to qualitatively describe complex networks: Randomness, heterogeneity and modularity (Figure 1); therefore, we show how these characteristics can be derived from Information Theory point of view.

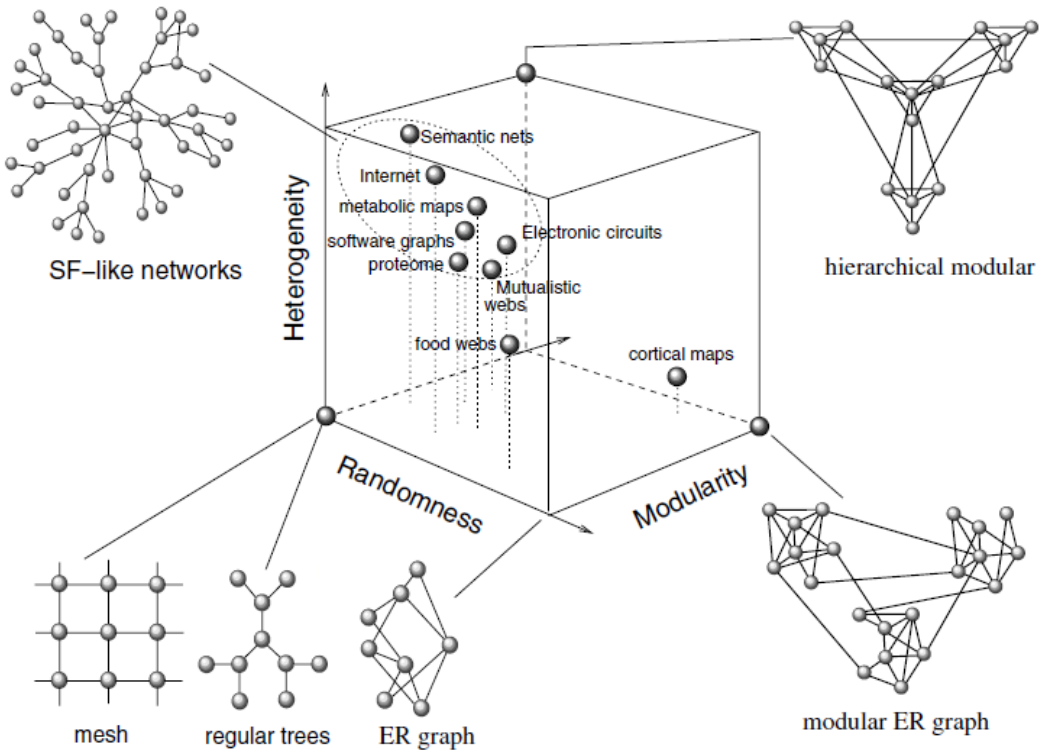

Figure 1: In this qualitative space, three relevant characteristics are included: randomness, heterogeneity and modularity. Image originally published in [5].

2.1 Measuring Correlations

Beyond the degree distribution and average statistical measures, correlation measures offer considerable insight into the structural properties displayed by complex networks. One particularly interesting is the Asortative Mixing (AM) [6], which implies that high degree vertices tend to be attached to other high-degree vertices. At the other extreme there are graphs displaying Dissortative Mixing (DM), or anti-correlation.

To identify AM or DM we will not focus in the degree distribution P_k but instead in the remaining degree $q(k)$ which refers to the number of vertices leaving the vertex other than the one we arrived along, the Figure 2 better illustrates the remaining degree, and the Equation (3) describes how that value can be calculated, where $\langle k \rangle = \sum_k kP_k$.

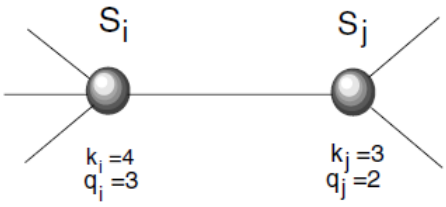

Figure 2: Here two given, connected nodes s_i, s_j, with different degrees k_i, k_j. The remaining degrees is indicated as q_i, q_j. Image originally published in [5].

$$q(k) = \frac{(k+1)P_{k+1}}{\langle k \rangle}, \tag{3}$$

In addition to $q(k)$ there is $q(k, k')$ which gives the probability of observing a vertex with k degree linked to another vertex of k' degree, this value is related to conditional probabilities $\pi(k|k')$ which measures the probability of observing a vertex with k edges leaving it provided that the vertex at the other end of the chosen edge has k' leaving edges. The Equation (4) summarizes these definitions, and represents the diversity of linked pairs with given remaining degrees:

$$q(k, k') = \frac{\pi(k|k')}{q(k')} \tag{4}$$

The definition of information within the context of communication channels implies additional statistical quantities (such as channel entropy and noise) that provide a detailed characterization of system's complexity.

2.2 Entropy and Information

By using the previous distribution $Q = (q(1), \ldots, q(i), \ldots, q(n))$, an entropy measure $H(q)$ can be defined, in this context, it provides an average measure of network's heterogeneity, since it measures the diversity of the link distribution, see Equation (5).

$$H(Q) = -\sum_{k=1}^{n} q(k) \log q(k). \tag{5}$$

The maximum entropy ($H_{\max}(Q) = \log n$) is obtained when $q(k) = 1/n (\forall k = 1, \ldots, n)$, what means that for each $k \in \{1, \ldots, n\}$ there is one node in the network. The minimum entropy ($H_{\min}(Q) = 0$) instead occurs when $Q = (1, 0, \ldots, 0)$, i.e. all nodes have the same remaining degree. Similarly the joint entropy can be defined by Equation (6).

$$H(Q, Q') = -\sum_{k=1}^{n} \sum_{k'=1}^{n} q(k, k') \log q(k, k'). \tag{6}$$

By now, we can also define the conditional entropy (or noise) using conditional probabilities $\pi(k, k')$, see Equation (7).

$$\begin{aligned} H(Q|Q') &= -\sum_{k=1}^{n} \sum_{k'=1}^{n} q(k, k') \pi(k, k') \log \pi(k, k') \\ &= -\sum_{k=1}^{n} \sum_{k'=1}^{n} q(k, k') \log \frac{q(k, k')}{q(k')} \end{aligned} \tag{7}$$

Consequently we can define Mutual Information as the difference between entropy and conditional entropy which can be extended to Equation (8), where $\pi(k|k')$ is the

probability of observing a vertex with k edges leaving it provided that the vertex at the other end of the chosen edge has k' leaving edges:

$$\begin{aligned} I(Q) &= H(Q) - H(Q|Q') \\ &= -\sum_{k=1}^{n} q(k) \log q(k) + \sum_{k=1}^{n}\sum_{k'=1}^{n} q(k,k') \log \frac{q(k,k')}{q(k')} \\ &= -\sum_{k=1}^{n}\sum_{k'=1}^{n} q(k,k') \log \frac{q(k)}{\pi(k|k')} \\ &= \sum_{k=1}^{n}\sum_{k'=1}^{n} q(k,k') \log \frac{q(k,k')}{q(k)q(k')} \end{aligned} \tag{8}$$

In this case, the maximum information transfer, for a given $\{q(k)\}$, is obtained when $H(Q|Q) = 0$, i. e. when the conditional probabilities are such that $\pi(k|k') = 1$ or 0 for all $k, k = 1, \ldots, n$. Another interesting case is given by $\pi(k|k') = \delta_{k,k'}$, which implies that $q(k,k') = q(k)\delta_{k,k'}$ which is precisely the case of perfectly assortative network.

By using the measures previously described we can characterize: (a) the amount of correlation between nodes, measured by the information; (b) the noise level, measured by the conditional entropy (also provides a measure of assortativeness) and (c) the entropy of the $q(k)$ distribution. In the next section we will demonstrate the use of these aforementioned functions, for complex networks characterization.

2.3 Model Networks

2.3.1 Lattices Networks

Lattices networks represent the highest degree of homogeneity,having no randomness; this type of networks is common in parallel computers. The Figure 3 illustrates a lattice-like network, in this case, all vertices have the same degree z, hence, considering the Kronecker's delta function, in Equation (9), we have $P_k = \delta_{k,z}$.

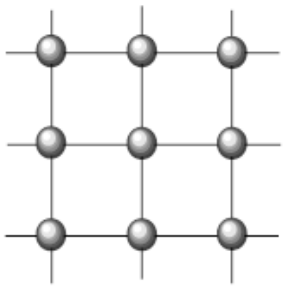

Figure 3: Lattice network where every node has $z = 4$ degree. Image originally published in [5].

$$\delta_{k,z} = \begin{cases} 1 & \text{if } k = z \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

Furthermore we can simplify the remaining degree (Equation 3) considering peculiarities of a lattice like networks, as a result:

$$\begin{aligned} q(k) &= \frac{(k+1)\ P_{k+1}}{\langle k \rangle} \\ &= \frac{(k+1)\ P_{k+1}}{\sum_{k=1}^{n}\ k\ P_k} \\ &= \frac{(k+1)\ \delta_{k,(z-1)}}{k\ \delta_{k,(z-1)}} \\ &= \frac{k\ \delta_{k,(z-1)}\ + \delta_{k,(z-1)}}{k\ \delta_{k,(z-1)}} \\ q(k) &= \delta_{k,(z-1)} \end{aligned} \tag{10}$$

In a lattice network there is not assortativeness taking effect, that is, every node can be seen as an independent random variable, what leads us to define Equation (11)[1]:

$$\begin{aligned} q(k,k') &= q(k)\ q(k') \\ &= \delta_{k,(z-1)}\ \delta_{k',(z-1)} \end{aligned} \tag{11}$$

Now, we can calculate through Equation (8) the amount of correlation between nodes in a lattice-like graph, as $q(k)\ q(k') = q(k,k')$ and $\log(1) = 0$ this value is $I(Q) = 0$:

$$\begin{aligned} I(Q) &= \sum_{k=1}^{n}\sum_{k'=1}^{n} q(k,k') \log \frac{q(k,k')}{q(k)q(k')} \\ &= \sum_{k=1}^{n}\sum_{k'=1}^{n} \delta_{k,(z-1)}\ \delta_{k',(z-1)} \log \frac{\delta_{k,(z-1)}\ \delta_{k',(z-1)}}{\delta_{k,(z-1)}\delta_{k',(z-1)}} \\ I(Q) &= 0 \end{aligned} \tag{12}$$

2.3.2 Random Graphs

Random graphs (or Erdös-Renyi graphs) are built joining every two nodes with some probability p, consequently their average degree is $\langle k \rangle \approx pn$, where n is the total number of nodes. This random building process implies a randomly chosen degree for each node, consequently, the degree distribution can be represented by a Poisson distribution [7], defined by the Equation (13), where x is a random variable with expected number of occurrences per unit equal to m, for random graphs that equation leads us to Equation (14).

$$P(x) = \frac{e^{-m}m^x}{x!} \tag{13}$$

$$P_k = \frac{e^{-\langle k \rangle}\langle k \rangle^k}{k!} \tag{14}$$

[1] If there was assortative mixing $q(k,k')$ would differ from the result of Equation (11).

The remaining degree equation can be obtained similarly, as we show next, its value is equal to P_k.

$$\begin{aligned} q(k) &= \frac{(k+1)\ P_{k+1}}{\langle k \rangle} \\ &= \frac{(k+1)}{\langle k \rangle} \frac{e^{-\langle k \rangle} \langle k \rangle^{(k+1)}}{(k+1)!} \\ &= \frac{e^{-\langle k \rangle}}{\langle k \rangle} \frac{\langle k \rangle^{(k+1)}}{k!} \\ q(k) &= \frac{e^{-\langle k \rangle} \langle k \rangle^{k}}{k!} \quad = \quad P_k \end{aligned} \tag{15}$$

The network is built by a random process, therefore, there is no intense assortativeness, i.e. $q(k, k') \approx q(k)q(k')$ for a large n, thus, the information transfer function should lead to a near zero value.

$$\begin{aligned} I(Q) &\approx \sum_{k=1}^{n} \sum_{k'=1}^{n} q(k)q(k') \log \frac{q(k)q(k')}{q(k)q(k')} \\ I(Q) &\approx 0 \end{aligned} \tag{16}$$

2.3.3 Star graph

A star graph defines an extreme situation which one node has degree $n-1$ and all the others have degree 1. This structure is commonly found in scale-free networks, and represents a strong evidence of assortativeness. The degree distribution in a star graph can be described by Equation (17), and the respective remaining degree by Equation (18).

$$\begin{aligned} P_k &= \frac{n-1}{n}\delta_{k,1} + \frac{1}{n}\delta_{k,(n-1)} \\ P_k &= \frac{1}{n}\left[(n-1)\delta_{k,1} + \delta_{k,(n-1)}\right] \end{aligned} \tag{17}$$

$$\begin{aligned} q(k) &= \frac{(k+1)}{\langle k \rangle}\ P_{k+1} \\ &= \frac{(k+1)}{\langle k \rangle n}\left[(n-1)\delta_{k,0} + \delta_{k,(n-2)}\right] \\ &= \frac{(k+1)n}{2(n-1)n}\left[(n-1)\delta_{k,0} + \delta_{k,(n-2)}\right] \\ &= \frac{(k+1)(n-1)\delta_{k,0} + (k+1)\delta_{k,(n-2)}}{2(n-1)} \\ &= \frac{(k+1)}{2}\delta_{k,0} + \frac{(k+1)}{2(n-1)}\delta_{k,(n-2)} \\ q(k) &= \frac{1}{2}(\delta_{k,0} + \delta_{k,(n-2)}) \end{aligned} \tag{18}$$

The joint probabilities of remaining degrees can, so, be deduced from Equation (4):

$$\begin{aligned} q(k,k') &= q(k')\pi(k|k') \\ &= \frac{1}{2}\left[\delta_{k,0} + \delta_{k,(n-2)}\right]\left[\delta_{k',0} + \delta_{k',(n-2)}\right] \\ &= \frac{1}{2}\delta_{k,0}\delta_{k',0} + \frac{1}{2}\delta_{k,0}\delta_{k',(n-2)}\frac{1}{2}\delta_{k,(n-2)}\delta_{k',0} + \frac{1}{2}\delta_{k,(n-2)}\delta_{k',(n-2)} \\ &= 0 + \frac{1}{2}\delta_{k,0}\delta_{k',(n-2)} + \frac{1}{2}\delta_{k,(n-2)}\delta_{k',0} + 0 \\ q(k,k') &= \frac{1}{2}(\delta_{k,0}\delta_{k',(n-2)} + \delta_{k,(n-2)}\delta_{k',0}) \end{aligned} \tag{19}$$

By these results we can calculate the mutual information using Equation (8):

$$\begin{aligned} I(Q) &= \sum_{k=1}^{n}\sum_{k'=1}^{n} q(k,k')\log\frac{q(k,k')}{q(k)q(k')} \\ &= \sum_{k=1}^{n}\sum_{k'=1}^{n} \frac{1}{2}(\delta_{k,0}\delta_{k',(n-2)} + \delta_{k,(n-2)}\delta_{k',0})\log\left(\frac{\frac{1}{2}(\delta_{k,0}\delta_{k',(n-2)} + \delta_{k,(n-2)}\delta_{k',0})}{\frac{1}{4}(\delta_{k,0} + \delta_{k,(n-2)})(\delta_{k',0} + \delta_{k',(n-2)})}\right) \\ &= 2\left[\frac{1}{2}\log\left(\frac{\frac{1}{2}(\delta_{k,0}\delta_{k',(n-2)} + \delta_{k,(n-2)}\delta_{k',0})}{\frac{1}{4}(\delta_{k,0} + \delta_{k,(n-2)})(\delta_{k',0} + \delta_{k',(n-2)})}\right)\right] \\ &= 2\left[\frac{1}{2}\log\left(\frac{2\delta_{k,0}\delta_{k',(n-2)}}{\delta_{k,0} + \delta_{k,(n-2)}}\right)\right] \\ I(Q) &= \log 2 = 1 \end{aligned} \tag{20}$$

2.4 Summary

In this section we discussed about the utilization of an Information-theoretic approach to measure some complex networks' characteristics such heterogeneity and correlation. Through the approach described network models were classified, including Erdös-Renyi random graphs, Lattice networks and Star graphs. These examples aimed make clear the power of Information Theory for describe complex networks, being a very interesting option when there is nonlinear relationships between nodes.

3 Collective Dynamics

Networks of dynamical systems have been used to model almost everything. Examples are present in physics (neutrino synchronization) [8], chemistry (chemical reactions) [9] and earth sciences(geophysics and geology) [10].

However, little can be said of the collective dynamics of such systems when they are coupled together [11]. Collective dynamics can be quite challenging, since they span a big number of disciplines from dynamical systems, statistical learning to neuroscience. Here the focus will be on synchronization with some brief explanations of learning networks and other collective dynamic systems.

3.1 Oscillators Synchronization and Other Behaviors

The most common studied synchronous behavior is the synchronization of oscillators. They are present in the brain, clock and other systems as well. Although common, there are other types of synchronization such as the spike synchronization, which is commonly found in the brain [12].

Here we will analyse systems with identical oscillators, non-identical oscillators and chaotic oscillators. And build a theoretical framework for the understanding and prediction of behavior of general synchronization.

3.1.1 Identical Oscillators

Identical oscillators are found in many biological systems like the waves of the brain, heart, intestine and nervous system [13]. When the set of identical oscillators are coupled by smooth interactions. They often synchronize or form patterns dependent of the network symmetry [14]. On the other hand, when oscillators communicate by sudden impulses as commonly found in biology such as neuron spikes. It was proved that N identical integrate-and-fire oscillators (IFO) connected in a all-to-all network will end up firing in unison independent of their initial state [15], [16]. With such simple models it is yet possible to observe systems which result in synchronization or self-organized criticality [17].

3.1.2 Self-Organized Criticality

The behavior of a system to self-organize around a critical point is called self-organized criticality (SOC) [18],[19]. The self-organize capacity is defined as a inner dynamic property of the system. Which independent of its parameters or interferences would drive the system toward a given state. Critical point has a varied number of definitions depending on the context it is applied. In mathematics it is the point where either the derivative is 0 or it is non differentiable [20]. In physics it is where a phase boundary (such as the vapor-liquid point) ceases to exist [21]. Sometimes, a broad definition definition is used defining a point where a system properties change suddenly to be a critical point. This concept was developed in 1987 but is still view with skepticism due to its ubiquitous application.

3.1.3 Non-identical Oscillators

In [9] almost identical limit-cycle oscillators were shown to exhibit completely desynchronized behavior when weakly couped. And a mixture of synchronized behavior and desynchronized behavior when strongly coupled.

3.1.4 Kuramoto Model

The Kuramoto model is composed of N oscillators in a complete graph (all-to-all connections, illustrated in Figure 4) and each one has the following dynamics [9]:

$$\frac{\partial \theta_i}{\partial t} = \omega_i + \frac{K}{N}\sum_{j=1}^{N}\sin(\theta_j - \theta_i), \qquad i = 1 \dots N, \tag{21}$$

where ω_i is the natural frequency of the ith oscillator, θ_i is its phase and K is the coupling strength (identical for all edges). The ω_i were drawn from a Lorentzian distribution defined by:

$$f(\omega;\omega_0,\gamma) = \frac{1}{\pi}\left[\frac{\gamma}{(\omega-\omega_0)^2+\gamma^2}\right], \tag{22}$$

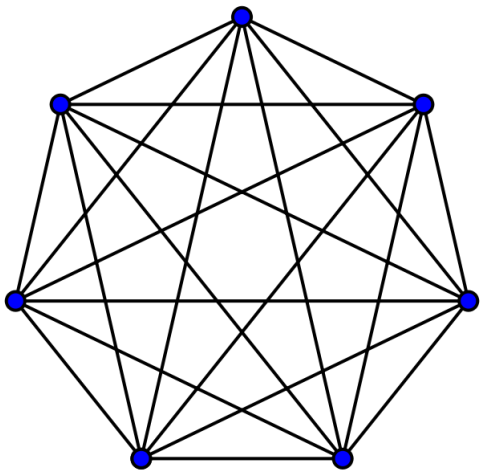

Figure 4: Diagram of a complete graph illustrating the scenario where the Kuramoto model is applied.

To solve this model analytically when $N \to \infty$, Kuramoto applied the following transformation:

$$re^{i\psi} = \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j}, \tag{23}$$

where ψ is the average phase and r measures the coherence of the oscillators. Resulting in the following equation:

$$\frac{\partial\theta_i}{\partial t} = \omega_i + Kr\sin(\psi-\theta_i) \tag{24}$$

In the limit (as $N \to \infty$ and $t \to \infty$), Kuramoto found that the following behavior occurs:

$$r = \begin{cases} 0, & K < K_c \\ \sqrt{1-\frac{K_c}{K}}, & K \geq K_c \end{cases}, \tag{25}$$

where $K_c = 2\gamma$. Therefore, when $K < K_c$ the oscillators are not synchronized. On the other hand when $K \geq K_c$ part of the oscillators synchronize, or more specifically lock their phases ($\frac{\partial\theta_i}{\partial t} = 0$) and part are rotating out of synchrony. When the coupling increases to the limit as $K \to \infty$, the oscillators become totally synchronized approximately to their average phase $\theta_i \approx \psi$ at the same time that $r \to 1$.

Extensions to the Kuramoto model are still in research [22], for the complete proof or more information on variations of the Kuramoto model, such as general frequency distributions, refer to [23].

3.2 Stability of Synchronized states

Frequently, the master stability function and Lyapunov exponents are used to infer the stability of the synchronization in a number of systems.Therefore, the following subsections will briefly define the topics and give sufficient information for understanding the later analysis.

3.2.1 Master Stability Function

The master stability function (MSF) is a framework developed to study the synchronization state complex of network topologies independent of the peculiarities of the oscillators. It is based on the premise that all nodes are identical dynamical units, described by the following equation:

$$\frac{\partial x_i}{\partial t} = F(x_i), \tag{26}$$

where x_i is a m-dimensional vector and $F(x_i)$ is its evolution function (which is the same for every node). The output of the system is described by $H(x_i)$, which is also identical for all N dynamical units and is a coupling map of the nodes. For example, in y-coupled oscillators $H(x_i) = (0, y, 0)$, where they are only communicating with each other through the y component. The complex network is defined by the node's adjacencies a_{ij} and their respective weights w_{ij} (for unweighted matrix consider $w_{ij} = 1$) expressed by:

$$\begin{aligned} \frac{\partial x_i}{\partial t} &= F(x_i) + \sigma \sum_{i=1}^{N} a_{ij} w_{ij} (H(x_j) - H(x_i)) \\ &= F(x_i) + -\sigma \sum_{j=1}^{N} G_{ij} (H(x_j)), \end{aligned} \tag{27}$$

where σ is the uniform coupling strength ($\sigma > 0$ for diffusive coupling) and G is a coupling matrix defined by:

$$G_{ij} = \begin{cases} -a_{ij} w_{ij} & i \neq j \\ \sum_{j=1}^{N} a_{ij} w_{ij} & i = j \end{cases}, \tag{28}$$

it follows that G has zero row-sum. Then the synchronization occurs when for all nodes the coupling term vanishes. When this occurs, the nodes will be based solely on their internal dynamics which happen to be the same for all N nodes. Resulting in a state called synchronization manifold, which happens when:

$$x_1(t) = x_2(t) = ... = x_N(t) = x_s(t). \tag{29}$$

The synchronization manifold occurs as a consequence of Equation 29 because the matrix is a zero row-sum and the function $H(x)$ is the same for any node. The stability of synchronization is secured when the system remains entirely inside the synchronization manifold. However, in the real world system are susceptible to perturbations, in this scenario the following equation holds:

$$\frac{\partial \tilde{x_i}}{\partial t} = [JF(x_s) - \sigma \lambda_i JH(x_s)] \tilde{x_i}, \tag{30}$$

where $\tilde{x_i}$ is the deviation from the $x_s(t)$ for $x_i(t)$, J is the Jacobian operator and λ is the eigenvalue of the respective m conditional Lyapunov exponents. Conditional Lyapunov exponents for each value of $\sigma\lambda_i$ can now be acquired.

3.2.2 Lyapunov exponent

The Lyapunov exponent is defined as the rate of separation of two orbits. It is a measure of sensitivity of a given dynamic system from its initial conditions, exactly calculated by:

$$\lambda = \lim_{t\to\infty} \lim_{\delta x_0 \to 0} \frac{1}{t} ln \frac{|\delta x(t)|}{|\delta x_0|}, \tag{31}$$

where δx_0 is the initial difference between the two orbits and $\delta x(t)$ is the difference between the two orbits in instant t. A dynamics system with m-dimensional phase space will have m exponents, forming the Lyapunov spectrum $\lambda_1, \lambda_2, ..., \lambda_m$. The largest value of the Lyapunov spectrum is known as the maximum Lyapunov exponent (MLE) or simply λ_{max}.

Usually, Lyapunov exponents are used in synchronization to verify the stability of a system under small perturbations of its orbit. To verify it, the behavior of λ_{max} is used. If $\lambda_{max} < 0$ the system is stable because all the remaining Lyapunov is also smaller than λ_{max} which is negative. On the other hand, if $\lambda_{max} \geq 0$ the system is said to be unstable, because one exponent is already positive and the system will have by definition an exponential divergence of orbits. It is important to note that although necessary the Lyapunov exponents do not offer a sufficient condition for the stability of the system. As it will be seen later in the unusual behavior of chaotic systems.

Extensions of the master stability function are still under research, recent results are the MSF near identical systems [24].

3.3 Unusual Behavior of Chaotic Oscillators

When analyzing chaotic systems, some interesting phenomena may appear and cause a desynchronization or characteristic behavior of the system. Here only the bursting and short-length bifurcation will be described.

3.3.1 Bursting

In some systems like the ones from [25] and [26] which are y-coupled Rössler systems. After reaching synchronous state above a coupling threshold (i.e., after the longest-wavelength mode becomes stable), they still present a difference between the average x and its component x shown in Figure 5. This difference is expected to be close to 0 in synchronized systems. To this effect was given the name of bursting and they may appear because of unstable periodic orbits (UPO).

3.3.2 Short-length Bifurcation and Size Effect

Although counterintuitive, there is a possible limit in the number of chaotic coupled oscillators which can produce synchrony. Coupled modified Rössler oscillators desynchronize with the increase in the coupling strength as shown in [27]. This is known as the short-wavelength bifurcation, because the higher modes go unstable first. In Figure 6 this is illustrated. This can happen in any system where each node has internal dynamics which are not directly coupled to other nodes. For example, this can happen in the biological cell scenario, where cells have internal chemical dynamics, but only communicate via

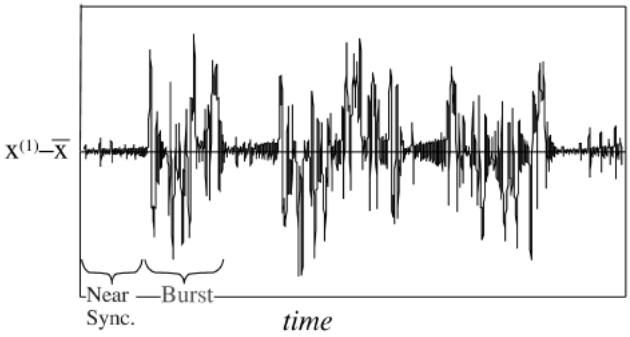

Figure 5: The difference from the average observed in the x variable of y-coupled Rössler systems. Image from [26].

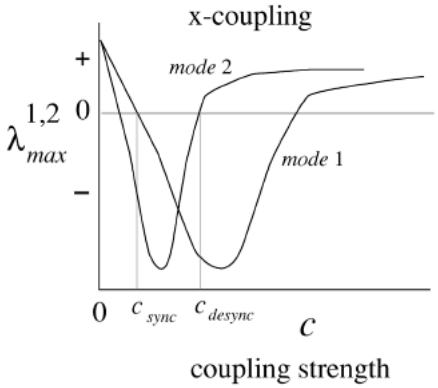

Figure 6: Stability diagram showing the short-length bifurcation of x-coupled Rössler systems. Image from [26].

hormones, voltages and other chemical exchanges. Another interesting behavior which contradicts the naive view that the Kuramoto model would be extensible to chaotic systems is the **size effect**. Which states that depending on the choice of oscillator and its respective variables a system may have a limit in the number of oscillators that can be coupled for it to achieve synchronous chaotic behavior [28]. This happens because the system does not have a coupling state where all modes are stable, since the short wavelength goes unstable when or before the longest wavelengths goes stable as shown in Figure 7.

3.4 Robustness of Scale-free Networks and its Synchronization Consequences

Random removal of nodes in a scale-free network will not affect greatly. Since there are exponentially more nodes with lower degree than higher degree nodes. However, scale-free networks are more susceptible to intentional attacks than random networks [29]. As noted by [30]:

> The bad news is that the inhomogeneous topology has its drawbacks as well:

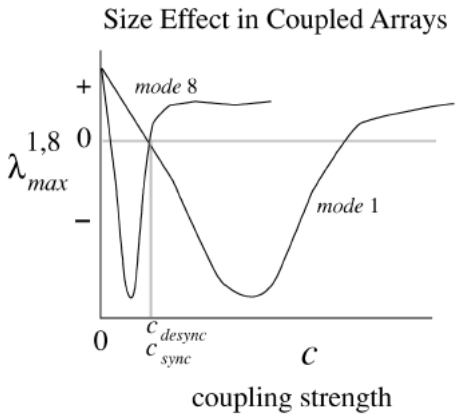

Figure 7: Stability diagram showing the size effect of 16 x-coupled Rössler systems. Image from [26].

> scale-free networks are rather vulnerable to attacks. Indeed, the absence of a tiny fraction of the most connected nodes will break the network into pieces. Calculations indicate that removing nodes in the order of their degree leads to the rather small critical threshold.

This robustness (or vulnerability) also affects the collective dynamics of scale-free networks, since synchronization inside scale-free networks can sustain random removal of 5% of its nodes with minor affects in behavior. While 1% of wisely selected nodes might destroy completely the synchronization [31], [32].

3.5 Synchronization of Robots, Vehicles and other Objects

Much of the last subsections can be extended easily to synchronization of particles, robots and other object related trajectories. Since, any object with an trajectory can be described in polar coordinates and then the same equations apply.

3.6 Network Learning Dynamics

Networks can be composed of memory and some interaction based on statistical learning [30]. And although recognized as learning systems, perceptrons and other learning machines are based on simple equations and memory. This equations are mostly based on rules for updating the weights and calculating the output of a network. Some of these models even derive from statistical physics [33], [34]. The nodes of such systems will be other networks or some other complex machine. For the sake of brevity, this section will not go into the detail or give any equation. It will give a rather overview of some different behavior as illustration of other possible collective behaviors.

3.6.1 Teacher and Student

The situation where a student node learn from a teacher node. Which is exactly the interaction defined in supervised learning [35]. The unique difference is that in the super-

vised learning setting the teacher is not specifically defined as a network node. Dynamics in this simple situation are well defined and studied, an example using two perceptrons is shown on Figure 8. Supposing the student has a multilayer-perceptron node, it is known

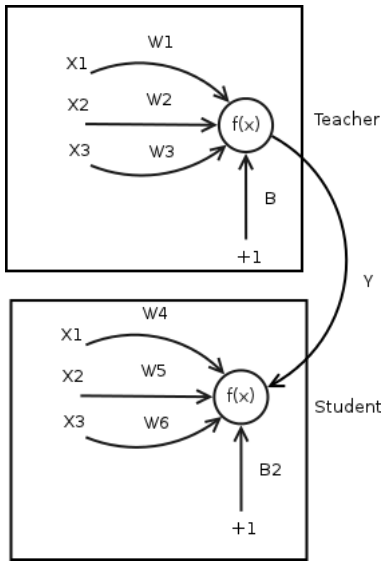

Figure 8: Example of dynamics between teacher and student. Where W variables are weights, X are input, B are the bias, $f(x)$ is the activation function and Y is the result from the teacher. The rectangles represent the nodes in this small network of two units.

that independent of the function present in the teacher node, the student will be able to learn it. Provided that the student has a sufficient number of hidden nodes using a back-propagation or equivalent learning algorithm [34].

3.6.2 Self-Interacting

Suppose now the situation where an already trained node is interacting with itself [30]. The node can be trained in any arbitrary sequence. This time, however, the node is learning the opposite of its own prediction. Which leads to a situation where its prediction error is 100% and the sequence produced by the node is close to random [36]. In the case of a Hebbian training with a Boolean perceptron in N dimensions, the bit sequence has an average length defined by Equation 32:

$$L \backsimeq 2.2^N. \tag{32}$$

When a second Boolean perceptron is added to predict the sequence, it has 78% of prediction error. Which is somewhat better than the self-interacting node, but still worse than a 50% random guessing [37]. The Figure 9 illustrates the example with two perceptrons, note that this example can be extended to other learning machines.

References

[1] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Rev. Mod. Phys.*, 74(1):47–97, Jan 2002.

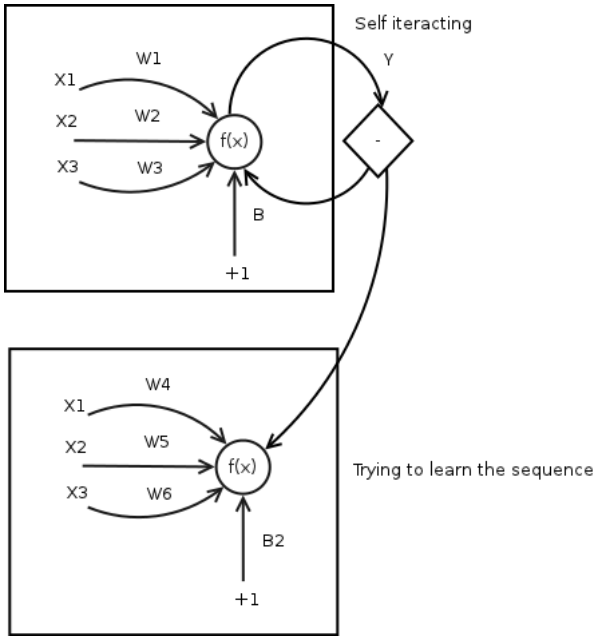

Figure 9: Example of dynamics a self teaching node and another perceptron trying to learn the output. Where W variables are weights, X are input, B are the bias, $f(x)$ is the activation function and Y is the result from the teacher. The rectangles isolate the nodes of the two units.

[2] C.E. Shannon and W. Weaver. The mathematical theory of information. 1949.

[3] T.M. Cover, J.A. Thomas, J. Wiley, et al. *Elements of information theory*, volume 306. Wiley Online Library, 1991.

[4] J.R. Pierce. *An introduction to information theory: symbols, signals & noise*. Dover Pubns, 1980.

[5] Ricard Solé and Sergi Valverde. Information theory of complex networks: On evolution and architectural constraints. In Eli Ben-Naim, Hans Frauenfelder, and Zoltan Toroczkai, editors, *Complex Networks*, volume 650 of *Lecture Notes in Physics*, pages 189–207. Springer Berlin / Heidelberg, 2004. 10.1007/978-3-540-44485-5_9.

[6] M. E. J. Newman. Assortative mixing in networks. *Phys. Rev. Lett.*, 89(20):208701, Oct 2002.

[7] D.R. Byrkit. *Statistics today: a comprehensive introduction*. Benjamin-Cummings Publishing Co., Inc., 1986.

[8] G.G. Raffelt and I. Tamborra. Synchronization versus decoherence of neutrino oscillations at intermediate densities. *Physical Review D*, 82(12):125004, 2010.

[9] Y. Kuramoto. *Chemical oscillations, waves, and turbulence*. Dover Pubns, 2003.

[10] D.L. Turcotte. *Fractals and chaos in geology and geophysics*. Cambridge Univ Pr, 1997.

[11] S.H. Strogatz. Exploring complex networks. *Nature*, 410(6825):268–276, 2001.

[12] M.N. Shadlen and J.A. Movshon. Synchrony unbound: Review a critical evaluation of the temporal binding hypothesis. *Neuron*, 24:67–77, 1999.

[13] A.T. Winfree. *The geometry of biological time*. Springer Verlag, 2001.

[14] JJ Collins and IN Stewart. Coupled nonlinear oscillators and the symmetries of animal gaits. *Journal of Nonlinear Science*, 3(1):349–392, 1993.

[15] C.S. Peskin. *Mathematical aspects of heart physiology*. Courant Institute of Mathematical Sciences, New York University, 1975.

[16] R.E. Mirollo and S.H. Strogatz. Synchronization of pulse-coupled biological oscillators. *SIAM Journal on Applied Mathematics*, 50(6):1645–1662, 1990.

[17] Á. Corral, C.J. Pérez, A. Diaz-Guilera, and A. Arenas. Self-organized criticality and synchronization in a lattice model of integrate-and-fire oscillators. *Physical review letters*, 74(1):118–121, 1995.

[18] P. Bak, C. Tang, and K. Wiesenfeld. Self-organized criticality: An explanation of the 1/f noise. *Physical Review Letters*, 59(4):381–384, 1987.

[19] D.L. Turcotte. Self-organized criticality. *Reports on progress in physics*, 62:1377, 1999.

[20] J. Stewart. *Calculus: early transcendentals*. Cengage Learning EMEA, 2008.

[21] Y.A. Cengel and M.A. Boles. *Thermodynamics: an engineering approach*, volume 988. McGraw-Hill Higher Education, 2006.

[22] EA Martens, E. Barreto, SH Strogatz, E. Ott, P. So, and TM Antonsen. Exact results for the kuramoto model with a bimodal frequency distribution. *Physical Review E*, 79(2):026204, 2009.

[23] J.A. Acebrón, L.L. Bonilla, C.J.P. Vicente, F. Ritort, and R. Spigler. The kuramoto model: A simple paradigm for synchronization phenomena. *Reviews of modern physics*, 77(1):137, 2005.

[24] J. Sun, E.M. Bollt, and T. Nishikawa. Master stability functions for coupled nearly identical dynamical systems. *EPL (Europhysics Letters)*, 85:60011, 2009.

[25] P. Ashwin, J. Buescu, and I. Stewart. Bubbling of attractors and synchronisation of chaotic oscillators. *Physics Letters A*, 193(2):126–139, 1994.

[26] L.M. Pecora, T.L. Carroll, G.A. Johnson, D.J. Mar, and J.F. Heagy. Fundamentals of synchronization in chaotic systems, concepts, and applications. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 7:520, 1997.

[27] J.F. Heagy, L.M. Pecora, and T.L. Carroll. Short wavelength bifurcations and size instabilities in coupled oscillator systems. *Physical review letters*, 74(21):4185–4188, 1995.

[28] L.M. Pecora. Synchronization conditions and desynchronizing patterns in coupled limit-cycle and chaotic systems. *Physical Review E*, 58(1):347, 1998.

[29] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin. Breakdown of the internet under intentional attack. *Physical Review Letters*, 86(16):3682–3685, 2001.

[30] S. Bornholdt, H.G. Schuster, and J. Wiley. *Handbook of graphs and networks*. Wiley Online Library, 2003.

[31] J. Lü, G. Chen, and D. Cheng. A new chaotic system and beyond: the generalized lorenz-like system. *Int. J. Bifurc. Chaos*, 14(5):1507–1537, 2004.

[32] X.F. Wang and G. Chen. Synchronization in scale-free dynamical networks: robustness and fragility. *Circuits and Systems I: Fundamental Theory and Applications, IEEE Transactions on*, 49(1):54–62, 2002.

[33] J. Hertz, A. Krogh, and R.G. Palmer. *Introduction to the theory of neural computation*. Westview press, 1991.

[34] A. Engel and C. Broeck. *Statistical mechanics of learning*. Cambridge Univ Pr, 2001.

[35] SB Kotsiantis, ID Zaharakis, and PE Pintelas. Supervised machine learning: A review of classification techniques. *Emerging artificial intelligence applications in computer engineering: real word AI systems with applications in eHealth, HCI, information retrieval and pervasive technologies*, 160:3, 2007.

[36] R. Metzler, W. Kinzel, L. Ein-Dor, and I. Kanter. Generation of unpredictable time series by a neural network. *Physical Review E*, 63(5):056126, 2001.

[37] H. Zhu and W. Kinzel. Antipredictable sequences: harder to predict than random sequences. *Neural computation*, 10(8):2219–2230, 1998.