The relationship between structure and function in locally observed complex networks

2013 New J. Phys. 15 013048

(http://iopscience.iop.org/1367-2630/15/1/013048)

New Journal of Physics
The open access journal for physics

The relationship between structure and function in locally observed complex networks

Cesar H Comin, Matheus P Viana and Luciano da F Costa[1]
Institute of Physics at São Carlos, University of São Paulo, São Carlos, SP 13560-970, Brazil
E-mail: ldfcosta@gmail.com

New Journal of Physics **15** (2013) 013048 (15pp)
Received 13 June 2012
Published 18 January 2013
Online at http://www.njp.org/
doi:10.1088/1367-2630/15/1/013048

Abstract. Recently, studies looking at the small scale interactions taking place in complex networks have started to unveil the wealth of interactions that occur between groups of nodes. Such findings make the claim for a new systematic methodology to quantify, at node level, how dynamics are influenced (or differentiated) by the structure of the underlying system. Here we define a new measure that, based on the dynamical characteristics obtained for a large set of initial conditions, compares the dynamical behavior of the nodes present in the system. Through this measure, we find that the geographic and Barabási–Albert models have a high capacity for generating networks that exhibit groups of nodes with distinct dynamics compared to the rest of the network. The application of our methodology is illustrated with respect to two real systems. In the first we use the neuronal network of the nematode *Caenorhabditis elegans* to show that the interneurons of the ventral cord of the nematode present a very large dynamical differentiation when compared to the rest of the network. The second application concerns the SIS epidemic model on an airport network, where we quantify how different the distribution of infection times of high and low degree nodes can be, when compared to the expected value for the network.

S Online supplementary data available from stacks.iop.org/NJP/15/013048/mmedia

[1] Author to whom any correspondence should be addressed.

Contents

1. Introduction

Given that complex systems are almost invariantly composed of a large number of interacting elements, they can be efficiently represented and studied in terms of complex networks [1–3]. In this representation, their structural and dynamical properties can be extracted and investigated. Typically, the structure of such networks is quantified in terms of several measures [4], reflecting different properties of the respective topology (e.g. node degree, shortest paths, centralities) and geometry (e.g. arc length distances, angles, spatial density).

A great deal of the investigations into the structure and function in complex systems have focused on trying to predict the dynamics from specific structural features [5–7]. Such an ability would provide the means for effectively controlling real-world systems [8]. Despite the growing number of studies devoted to this problem, knowledge about the relationship between the structural and dynamical properties remains incipient due to three main reasons: (a) dynamics are often summarized in terms of global statistics, which overlooks their intricacies; (b) the investigation often focuses on linear relationships such as correlations between the structural and dynamical features; and (c) several effects, such as the initial conditions, network topology, stochasticity or dynamical differences from node to node are not selectively fixed or controlled. Still, there are some notable examples of local analysis previously done on complex networks. Gómez-Gardeñes *et al* [9] studied how synchronization takes place on heterogeneous random systems, in comparison to their homogeneous counterparts. They found that the systems differ by the way smaller synchronized groups are formed while increasing the coupling strength of the Kuramoto oscillators. Kitsak *et al* [10] used the *k*-shell index to find nodes having a high potential to spread a disease in a network. Their fundamental result was that it is possible to predict the number of infected nodes of the entire system by knowing only the *k*-shell of the initial infected node. Still on the subject of epidemics, there are other works concerning local analysis [11, 12] as well as the use of temporal networks to show the importance of initial conditions [13]. Another good example

of local structural analysis that can be translated to dynamical behavior is the study of motifs [14, 15].

In this paper we propose a novel methodology to quantify how much the dynamics at each node differentiates from the dynamics at the other nodes as a consequence of specific aspects (e.g. local anisotropies) of the network structure. This is accomplished by simulating the investigated dynamics for a large number of initial conditions and checking how much a given property of the dynamics at a node (e.g. entropy of the time series at that node) deviates from the overall dynamics. The level at which a node i 'feels' the structure differently from the other nodes is quantified in terms of a parameter α_i. In this way, the proposed methodology addresses the three shortcomings mentioned above by: (i) being local, i.e. it is applied for each individual node; (ii) by not imposing any specific kind of relationship between the dynamical and structural features; and (iii) isolating each (above mentioned) condition that can affect the dynamics.

Several important findings have been obtained using this methodology. Our results show that the nodes feel rather distinctly the structure in most of the considered situations. While the Erdős–Rényi (ER) model [16] does not show any dynamical differentiation, Waxman's [17] geographic model presents fluctuations that naturally create different dynamical groups through the density of connections. The Barabási–Albert (BA) [18] model shows a rather distinct behavior for the highly connected nodes of the network, which end up having a very distinct dynamics related to the rest of the network, being even more extreme than the topological differences. When considering a real network of the nematode *Caenorhabditis elegans*, we find that there exists a group of neurons where the local topology influences the spike rate of the signals rather distinctly. Finally, the study concerning epidemic dynamics shows that the first infection time of a node can have different levels of variability depending on the degree of the node.

2. Methodology

2.1. Measuring the differentiation

In order to apply our methodology we begin with a network having N nodes, which will be fixed through the entire process, and execute M times a dynamics on it. Each execution starts with a randomly sorted initial condition. It is important to note that depending on the dynamics being studied we can have a specific set of initial conditions that take the system to a particular state, so this state will rarely be accessed by sampling. This is not a problem in our method because we are analyzing the dynamics for the set of initial conditions imposed, that is, we are studying the signals that are in fact observed.

In possession of the dynamic signals for the M realizations, we need a mechanism to represent them in order to compare their behavior. In our case we use dynamical measures that try to extract the most relevant information about the signals. Let F be one of such measures, after the many realizations we get a set of observed values $f_{i,r}$, where i is the node index and r the realization. These values can vary through four distinct mechanisms: (a) initial condition, (b) network topology, (c) particular dynamics and (d) stochasticity. Our objective here is to study only the relation between the initial condition and the topology, so we use dynamics having identical equations for every node and consider stochastic variations to be small. With these restrictions the only variations we can observe on the values of F are: (a) fluctuations

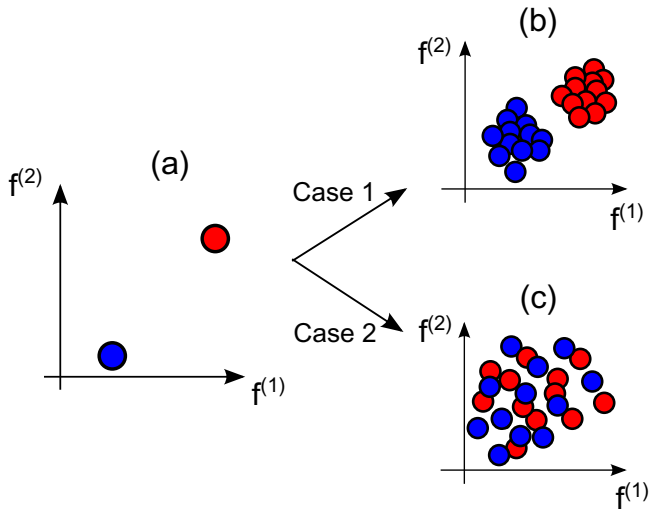

Figure 1. Example of topological differentiation for two nodes. Projecting the signal characteristics for a single realization of the dynamics (a) is not sufficient to discern the topological influence. We need to consider many distinct initial conditions in order to infer if the topology differentiates (b) or not (c) the dynamics of the nodes.

on the dynamical values of a given node, which given the fact that the topology is static, *can only be caused by the variation of the initial condition*. (b) Differences on the mean values of F for distinct nodes, that because of the properties assumed *can only be caused by the topological differences of the nodes*. The term *topological difference* needs to be used with care, because unless in very specific cases where the network is perfectly symmetric (e.g. a lattice with toroidal boundary), the topology of two given nodes is rarely identical, that is, we can always find a structural characterization that will have distinct values for them. Nevertheless, since the networks we use are not regular, every significant dynamical difference we observe must be caused by the topology. It is also important to note that the reverse is not true, if the dynamics of two nodes appear to be the same, their topologies are still distinct, what happened is that both nodes *felt* the topology in the same manner. An example of this last case is the diffusion dynamics on graphs [19], in which the equilibrium behavior depends only on the degree of the nodes, that is, although the nodes possess distinct general topology, the localized characteristic of the dynamics allows only the degree to differentiate the nodes. The two cases we may come across are shown in figure 1.

In order to quantify the difference of the values obtained for each node, we use a statistical test that will be defined on the next section. Through this statistic we can identify if the difference of the means of the dynamical values obtained are in fact significant, that is, we are quantifying the difference between the dynamics of the nodes normalized by the intrinsic fluctuations caused by the initial condition. The distance between every pair of nodes is then represented by the matrix Ξ, where each line i and column j represents the distance obtained between the nodes i and j, that is

$$\Xi_{ij} = d_h(i, j), \tag{1}$$

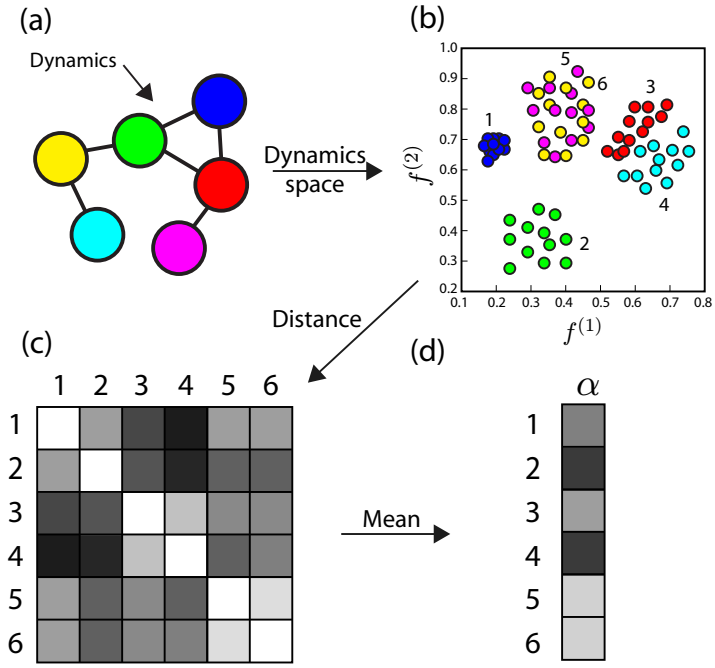

Figure 2. Example application of the methodology. (a) We execute 12 runs of the dynamics with different initial conditions and (b) project the obtained signals on the measure space, which in this case is two dimensional. (c) The matrix Ξ_{ij} is obtained using equation (6), and (d) its mean is taken in order to obtain the α of each node.

where $d_h(i, j)$ is the distance defined by equation (6). Finally, we can define our mean dynamical differentiation measure, α, as the mean values of each line of this matrix

$$\alpha_i = \frac{1}{N-1} \sum_{j=1}^{N} \Xi_{ij}. \tag{2}$$

The standard procedure now would be to calculate the statistical significance of the observed values of α, but since we are concerned with the comparison between the distances and not their absolute values, this does not need to be performed. To carry out the comparison, we construct a histogram of the obtained values of α. Having in mind that α is relative to some dynamical characteristic, we can have distinct histograms relative to the desired characterization.

What we search for are particular behaviors of the histograms, for example, it is expected that a single node with very distinct dynamics compared to the rest of the network will have a very large α value. It is important to observe that although we presented the methodology for a single measure, nothing prevents us from calculating the distances using simultaneously various dynamical measures. In figure 2 we show an example application of the presented model.

2.2. Distance measure

In this section we present and motivate the statistical test that will be used throughout the paper. Suppose that we have a set of points inserted in a m-dimensional space and these points form

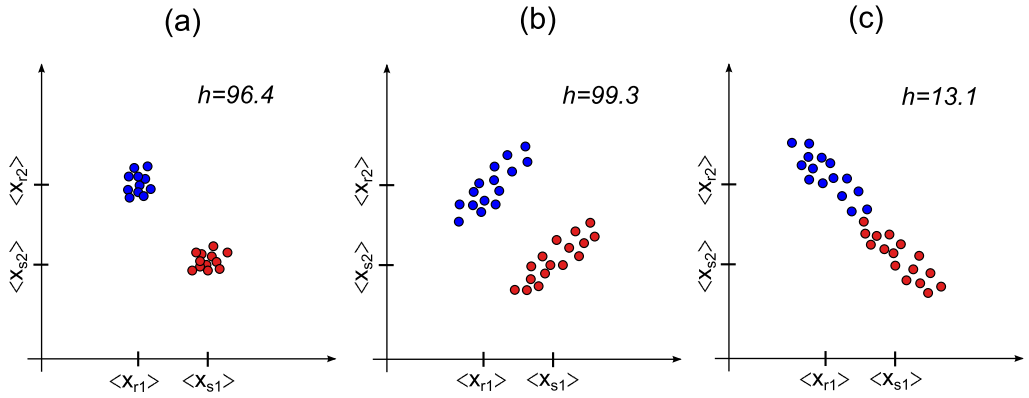

Figure 3. Example of Hotelling distances. The three figures present the same Euclidean distance between the mean point of the two groups, but in cases (a) and (b) the statistical distance is more significant than in (c). The respective Hotelling distances are shown in each figure.

two distinct groups (see figure 3). An immediate way to quantify the separation between the groups is by using the Euclidean distance between the center of mass of each group, given by

$$d(r,s) = \sqrt{\sum_{i}^{m} (\langle x_{ri}\rangle - \langle x_{si}\rangle)^2}, \tag{3}$$

where r and s are the indices of the groups and $\langle x_{ri}\rangle$ represents the mean (or center of mass) of the points in group r on the ith dimension. The disadvantage of the Euclidean distance is that it does not take into account the dispersions of the groups from which the distance is being measured. Following figure 3, if we have two random variables X_r and X_s with the realized values $\{x_r\}$ and $\{x_s\}$, marked respectively in blue and red in the figure, clearly the two cases shown in figures 3(a) and (b) have a more significant distance between their mean than the case in figure 3(c), although the Euclidean distance is the same. There are many ways to take into account the dispersions of the groups, here we use the Hotelling statistic [20], which considers the variance of each group in the direction defined by the line that passes between the two means. The distance, h, between two groups is defined by

$$h^2 = \frac{n_r n_s}{n_r + n_s} (\langle \vec{x}_r\rangle - \langle \vec{x}_s\rangle)' \Sigma_m^{-1} (\langle \vec{x}_r\rangle - \langle \vec{x}_s\rangle), \tag{4}$$

where $\langle \vec{x}_r\rangle$ is the average position of group r. The variable Σ_m is the estimation of the equivalent covariance matrix of each group, given by

$$\Sigma_m = \frac{n_r \Sigma_r + n_s \Sigma_s}{n_r + n_s - 2}, \tag{5}$$

where Σ_r and Σ_s are the covariance matrix of the groups. In figure 3 we show the values of h for each case.

The distance we considered is a hybrid version of both Euclidean and Hotelling metrics. This is required in order to avoid singularities for a set of nodes with zero variance. Since we are not interested in the absolute value of the distance, but in the comparison of values between

groups, we can define a new hybrid distance, given by

$$d_h(r, s) = \frac{h}{h+1} d(r, s), \tag{6}$$

where $d(r, s)$ is the usual Euclidean distance between the center of mass of groups r and s.

Finally, in cases where we have to calculate the distances through more than one variable, it is necessary to normalize them so as to give a fair comparison. In order to do so we calculate the standard score of each measure x, given by

$$\hat{x} = \frac{x - \langle x \rangle}{\mathrm{std}(x)}, \tag{7}$$

where $\mathrm{std}(x)$ is the standard deviation of x. Since throughout this paper we always use the standardized version of the values, we simplify the notation by calling $\hat{x}$ just by x.

3. Studied systems

In this section we briefly present the topologies and dynamics where our methodology was applied.

3.1. Network models

In our study we use three network models to compare the methodology applied to three distinct scenarios. The first is the traditional ER [16] model that connects every possible pair of nodes with a probability p, originating a Poisson degree distribution representing a completely random graph.

The second scenario is when every connection has a cost associated with it, usually represented by a geographic network where the nodes have a spatial position. A commonly used mechanism to model this behavior is due to Waxman [17], who randomly placed the nodes in a [1,1] grid and, for every pair (i, j) of nodes, defined a probability of this pair being connected, given by

$$P(i, j) = \beta \mathrm{e}^{-d(i,j)/d_0}, \tag{8}$$

where $d(i, j)$ is the Euclidean distance between nodes i and j, β tunes the density of edges, and d_0 sets the typical size of the connections. This generates a topology where a long range shortcut rarely occurs. Because of this, one of the main characteristics of this model is that, for the most commonly used values of β and d_0, it does not exhibit the small world behavior.

Another class of networks, usually called power-law, is greatly represented by the BA [18] model, where two principles, namely growth and referential attachment, are used to model the usual power-law degree distribution ($P_k \approx k^{-3}$) that is observed in many real systems [21]. Due to the power-law distribution, the degree of the nodes can show large fluctuations.

3.2. Dynamics used

3.2.1. Integrate-and-fire dynamics. Our first application of the method will be related to the transmission of neuronal signals, modeled by the integrate-and-fire dynamics [22].

This model treats the neuron as an integrator with a hard threshold limit, $\mathcal{T}$. The actions of a given neuron i along the time is stored in the binary time series $s_i(t)$, which indicates that the neuron spikes at instant t whether $s_i(t) = 1$. Using this time series, we define the spike rate of a neuron as

$$r_i = \frac{1}{T_{\text{sim}} - T_{\text{est}}} \sum_{t=T_{\text{est}}}^{T_{\text{sim}}} s_i(t), \tag{9}$$

where T_{sim} is the total simulation (or experiment) time and T_{est} is a long enough time, found empirically through prior simulations, so that r_i do not significantly change after T_{est} (we say that the dynamics stabilized). This measure corresponds to the average number of spikes during the considered interval, and it is widely used in neuroscience, since many neurons codify the stimulus amplitude through the rate of spikes [23].

3.2.2. SIS model. In the SIS (susceptible–infected–susceptible) model each node can be in one of two states: infected or susceptible. The spread of the disease between neighbors happens with a rate β, interpreted as the probability per time step that the disease will spread from an infected node to a susceptible one. Each node returns to the susceptible state with a rate γ which, without loss of generality, we define as being $\gamma = 1$. There are many ways to simulate the start of this disease on a network. Here we chose to randomly select with equal probability a single node and turn it to the infected state. After iterating the dynamics for a sufficiently long time, we keep the simulation data if the disease has spread to the entire network, otherwise we start another simulation. The measure we use in this case is the first infection time, I_{f}, defined as the iteration where the node got its first infection. The value of I_{f} for a node will strongly depend on the initial condition, so it will be a good case study for our method.

4. Results

4.1. Integrate-and-fire on random network models

We compare the differentiation relative to the spike rate feature, which we call α_r for different network topologies, namely ER [16], BA [18] and Waxman's geographic model [17]. In figure 4(a) we show the result obtained for the geographic model with $N = 1000$ and $\langle k \rangle = 10$ and an integrate-and-fire dynamics with $\mathcal{T} = 8$ taking place on the network. To obtain statistical significance we use 100 different generated networks; each network is subjected to $M = 1000$ realizations of the dynamics with different initial conditions. We construct histograms of α_r obtained for each generated network and show in figure 4(a) the mean value and standard deviation the set presents. We see that the frequency of nodes with small α_r has a large variation, which is caused by the intrinsic fluctuations in the dynamics of the many nodes with a similar spike rate present on the network. It is feasible to think that this fluctuation would decay as α_r increases, but this is true only for intermediate α_r, while at high values of α_r we observe a sudden increase of fluctuation. Additionally, the mean value stays at an almost constant value for α_r in the range [0.9, 1.5]. This is caused by the high potential of the geographic model to display structural fluctuations, originating in regions with a higher density when compared to the rest of the network. These regions significantly alter the spike rate of the nodes and highly differentiated groups appear. In figure 4(b) we apply the same procedure to calculate α_r in the geographic model, only changing the dynamic threshold to $\mathcal{T} = 10$. It is clear that the groups

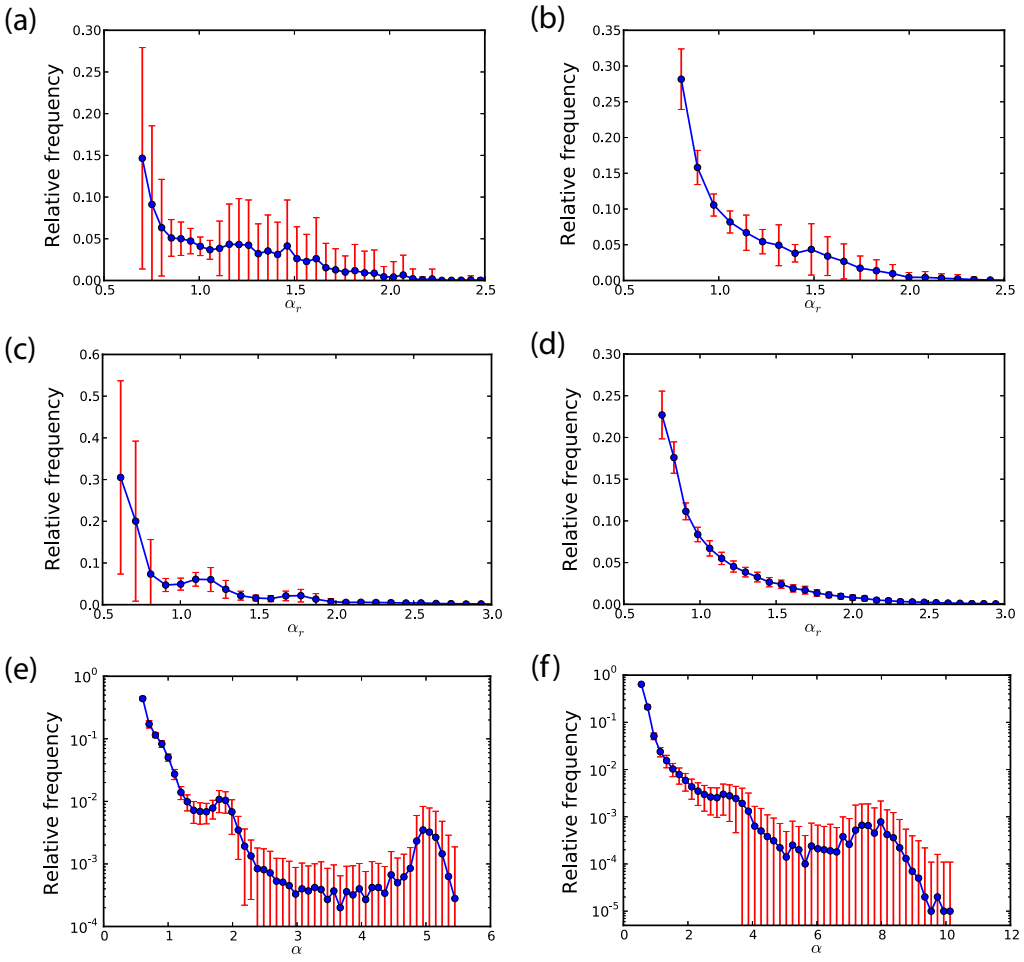

Figure 4. α values obtained for the integrate-and-fire dynamics on random networks. We generated 100 networks for each case and obtained the histogram of α for each network. The plots show the mean and standard deviation of these histograms. The networks were generated using (a) and (b) for the geographic model, (c) and (d) for the ER model and (e) and (f) for the BA model. The graphics on the left were obtained using $\mathcal{T} = 8$ and the ones on right with $\mathcal{T} = 10$.

are no longer distinguishable. This is so because the threshold is now so large that even the topological fluctuations cannot differentiate a significant number of nodes, when compared to the ones with small α_r.

We apply the same procedure used for the geographic model to the ER networks with $N = 1000$ and $\langle k \rangle = 10$. In figure 4(c) we show the information about the obtained histograms for $\mathcal{T} = 8$, which makes it clear that this model exhibits much smaller fluctuations. This is caused by the much shorter geodesic distances that the model exhibits, when compared to the geographic counterpart, which creates a more compact network. We also show in figure 4(d) the case $\mathcal{T} = 10$ for the ER model, where we see a decaying behavior expected for a random Poissonian system.

The third investigated model is the BA with $N = 1000$ and $\langle k \rangle = 6$. Figures 4(e) and (f) show the log-scale histogram for, respectively, $\mathcal{T} = 8$ and 10. In both cases we observe a

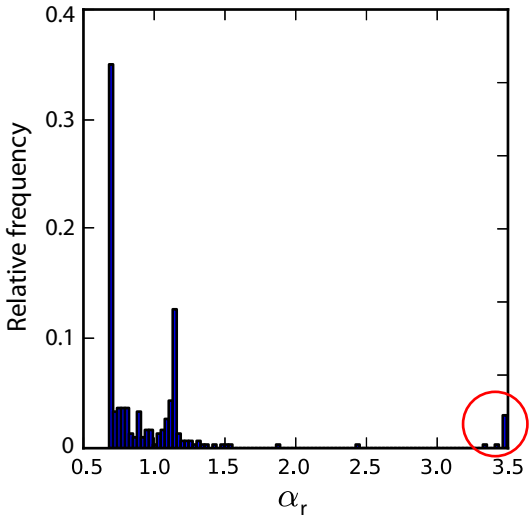

Figure 5. Histogram of the α values obtained for the integrate-and-fire signals taking place on the *C. elegans* network. A total of 1000 realizations of the dynamics were made with different initial conditions. The α was calculated using the spike rate measure.

significantly high peak for large α_r, which we found to be related to the network hubs (nodes with a very high degree). This result was expected, given our observations of large fluctuations on the geographic network, but this is not the main result for the BA model. The important result is that although the power-law degree distribution of the model has a continuous decaying behavior, the histogram of α_r shows small values for intermediate α_r and increases for large α_r. This behavior can be interpreted as follows: the dynamical differentiation of a hub is, as expected, very large, but a node with an almost equal degree can end up with a much smaller differentiation, having dynamics more similar to the low degree nodes. This result confirms the important role that hubs have in complex systems, not only in the sense of being central, but also in having a different purpose to the network dynamics.

4.2. Integrate-and-fire on the network of the Caenorhabditis elegans

Although many interesting properties arise when studying random network models, it is on real networks that the dynamical differentiation analysis can show its real potential. To show this we now apply the methodology to the *C. elegans* neuronal network. In this network, each node represents a neuron and two nodes are connected if there exists some kind of directed communication between them (e.g. synapses, gap junctions, etc). The data was compiled by Chen *et al* [24, 25] and obtained from [26]. The network has 279 nodes and $\langle k \rangle = 22.4$.

Although we motivate the method with a real network, it is important to note that our dynamics do not take into account many signal particularities that arise for real neurons [27], therefore we are looking for a coarse grained description of the neurons inside the network. We will show that, even with this simplified description, it is still possible to observe some interesting phenomena.

We begin by showing in figure 5 the histogram of α relative to the spike rate, α_r. An immediate result observed is that there is a group with high differentiation, indicated with a

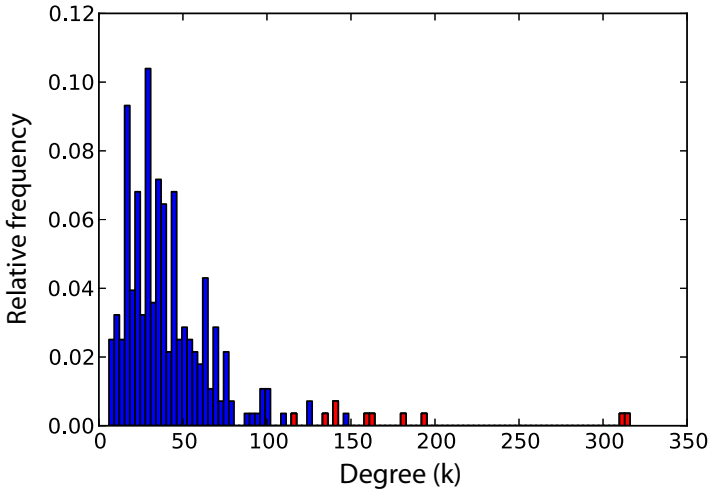

Figure 6. Degree distribution of the *C. elegans* network. Red bars indicate the nodes present in table 1, i.e. the nodes indicated in figure 5 as having the highest α values.

red circle, which is somewhat similar to that observed for the BA random model. With this in mind, we plot in figure 6 the degree histogram of the network, indicating in red the nodes inside the observed group. We see that the nodes with high differentiation possess a high degree in the network, but there are some high degree nodes that do not show a distinct dynamics. It is clear that the topology influence does not occur merely by the degree of the nodes. There is a particular relation between these high differentiated nodes that make their dynamics very peculiar when compared to the rest of the network. The nodes inside the red circle in figure 5 are known as the interneurons of the ventral cord of the *C. elegans*. They are well recognized for possessing a high number of synapses [28], given that they make the bridge between sensory and motor neurons without much restriction on the type of transmitted signals (some classes of interneurons are known for receiving only a specific type of signal).

In table 1 we present the traditional names of these neurons and some information about their spatial and topological distance. Position refers to the spacial localization of each neuron relative to the axis that goes from the head (value 0) to the tail (value 1) of the nematode. We see that the majority of the described neurons are on the head (more specifically, in the nerve ring of the nematode [29, 30]), with the exception of PVCL and PVCR that are on the tail. In the table, $D1$ is the mean topological distance between these neurons, that is, given a node i we measure how many edges we need to travel in order to go to node j and we take the mean of this distance for all j inside the differentiated group. $D1 = 1$ means that the neuron is a neighbor, or receives a direct signal, of all the other neurons shown in the table. The feature $D2$ complements $D1$ as it shows the topological distance between the given neuron and all other neurons, excluding those present in the table. We see that in all cases the differentiated nodes are closer between themselves than with the rest of the network, an effect partially provoked by their high degree. That is, besides having a high degree, these nodes are well connected between themselves, originating a high capacity of communication inside the group and rendering their dynamics very distinct in comparison to the rest of the network. In order to illustrate these neurons, we show in figure 1 of the supplementary material (available from stacks.iop.org/NJP/15/013048/mmedia) the position of each neuron inside the nematode.

Table 1. Calculated topological values for the interneurons of the ventral cord of *C. elegans*. See the text for explanation about $D1$ and $D2$.

Neuron	AVAL	AVAR	AVBL	AVBR	AVDL	AVDR	AVEL	AVER	PVCL	PVCR
Position	0.13	0.13	0.15	0.15	0.16	0.16	0.13	0.14	0.82	0.82
$D1$	1.00	1.00	1.00	1.44	1.33	1.22	1.33	1.22	1.00	1.00
$D2$	1.76	1.77	1.81	1.80	1.97	1.87	1.90	2.09	2.04	2.03

4.3. Epidemics on the airport network

In order to show that our approach can be applied to a completely different system, we also study the first infection time, I_f, of nodes going through an epidemics dynamic. The network we used is the world-wide airport network, which describes the flying routes between a large number of airports throughout the world. Each node is an airport (3302 in total) and an edge indicates that there is a flight between two airports. The data was obtained from www.openflights.org/data.html.

The value of I_f for a node strongly depends on its relative position with respect to where the epidemic has started, so for each initial condition the node will have many different I_f values. Still, the local topology of the node can influence the mean time it takes for the infection to arrive. Suppose we take two nodes with similar local topologies. It is expected that their I_f will be slightly different, but how different are they really? If the topologies of these nodes completely define their dynamics, i.e. I_f does not vary for the different initial conditions, then the difference in their dynamics, caused by the topology, is significant. On the contrary, if their mean values are close but the topology does not have a strong influence, then with different initial conditions it is possible to reach many different dynamical states. This means that the nodes are not significantly differentiated by their local topology. We reinforce that this is different from simply comparing their mean values, as we are taking into account both the variation caused by the initial condition and the network topology.

We performed 1000 simulations of the SIS epidemics dynamics and calculated the respective α values; the result is shown in figure 7. The first noticeable feature is the small number of nodes having a large value of α, these nodes are probably in regions of the network where the epidemic is far from its average behavior. They could be isolated from the network and so would have a very small chance of receiving the disease, or they could be hubs of the network: it does not matter where the disease started, they always catch it immediately. We found that the node types are actually mixed, with a little advantage to the hubs. In figure 8(a) we plot α versus the degree of the nodes, where we see that a node having high degree is guaranteed to have large α, while low degree nodes can have a range of α values. This means that high degree nodes depend only on their immediate neighborhood, i.e. the local topology completely defines their dynamics. On the other hand, as low degree nodes are strongly influenced by their higher order neighborhood, the local topology is not sufficient to define the dynamics. In order to show this more precisely, we plot in figure 8(b) the relationship of α with the closeness centrality of the nodes [31]. This topological measure takes into account all network nodes on its calculation, quantifying how central the node is. We see that the closeness defines much better the dynamics of the less connected nodes, showing that they are differentiated by the topology, but the difference is caused by the influence of the entire network.

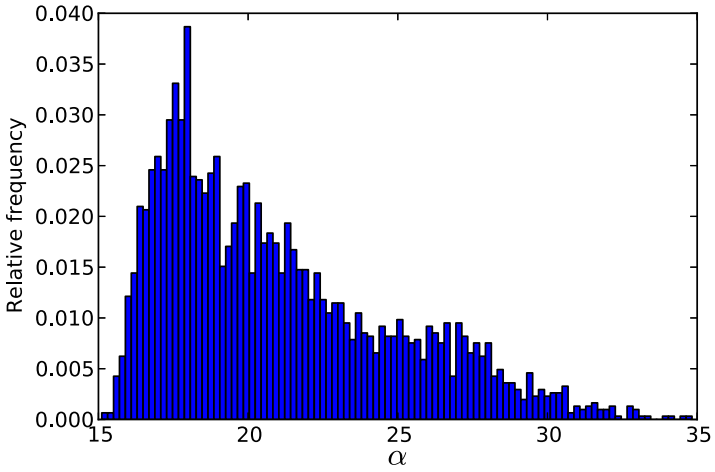

Figure 7. Histogram of α values related to the SIS epidemic model taking place on the airport network. The first infection time of the nodes was taken as a dynamical feature to calculate α.

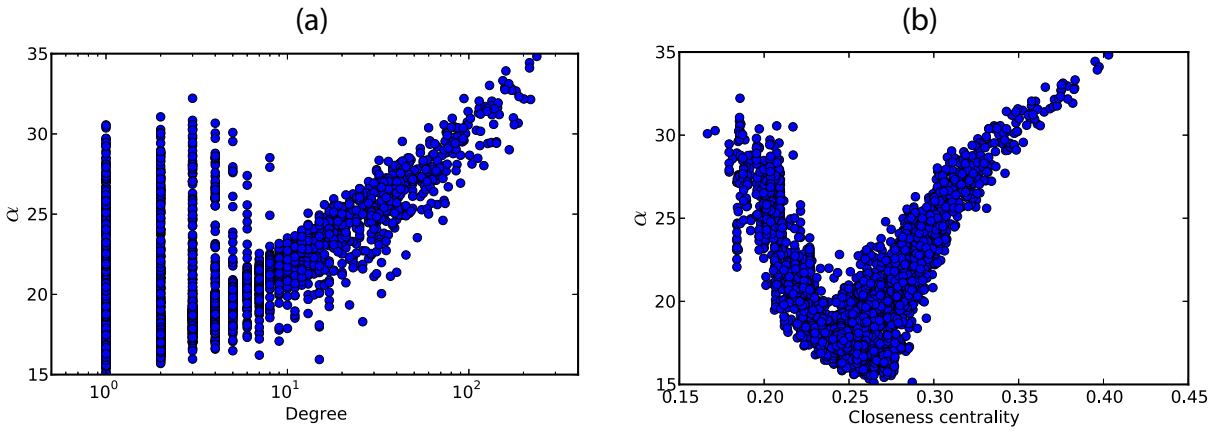

Figure 8. Scatter plots between α and topological features of the nodes. The two features used were (a) the degree and (b) the closeness centrality.

5. Conclusions

In a dynamical system underlain by a completely regular topology (e.g. a toroidal lattice), every node behaves identically regarding its influence on the overall dynamics. It remains an important question to quantify how local heterogeneities in the topology, which can be understood as structural symmetry breaks, may influence the unfolding of the respective dynamics. Despite continuing interest in this area, relatively incipient results have been obtained as a consequence of the fact that several elements that can interfere with the overall dynamics—such as initial conditions, stochasticity and parameter configurations—are not kept constant while inferring individual effects. This paper has addressed this problem by proposing a framework for quantifying to which extent the topology around each node contributes to differentiating the dynamics. Moreover, the method is devised in such a way as to not require the consideration of any specific topological or structural metric. Though the method can be applied with respect to

any of the potentially interfering elements, in this work we restrict our attention to the effect of the initial conditions.

We illustrated the potential of the reported methodology with respect to a range of different topologies, namely the ER, BA and geographic random models, the neuronal connections of the nematode (*C. elegans*) and the world-wide airport network. Several interesting findings are reported, including the fact that the nodes in ER networks are not significantly differentiated regarding their respective time series. The geographic model exhibits groups of nodes that are highly differentiated in comparison to the majority of network nodes, a consequence of the high statistical fluctuations present in the network construction. The result for the BA model showed that the topological particularities of the hubs are amplified in the dynamics taking place on the system.

Regarding the *C. elegans* network, we found that some nodes are highly differentiated by their spiking rate. While all these nodes have been found to be well connected nodes, there are well connected nodes that are not in this group, indicating the presence of additional topological influences besides the node degree. We identified these highly active nodes as corresponding to interneurons of the ventral cord of the nematode. For the epidemic dynamics we found that nodes having a high degree are isolated from the network influence, in the sense that their dynamics show little variation with different initial conditions. On the other hand, in order to predict the dynamics of low degree nodes, we need information that it is not available on their local neighborhood.

Several future developments are possible, including the consideration of other types of dynamics, other models of networks, as well as investigating the effect of stochasticity and varying the parameters or dynamics at each node.

Acknowledgments

LFC is grateful to FAPESP (05/00587-5) and CNPq (301303/06-1 and 573583/2008-0) for the financial support. CHC is grateful to FAPESP for sponsorship (2011/22639-8). MPV thanks FAPESP for their financial support (2010/16310-0).

References

[1] Newman M E J 2003 The structure and function of complex networks *SIAM Rev.* **45** 167–256
[2] Albert R and Barabási A L 2002 Statistical mechanics of complex network *Rev. Mod. Phys.* **74** 47–97
[3] Costa L F, Oliveira O N, Travieso G, Rodrigues F A, Boas P R V, Antiqueira L, Viana M P and Rocha L E C 2011 Analyzing and modeling real-world phenomena with complex networks: a survey of applications *Adv. Phys.* **60** 329–412
[4] Costa L F, Rodrigues F A, Travieso G and Boas P R V 2007 Characterization of complex networks: a survey of measurements *Adv. Phys.* **56** 167–242
[5] Barrat A, Barthelemy M and Vespignani A 2008 *Dynamical Processes on Complex Networks* (Cambridge: Cambridge University Press)
[6] Boccaletti S, Latora V, Moreno Y, Chavez M and Hwang D U 2006 Complex networks: structure and dynamics *Phys. Rep.* **424** 175–308
[7] Dorogovtsev S N and Goltsev A V 2008 Critical phenomena in complex networks *Rev. Mod. Phys.* **80** 1275–335
[8] Liu Y-Y, Slotine J-J and Barabási A L 2011 Controllability of complex networks *Nature* **473** 167–73

[9] Gómez-Gardeñes J, Moreno Y and Arenas A 2007 Paths to synchronization on complex networks *Phys. Rev. Lett.* **98** 034101
[10] Kitsak M, Gallos L K, Havlin S, Liljeros F, Muchnik L, Stanley H E and Makse H A 2010 Identification of influential spreaders in complex networks *Nature Phys.* **6** 888–93
[11] Barthélemy M, Barrat A, Pastor-Satorras R and Vespignani A 2004 Velocity and hierarchical spread of epidemic outbreaks in scale-free networks *Phys. Rev. Lett.* **92** 178701
[12] Pastor-Satorras R and Vespignani A 2002 Immunization of complex networks *Phys. Rev.* E **65** 036104
[13] Rocha L E C, Liljeros F and Holme P 2011 Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts *PLOS Comput. Biol.* **7** e1001109
[14] Milo R, Shen-Orr S, Itzkovitz S, Kashtan N, Chklovskii D and Alon U 2002 Network motifs: simple building blocks of complex networks *Science* **298** 824–7
[15] Sporns O and Kotter R 2004 Motifs in brain networks *PLOS Biol.* **2** e369
[16] Erdős P and Rényi A 1960 On the evolution of random graphs *Publ. Math. Inst. Hungar. Acad. Sci.* **5** 17–61
[17] Waxman B M 1988 Routing of multipoint connections *IEEE J. Sel. Area Commun.* **6** 1617–22
[18] Barabási A L and Albert R 1999 Emergence of scaling in random networks *Science* **286** 509–12
[19] Lovász L 1993 Random walks on graphs: a survey *Combinatorics, Paul Erdos is Eighty* vol 1 (Keszthely: Janos Bolyai Mathematical Society)
[20] Hotelling H 1931 The generalization of Student's ratio *Ann. Math. Stat.* **2** 360–78
[21] Clauset A, Shalizi C R and Newman M E J 2009 Power-law distributions in empirical data *SIAM Rev.* **51** 661–703
[22] Lapicque L 1907 Recherches quantitatives sur l'excitation électrique des nerfs traitée comme une polarization *J. Physiol. Pathol. Gen.* **9** 620–35
Brunel N and Rossum M C W V 2007 Quantitative investigations of electrical nerve excitation treated as polarization *Biol. Cybern.* **97** 341–9 (Engl. transl.)
[23] Rieke F, Warland D, De R R and Steveninck Bialek W 1999 *Spikes: Exploring the Neural Code* (Cambridge, MA: MIT Press)
[24] Chen B L, Hall D H and Chklovskii D B 2006 Wiring optimization can relate neuronal structure and function *Proc. Natl Acad. Sci. USA* **103** 4723–8
[25] Varshney L R, Chen B L, Paniagua E, Hall D H and Chklovskii D B 2011 Structural properties of the *C. elegans* neuronal network *PLOS Comput. Biol.* **7** 1–21
[26] WormAtlas, Altun Z F, Herndon L A, Crocker C, Lints R and Hall D H (ed) 2002–2012 www.wormatlas.org
[27] Izhikevich E M 2004 Which model to use for cortical spiking neurons? *IEEE Trans. Neural Netw.* **15** 1063–70
[28] White J G, Southgate E, Thomson J N and Brenner S 1986 The structure of the nervous system of the nematode *Caenorhabditis elegans Phil. Trans. R. Soc.* B **314** 1–340
[29] Ware R W, Clark D, Crossland K and Russel R L 1975 The nerve ring of the nematode *Caenorhabditis elegans*: sensory input and motor output *J. Comp. Neurol.* **162** 71–110
[30] Wood W B 1988 *The Nematode Caenorhabditis elegans* (Woodbury, NY: Cold Spring Harbor Laboratory Press)
[31] Newman M E J 2010 *Networks: An Introduction* (Oxford: Oxford University Press)