Journal of Physics: Complexity

PAPER • **OPEN ACCESS**

Autorrelation and *cross-relation* of graphs and networks

To cite this article: Luciano da Fontoura Costa 2022 *J. Phys. Complex.* **3** 045009

View the article online for updates and enhancements.

You may also like

- Evaluation of the learning strategies of the 10th-grade students in a private school of Bucaramanga
E P Carvajal Valero, D M Pabon Gelves, N T Castillo et al.
- Multifractal analysis of financial markets: a review
Zhi-Qiang Jiang, Wen-Jie Xie, Wei-Xing Zhou et al.
- Problem solving in the physics classroom. An analysis with secondary school students
C A Hernández-Suarez, L S Paz-Montes and W R Avendaño Castro

This content was downloaded from IP address 143.107.180.158 on 22/12/2022 at 14:50

IOP Publishing J. Phys. Complex. 3 (2022) 045009 https://doi.org/10.1088/2632-072X/aca57c

Journal of Physics: Complexity

OPEN ACCESS

RECEIVED
1 July 2022

REVISED
8 November 2022

ACCEPTED FOR PUBLICATION
23 November 2022

PUBLISHED
6 December 2022

Original Content from this work may be used under the terms of the Creative Commons Attribution 4.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

PAPER

Autorrelation and *cross-relation* of graphs and networks

Luciano da Fontoura Costa

São Carlos Institute of Physics—DFCM, University of São Paulo, Av. Trabalhador São Carlense, 400, São Carlos, SP 13566-590, Brazil

E-mail: luciano@ifsc.usp.br

Keywords: autorrelation, cross-relation, graphs, networks, complex networks, complex systems, network structure

Abstract

The concepts of auto- and cross-correlation play a key role in several areas, including signal processing and analysis, pattern recognition, multivariate statistics, as well as physics in general, as these operations underlie several real-world structures and dynamics. In the present work, the concept of multiset similarity, more specifically the coincidence similarity index, is used as the basis for defining operations between a same network, or two distinct networks, which will be respectively called autorrelation and cross-relation. In analogous manner to the autocorrelation and cross-correlation counterparts, which are defined in terms of inner products between signals, the two operations suggested here allow the comparison of the similarity of nodes and graphs respectively to successive displacements along the neighborhoods of each of the constituent nodes, which therefore plays a role that is analogue to the lag in the class correlation. In addition to presenting these approaches, this work also illustrates their potential respectively to applications for the characterization of several model-theoretic and real world networks, providing a comprehensive description of the specific properties of each analyzed structure. The possibility of analyzing the obtained individual autorrelation signatures in terms of their respective coincidence similarity networks is also addressed and illustrated.

1. Introduction

Thanks to continuing advances in graph theory (e.g. [1–3]) and network science (e.g. [4–7]), these two areas now underly a wide range of concepts and applications, extending from systems biology to transportation systems. In this work, we will understand complex networks, or *networks* for short, as corresponding to particularly intricate types of *graphs* (e.g. [8, 9]).

One of the most interesting characteristics of graphs and complex networks is their ability to represent and model virtually every discrete system and phenomenon. Indeed, from the point of view of data structures (e.g. [10–12]), graphs can be understood as providing one of the most general possible resources for representing structures including lists, trees, lattices, etc.

The intrinsic representational generality of graphs and networks has motivated a directly related interesting issue, namely how to mathematically handle, combine and compare the represented structures. This bears a direct analogy with the manipulation of mathematical structures such as functions, fields, vectors and matrices by using arithmetic and algebraic concepts and methods including addition, product, inner products, correlation, etc.

Among the several operations between mathematical structures underlain by respective data structures (e.g. vectors and matrices), the operations of auto- and cross-correlation, as well as the closely related auto- and cross-convolution counterparts (e.g. [13–17]), play a key role in several important areas, including signal and image processing and analysis, mathematic and computational physics, computer vision, control theory, and electronic engineering, to name but a few possibilities. For instance, the blurring of an image is frequently obtained by convolving the image with a Gaussian kernel (e.g. [14, 18]), and the recognition of an instance of a pattern within a signal can be approached in terms of the respective cross-correlation (i.e. template matching, e.g. [14, 19, 20]).

While auto- and cross-correlations are well-defined respectively to vectors and matrices, their extension to graphs and networks is not so straightforward. Related approaches include the concept of *graph Fourier*

© 2022 The Author(s). Published by IOP Publishing Ltd

transform, which involves the spectral analysis (eigenvalues and eigenvectors) of the Laplacian matrix of a given graph or network (e.g. [21–23]). The term *Fourier transform* here tends to be used as an analogy, not literally. Another approach related to the relationship between portions of a graph is *correlation clustering*, aimed at finding cuts that optimize the allocation of edges with large weights to result within respective clusters (e.g. [24–26]). Statistical-based approaches to correlation between graphs have also been suggested (e.g. [27, 28]).

The present work sets out at investigating the possibility to use multiset similarities, in particular the coincidence similarity index [29–31], as a means for obtaining analogues of the auto- and cross-correlation operations respectively to graphs and networks. The main justification for this approach is the fact that the coincidence similarity provides a selective and sensitive means for strictly comparing any mathematical structure (e.g. [30–33]).

In particular, the coincidence similarity operation is here employed in analogy with the inner product (see also [34, 35]) for gauging the similarity between nodes belonging to a same graph or to two graphs or networks. Then, in order to implement the operation analogous to the relative shifting of one signal respectively to the other, we consider the mean coincidence similarity between topological properties of each node of the networks, which is understood as a reference, and those of the nodes belonging to successively defined neighborhoods, or hierarchies (e.g. [36–38]), at topological distance δ from the reference node. Thus, in this work the variable d plays an analogous role to the *lag* variable in traditional correlation analysis.

The resulting operations between two graphs are here called *autorrelation* and *cross-relation* so as to indicate their analogy with the classic *autocorrelation* and *cross-correlation*, but at the same time distinguishing from the counterparts, because they are based on different concepts: while the latter two traditional operations are based on lagged *inner products* between functions or vectors, the *autorrelation* and *cross-relation* are founded on the lagged *coincidence similarity* between features representing the topological (or other types) of characteristics of the considered network nodes.

In addition to presenting the thus motivated auto- and cross-relation between two graphs or networks (a graph with itself in the case of autorrelation), we illustrate the potential of the suggested concepts and methods respectively to the identification of lag-based relationships between model-theoretic and real-world networks. In particular, it is shown that the average autorrelation functions can provide markedly specific signatures that can be used as a subsidy for inferring to which model a given network may adhere. In addition, we also address the possibility to derive autorrelation networks from the coincidence similarity between individual autorrelation curves or signatures, with promising results that can reflect in a particularly effective and detailed manner the intricacy of the topology around the nodes of the original networks.

The obtained results indicate that the autorrelation and cross-relation between networks reflect several global and local properties of the analyzed networks, therefore constituting an interesting resource for network characterization, as well as for studying the overall topological organization of networks.

We start this work by presenting the several involved concepts, and then introduce the suggested autorrelation and cross-relation operations and illustrate their potential for the characterization of global and local properties of several model-theoretical and real-world networks.

2. Basic concepts

Given two column vectors $\vec{x}$ and $\vec{y}$ with N elements each, their *inner product* can be expressed as:

$$\langle \vec{x}, \vec{y} \rangle = \vec{x}^{\mathrm{T}} \vec{y} = \sum_{i=1}^{N} x_i \, y_i. \tag{1}$$

If the magnitudes $|\vec{x}|$ and $|\vec{y}|$ of the two vectors are kept constant, or if they are somehow normalized, the respective inner product will provide a measurement of the similarity between these two vectors.

The *cross-correlation* between two real-valued vectors (or signals) $\vec{x}$ and $\vec{y}$ with sizes N can be written as:

$$(\vec{x} \otimes \vec{y})[\ell] = \sum_{i=1}^{N-1} x_i \, y[\ell - i] \tag{2}$$

where vector $\vec{y}$ is suitably padded with zeros and $\ell = 0, 1, \ldots, N-1$, so that the result has size $N-1$.

The *convolution* (e.g. [13, 14]) between two vectors (or signals) $\vec{x}$ and $\vec{y}$ with sizes N can be written as:

$$(\vec{x} * \vec{y})[\ell] = \sum_{i=1}^{N-1} x_i \, y[i - \ell] \tag{3}$$

where vector $\vec{y}$ is suitably padded with zeros and $\ell = 0, 1, \ldots, N-1$, so that the result has size $N-1$.

Let $g(t)$ be a function, with respective Fourier transform $G(f)$. Its *autocorrelation* (e.g. [13, 14]) can be calculated as:

$$g(t) \otimes g(t) \longleftrightarrow G^*(f)G(f) = |G(f)|^2 \tag{4}$$

where * is the complex conjugation, $\longleftrightarrow$ means a time-frequency Fourier pair, and $|G(f)|^2$ is often called the *power spectrum* of $f(t)$.

Given a non-directed graph, the *degree* of each of its nodes corresponds to the respective number of links. The node degree, in general, corresponds possibly to the most important local topological information about the local topology around that node. The *clustering coefficient* or *transitivity* of a node corresponds to the ratio between the total number of links existing between the neighbors of that node and the total maximum number of links that would be respectively possible (e.g. [4–6]).

Given a graph (or network), and one of its nodes i, the $\mathcal{N}_\delta(i)$ *neighborhood* of i corresponds to the set of all nodes that are at topological distance smaller or equal to δ (e.g. [37, 39–41]). The nodes that are precisely at distance δ from a reference node will be henceforth called as a *layer*, being expressed as $\mathcal{L}_\delta(i)$. These successive neighborhoods and layers can be understood as defining a respective hierarchy respectively to i.

Given two networks g and h with N nodes each, we henceforth say they are *aligned*, or *in correspondence*, whenever for each node of g we know which is the respectively corresponding node in h, and vice-versa. In other words, the nodes in g and h are pairwise related through respective bijective binary relations. Observe that two aligned networks may have distinct respective topologies.

The Jaccard index (e.g. [42, 43], see also [44]) applied to two traditional sets A and B, is calculated by dividing the number of elements (cardinality) in their respective intersection by the number of elements obtained in the respective union. Therefore, the original Jaccard index varies in the interval $[0, 1]$, with the largest value indicating maximum similarity (identity) between the two compared sets.

The real-valued coincidence similarity index has been found to provide a more strict comparison between two non-empty multisets than the Jaccard or overlap indices taken separately, which has paved the way to enhanced selectivity and sensitivity (e.g. [30, 32, 33]). More recently, the coincidence similarities established between each node of the network and all the others was suggested [45] as an feature that can be used to effectively complement to the characterization of the topology around each node.

Multisets (e.g. [10, 46–50]) can be understood as a generalization of traditional sets in the sense that elements are allowed to appear repeatedly, with the number of repetitions being called the respective *multiplicity*. Multisets inherits most of the main operations from traditional set theory, with the union and intersection between two sets being replaced by the minimum and maximum between the respective multiplicities of the elements of the two multisets.

The multiset coincidence similarity index has been suggested as a means to complement the Jaccard index (e.g. [42, 43, 51–53]) by also taking into account the interiority between the two compared mathematical structures [29]. It is henceforth assumed that all network measurements are non-negative. Otherwise, the real-valued version of the coincidence index (e.g. [29, 30]) should be considered.

If $\vec{x}$ and $\vec{y}$ are two non-zero vectors of size N, with elements not taking negative real values, the respective multiset Jaccard similarity (e.g. [29–31]) can be obtained as:

$$\mathcal{J}(\vec{x}, \vec{y}) = \frac{\sum_{i=1}^{N} \min\{x_i, y_i\}}{\sum_{i=1}^{N} \max\{x_i, y_i\}} \tag{5}$$

where the maximum and minimum operations are taken between each of the involved multiplicities x_i and y_i. Observe that, as with the original Jaccard similarity index, we also have that $0 \leqslant \mathcal{J}(\vec{x}, \vec{y}) \leqslant 1$.

The respective multiset *interiority* index (also overlap, e.g. [54]) can be expressed as:

$$\mathcal{I}(\vec{x}, \vec{y}) = \frac{\sum_{i=1}^{N} \min\{x_i, y_i\}}{\min\{S_x, S_y\}} \tag{6}$$

with $0 \leqslant \mathcal{I}(\vec{x}, \vec{y}) \leqslant 1$
where:

$$S_x = \sum_{i=1}^{N} x_i; \quad S_y = \sum_{i=1}^{N} y_i. \tag{7}$$

This index provides a quantification of how much any of the two multisets is contained within the other multiset. In case the two compared multisets are identical, we have $\mathcal{I}(\vec{x}, \vec{y}) = 1$.

The *multiset coincidence similarity* between two non-null vectors $\vec{x}$ and $\vec{y}$ then corresponds to the product of the respective real-valued Jaccard and interiority indices [29], i.e.:

$$\mathcal{C}(\vec{x}, \vec{y}) = \mathcal{J}(\vec{x}, \vec{y})\, \mathcal{I}(\vec{x}, \vec{y}) \tag{8}$$

with $0 \leqslant \mathcal{C}(\vec{x}, \vec{y}) \leqslant 1$.

As such, the coincidence similarity requires not only the compared vectors to be similar in the sense of the Jaccard index, but also to take into account the overlap between them in the sense implemented by the interiority index. Therefore, the coincidence similarity index in general results more strict than either the Jaccard or overlap indices taken separately. The interesting properties of the coincidence similarity have been applied with encouraging success to several situations, including characterization of cities [55] and automated identification of motifs in networks [56].

The real-valued coincidence similarity between the two vectors $\vec{x}$ and $\vec{y}$ can be understood as being analogous the Pearson correlation coefficient between these two vectors.

3. Autorrelation and cross-relation between two graphs

Having revised the main basic concepts, we are now in position to mathematically formalize the concepts *Autorrelation* and *Cross-Relation* between two graphs or networks.

Given two aligned networks g and h with N nodes, as well as a set of M topological features x_j, $j = 1, 2, \ldots, M$, we can associate each node $i = 1, 2, \ldots, N$ of those networks to the respective features, therefore obtaining the following *feature vectors*:

$$\vec{X}_{g,i} = \begin{bmatrix} x_{g,1,i} & x_{g,2,i} & \ldots & x_{g,M,i} \end{bmatrix}^{\mathrm{T}} \tag{9}$$

$$\vec{X}_{h,i} = \begin{bmatrix} x_{h,1,i} & x_{h,2,i} & \ldots & x_{h,M,i} \end{bmatrix}^{\mathrm{T}} \tag{10}$$

where $x_{g,j,i}$ means the value of the feature x_j of node i in graph g.

The respective *cross-relation* of those two aligned vectors can now be defined as:

$$(g \boxtimes h)[i, \delta] = \frac{1}{|\mathcal{L}_\delta(i)|} \sum_{k \in \mathcal{L}_\delta(i)} \mathcal{C}\left(\vec{X}_{g,i}, \vec{X}_{h,k}\right) \tag{11}$$

where $\delta = 1, 2, \ldots, \Delta$ is the considered sequence of lags (distances from each reference node) and $|\mathcal{L}_\delta(i)|$ is the number of nodes in the topological layer δ relatively to the reference node i.

In case $g = h$, the above expression becomes the *autorrelation* of a node i of the network g:

$$(g \boxtimes g)[i, \delta] = \frac{1}{|\mathcal{L}_\delta(i)|} \sum_{k \in \mathcal{L}_\delta(i)} \mathcal{C}\left(\vec{X}_{g,i}, \vec{X}_{g,k}\right). \tag{12}$$

Observe that $(g \boxtimes g)[i, \delta]$ with $\delta = 1, 2, \ldots, \Delta$ can be understood as *individual signatures* of each of the network nodes $i = 1, 2, \ldots, N$, corresponding to how much that node in one network is similar in the average, as far as the coincidence index is concerned, to the other network nodes at each successive layer $\mathcal{L}_\delta(i)$. A similar interpretation holds for the individual signatures $(g \boxtimes g)[i, \delta]$, but here the comparison takes place between nodes of the same network g.

The basic underlying assumption is that a system of progressive layers around each node i is therefore defined respectively to successive values of the lag parameter δ, starting at that node when $\delta = 0$. Therefore, N lag variables are established in this manner, allowing one of the networks to be shifted respectively to the other in this highly dimensional space. Observe that, as the lag increases, the number of paths tends to initially increase in a possibly combinatorial manner.

As such, the autorrelation of a network g can be understood in analogy to the autocorrelation of a function or signal, while a similar relationship can be established between the cross-relation between two networks g and h and the cross-correlation between two signals or functions.

As defined above, the autorrelation and cross-relation between two graphs will involve the coincidence comparison between the features of each node with those of nodes in successive neighborhoods, whose number tends to increase steadily in number as they radiate more links, and then decrease as a consequence of reaching the border of the network. This implies that the average coincidence is taken respectively to varying numbers of nodes in each successive neighborhood.

There are several ways, in addition to the henceforth adopted consideration of the average coincidence between the reference node and the nodes at each of its successive neighborhoods, in which the autorrelation and cross-relation between two graphs can be modified so as to make the resulting values more commensurate. One such manner consists in comparing between each node in one network and the nodes in a *walk* along the nodes in the other network. These walks can be specified in many ways, including the more traditional random walks and self-avoiding random walks (e.g. [57, 58]).

The autorrelation (cross-relation) presents potential for characterizing how much one (two) graphs are similar respectively to a sequence of respective *topological lags* δ. In the remainder of this work, we illustrate the application of the autorrelation and cross-relation concepts respectively to several model-theoretical and real-world structures.

An important aspect of the autorrelation and cross-relation between graphs as suggested above concerns the fact that these operations are relative to the set of adopted node features, which can be of topological or complementary types. This flexibility to define specific set of features allows the lag-based comparison of networks to be investigated from the point of view of several choices of measurements, aimed at reflecting network properties of specific interest in each research situation.

For simplicity's sake, the measurements of node degree and clustering coefficient will be adopted henceforth. Though the three types of network models considered in the present work tend to have relatively small cluster coefficients, they have been still been verified to influence the individual autorrelation signatures, also moderately contributing to enhancing the separation between different models, especially in the Watts-Strogatz case.

4. Results and discussion

In this section, the concepts of autorrelation and cross-relation between two graphs are illustrated respectively to model-theoretical and real-world examples. In all cases, networks that happen to be directed are first symmetrized, and only the largest connected component is then taken into account. In addition, given that maximum autorrelation value (equal to one) is always observed for $\delta = 0$, these values are not shown henceforth for simplicity's sake. Observe that this property does not necessarily hold for the cross-relation.

We start the study of the application of the autorrelation and cross-relation concepts respective to three model-theoretical complex networks (e.g. [4–6]), namely: Erdős–Rényi—ER, Barabási–Albert—BA, and Watts–Strogatz—WS in its ring configuration. Consequently, we have the possibility to consider uniformly random interconnectivity (ER), scale free degree distribution (BA), as well as the model WS which varies, in terms of a respective rewiring probability, from a perfect lattice to an ER configuration. The ER, BA and WS, the latter with high reconnection probability, are all characterized by having relatively small average shortest path distances. At the same time, the WS structures with relatively small rewiring probability tends to have larger average shortest path distances.

Figure 1(a) presents the autorrelation signatures obtained for an ER network with $N = 300$ nodes and average degree $\langle k \rangle = 4$, considering topological lags $\delta = 1, 2, \ldots, 15$. Shown are the several autorrelation signatures obtained for each of the 300 nodes (several random colors), as well as the respective mean signature (shown in black). Interestingly, the autorrelation function of this network can be found to decrease monotonically with two negative slopes: one with smaller magnitude taking place from δ equal to 1 to 4.5, which is followed by increased slope magnitude thereon.

The autorrelation function of a BA network, also with $N = 300$ nodes and $\langle k \rangle = 4$, is presented in figure 1(b). Compared to the ER case, a completely different result has been obtained. Now, the autorrelation values increase progressively initially, and then undergo an abrupt decrease for lag $\geqslant 6$. Similarity to the ER, though, the final decrease again coincides with the neighborhoods reaching the network border. The increase of autorrelation values, however, is mostly a phenomenon specific to this type of network among the three considered types of structures. It follows from the fact that, in a BA network, although only a few nodes are hubs, they are soon reached by the neighborhood progressions starting from the other nodes. Because the hubs are not similar to the other nodes, while the other nodes are mutually similar, the autorrelation tends to increase *after reaching* the hubs, until the border is met. The lower individual signatures have been found to be associated to the hubs of the network, which have markedly distinct topological features respectively to the other nodes.

Figure 1(c) depicts the autorrelation signatures obtained for a WS network (ring configuration) with similar parameters as before, obtained by using a rewiring probability of $p_w = 0.05$ and neighborhood of two nodes. A third type of autorrelation function can be observed for this network model, characterized by an initial plateau, followed by the an accentuated decrease which again is a consequence of the finite size of the network. The obtained plateau derives from the intrinsic mutual similarity between most nodes of this network.

In order to explore in more detail the autorrelation structure of the WS networks, figure 2 presents the results respectively obtained for WS networks with $N = 300$ nodes and rewiring probabilities $p_w = 0.005, 0.01$, and 0.02.

It can be readily observed that the increase of the rewiring probability not only reduces the range of lags characterized by non-zero autorrelations, but also influences the sharpness of the transition from core to

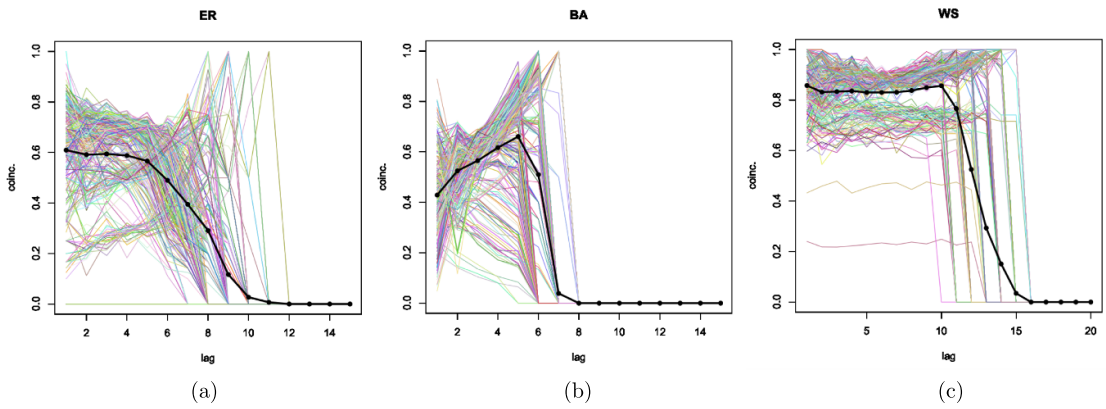


Figure 1. The autorrelation functions in terms of successive lag values δ for ER, BA and WS networks with $N = 300$ nodes and $\langle k \rangle = 4$. The WS network was obtained by using rewiring probability $p_w = 0.05$. Corresponding to the longest shortest path between any two nodes of a network, the diameters of the ER, BA and WS networks in this figure are 11, 7, and 12, respectively.

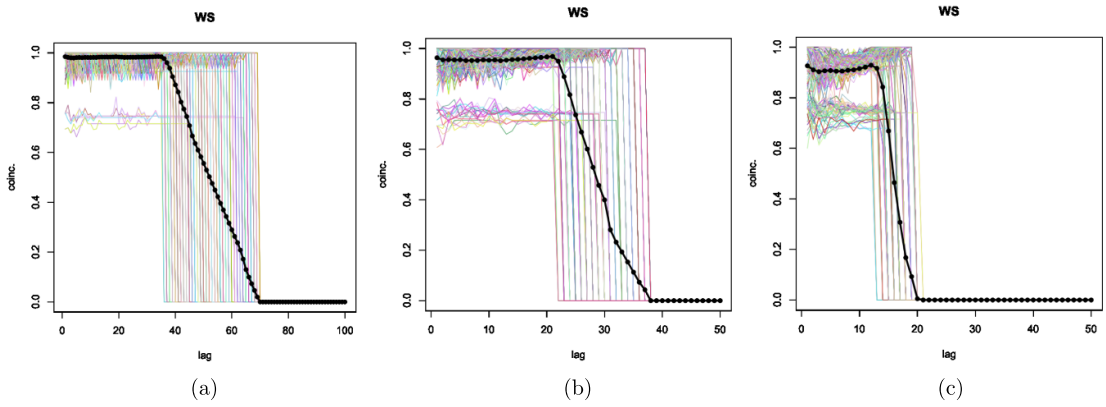


Figure 2. The autorrelation functions of an WS networks with $N = 300$ nodes, $p = 4$ and $p_w = 0.005$ (a), 0.01 (b), and 0.02 (c), respectively to topological lags $\delta = 1, 2, \ldots, 30$.

Table 1. The averages of the Kullback–Leibler divergences between each of the three types of networks considered in figure 1.

	ER	BA	WS
ER	0.0187	0.616	0.62
BA	0.181	0.0060	0.955
WS	2.15	4.99	0.0394

border, as well as the height and extent of the plateaux. Also of interest are the lower signatures, which correspond to outlier nodes originating from the network rewirings.

The markedly distinct average signatures presented and discussed above, especially those in figure 1, suggest that the autorrelation approach has potential to provide an effective means for distinguishing between different types of complex networks, as well as for estimating a likely model for a given specific network. In order to consider these possibilities further, we describe two related studies in the following.

First, we calculate the Kullback–Leibler (KL) divergence [59] between the individual signatures obtained for 30 networks of each of the three models ER, BA, and WS having $\langle k \rangle = 4$ and $N = 300$ nodes and three prototype signatures representing each of these types of networks. For this specific purpose, the individual autorrelation signatures obtained for each group were normalized as statistical distributions, so that the sum of their values is equal to one. After being similarly normalized, the averages of the individual signatures within each of the networks in each of these three types (please refer to figure 1) are taken as respective prototypes.

Table 1 presents the average of the obtained KL divergence values. Each entry $[i, j]$ in this table corresponds to the average distance between the individual signatures in group of type j respectively to the prototype specified by i. The values obtained along the diagonal are noticeably smaller compared to the other

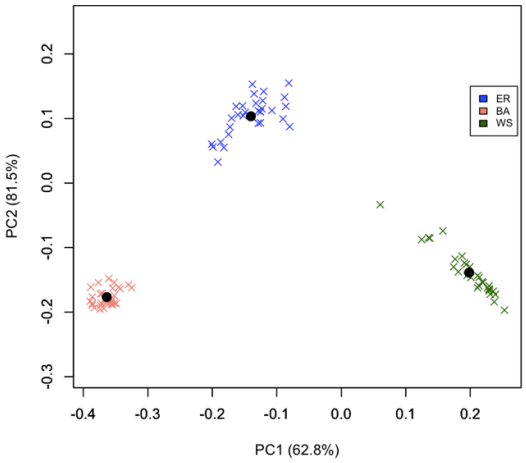


Figure 3. PCA projection of the derivatives of the individual signatures of 30 networks of each of the types ER, BA, and WS, as well as the respective group averages (shown as black disks). Three completely separated groups have been obtained, corresponding respectively to the ER, BA and WS considered types of networks. The cumulative percentage of variation explained by the two principal components are indicted in the respective axes.

entries, substantiating the potential of the above described autorrelation approach to differentiate between different types of complex networks.

In order to further investigate the potential of the autorrelation for distinguishing between network types, we also performed a two-dimensional (2D) projection by principal component analysis (PCA, e.g. [60, 61]) of the numerical derivatives of the average autorrelation functions obtained for 30 networks of each of the three models ER, BA, and WS having $\langle k \rangle = 4$ and $N = 300$ nodes. The numerical derivatives of the individual autorrelation functions have been considered in this experiment because they emphasize the intrinsic variations along each of the individual signatures.

By projecting the high dimensional autorrelation signatures into just two-dimensions, it becomes possible to visually appreciate the possible separation between the three network models. As can be seen from the results shown in figure 3, three well-separated groups have been obtained, corresponding to the ER, BA, and WS types of structures. These results corroborate further the potential of the autorrelation approach for distinguishing, as well as possibly classifying, networks among pre-determined types.

A noticeably interesting possibility is to illustrate the autorrelations between the nodes in a given network in terms of a respective coincidence similarity network [31]. Examples of this type of coincidence network—obtained for the above ER, BA and WS networks, are shown in figures 4(a)–(c), respectively. The size of the nodes reflect the degree of the nodes in the original network. The colors (from cyan to magenta) indicate the transitivity of those nodes.

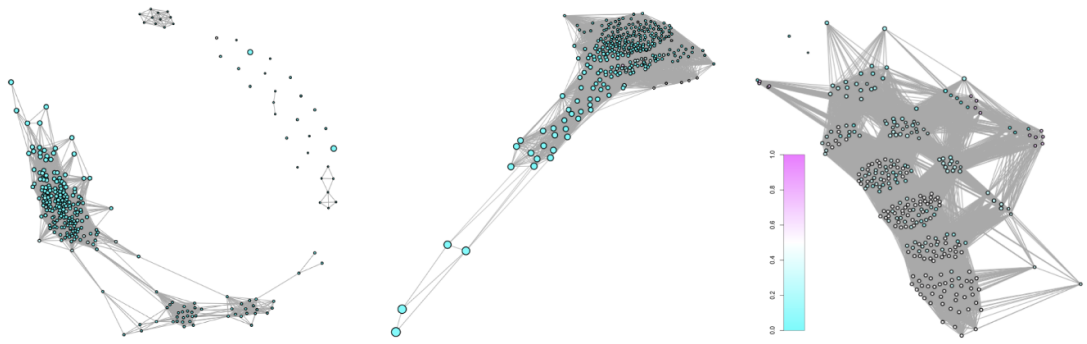
As could be expected, a relatively uniform autorrelation network has been obtained for the ER network, except for some interesting substructures arising from statistical fluctuations. In addition, most of the nodes are shown in cyan, reflecting the small transitivity of the original respective reference nodes.

An interesting structure can also be observed in the case of the BA network, with the hubs resulting in one of the extremities of the larger connected component, which are then successively followed by nodes of signatures obtained for medium and small degree original nodes. This type of structure reflects the fact that, in a BA network, the hubs are markedly different from the other nodes regarding their topological properties along the successive neighborhoods, but are mutually similar. As could be expected, most original nodes have very small transitivity (most nodes are in cyan).

A more uniformly connected autorrelation network has been obtained for the WS case, reflecting the predominant regularity of the original network. Interestingly, the shortcuts implemented by the rewiring with $p_w = 0.05$ implied a segmentation of the autorrelation network into respective modules, or communities. Observe that several nodes are shown in colors close to light magenta, indicating transitivity values relatively higher that those obtained for the ER and BA cases.

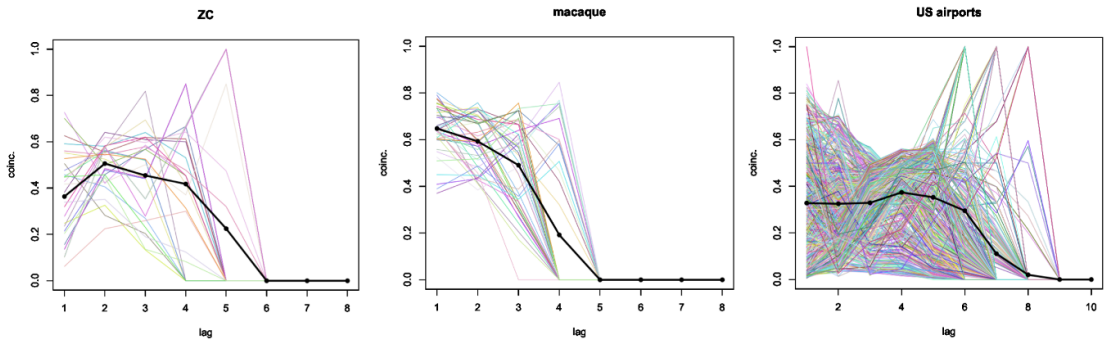
All in all, the autorrelation signatures obtained for the three model-theoretical networks confirm the ability of this approach to reflect several of the important properties of the respective networks, as well as their parameter configurations. The individual signatures also provide an effective means for identifying special types of nodes.

We now proceed to study the types of autorrelation functions that can be obtained from some real-world networks.



(a) (b) (c)

Figure 4. The coincidence networks obtained from the individual autorrelation signatures of the ER (a), BA (b), and WS (c) networks, shown after thresholding by $T = 0.8$, $T = 0.6$, and $T = 0.8$, respectively. Distinct values of threshold have been chosen in order to enhance the respective visualizations. The size of the nodes corresponds to the degree of the original network nodes providing the reference for the autorrelation analysis. The width of the edges is proportional to the respective coincidence value between the respective individual autorrelation signatures. The colors (from cyan to magenta) indicate the transitivity of the original nodes. The particularly dense patterns of interconnections observed in the WS structure in (c), as well as among the nodes with smaller degree in (b), are mostly a consequence of the relatively high similarity values between the nodes in these cases.



(a) (b) (c)

Figure 5. The average (in black) and individual (random colors) autorrelation signatures obtained for the Zachary karate club network [62, 63] (a); *macaque* network [63, 64] (b); and *US airports* network [63] (c). The diameters of the three networks in this figure are: 5, 4, and 8, respectively.

Figure 5(a) presents the autorrelation signature obtained for the Zachary karate club network [62, 63]. Despite the relatively small size of this network, the respective autorrelation signatures suggest that it has hubs, possibly possessing an approximately scale-free node distribution.

The autorrelation signatures obtained for the *macaque* network [63, 64], shown in figure 5(b), is characterized by relatively high initial autorrelation values, which is followed by a mostly gradual respective decrease, defining an overall shape that is similar to that obtained for the ER network (figure 1(a)).

The results obtained for the *US airports* network [63] are presented in figure 5(c). The average signature, which involves relatively small values, indicates a relatively more complex autorrelation structure for this network, involving the three following regimes: (a) an initial plateaux indicating sustained mutual similarity between each node and those at the nearest respective neighborhoods; (b) a following peak of autorrelation values, possibly related to presence of hubs in this network; and (c) a moderate decrease implied by the neighborhood progressions hitting the border of this network. An impressive variety of individual autorrelation signatures can be identified, some characterized by particularly small autorrelation values, indicating the presence of several potentially interesting outlier nodes.

Figure 6 depicts the coincidence similarity network obtained by considering the individual autorrelation signatures of the *US airports* network. The resulting coincidence network has been thresholded at $T = 0.87$ for the sake of enhanced visualization.

Several interesting aspects can be observed from the coincidence network obtained for the US airports network. First, we have that the signatures have been organized along a sequence of interconnections related to the degree of the original nodes. The hubs, being most different from the other nodes, but tending to be similar each other, resulted at one of the extremities of the main obtained structure. As the degree of the

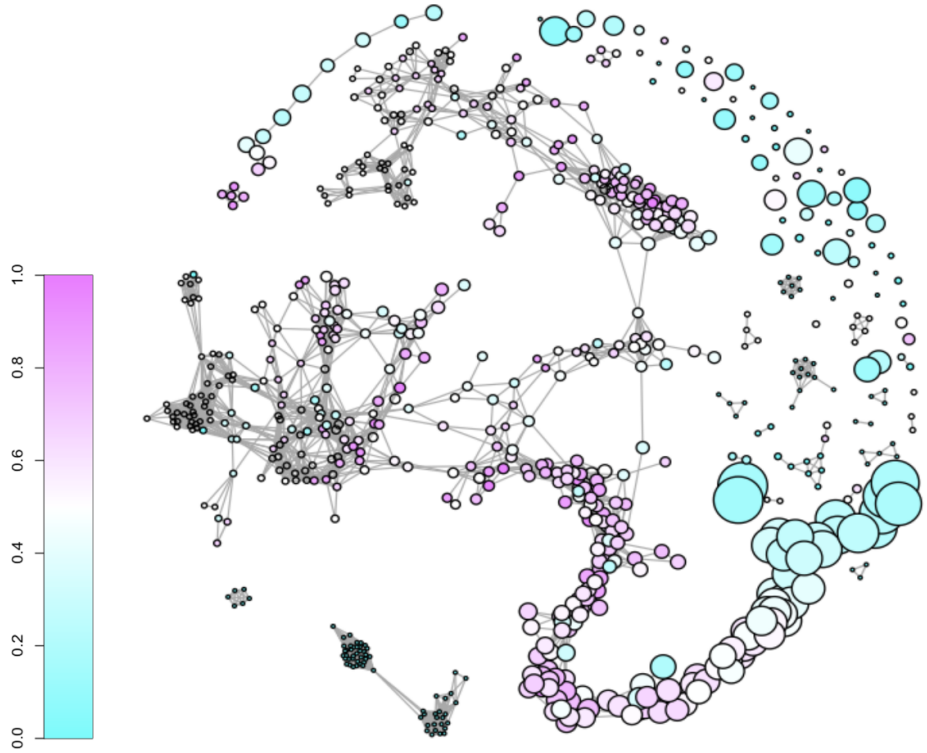


Figure 6. The coincidence similarity network obtained from the individual autorrelation signatures of the *US airports* network, shown after thresholding by $T = 0.87$. The size of the nodes corresponds to the degree of the original network nodes. The width of the edges is proportional to the respective coincidence value between the respective individual autorrelation signatures. The colors (from cyan to magenta) indicate the transitivity of the nodes in the original network.

nodes becomes smaller, the observed structure initially defines a predominantly sequential, chained structure, up to a point where two main branches and more intricate smaller structures start to emerge. This transition is possibly associated to the predominantly scale-free interconnections characteristic of the larger (in the sense of having more flights) airports into the more geographical interconnection patterns between smaller (as indicated by the respective small node diameters associated to smaller original degrees), spatially adjacent/closer airports.

Of particular interest is the separation of this initially string-like structure into two major branches, each of which presenting complex interconnections, including ring-like substructures. At the same time, several smaller disconnected components involving strongly interconnected nodes with smaller degree can be also observed. Also of particular interest are the relatively small string component observed at the top left-hand side of the visualization in figure 6, as well as the nodes that remained disconnected given the adopted parameter configuration. Most of these nodes are characterized by low transitivity.

It is also worth comparing the coincidence network obtained for the *US airports* individual autorrelation signatures with that of a BA network, as illustrated in figure 4(b). Though these two networks share some important properties, such as the chained/string-like connections with a gradual progression of the degree of the original nodes, they also present important differences, including the markedly more intricate structure of the *US airports* network for smaller values of original node degree, possibly as a consequence of the more local and geographical nature of the interconnections between smaller airports, thus suggesting that this network involves a mixture of structural organization principles varying along topological scales.

In order to verify the effect of the extension of the considered range of lags taken into account while building the coincidence network from the individual autorrelation signatures, we show in figure 7 respective autorrelation networks derived from the *US airports* individual signatures considering increasing lag intervals. The effect of taking into account longer lag extensions contributes to obtaining progressively more detailed coincidence networks. Interestingly, the complex topological structure of this network could not be discerned from the coincidence similarity networks considering narrow lag extensions. At the same time, the effect of incorporating additional lags tends to be larger for smaller lag extensions, with progressively more similar coincidence networks being obtained as the maximum lag approaches the borders of the network.

The above examples and results illustrate the potential of the autorrelation approach for revealing detailed information about the network being analyzed from the perspective of its hierarchical similarity structure.

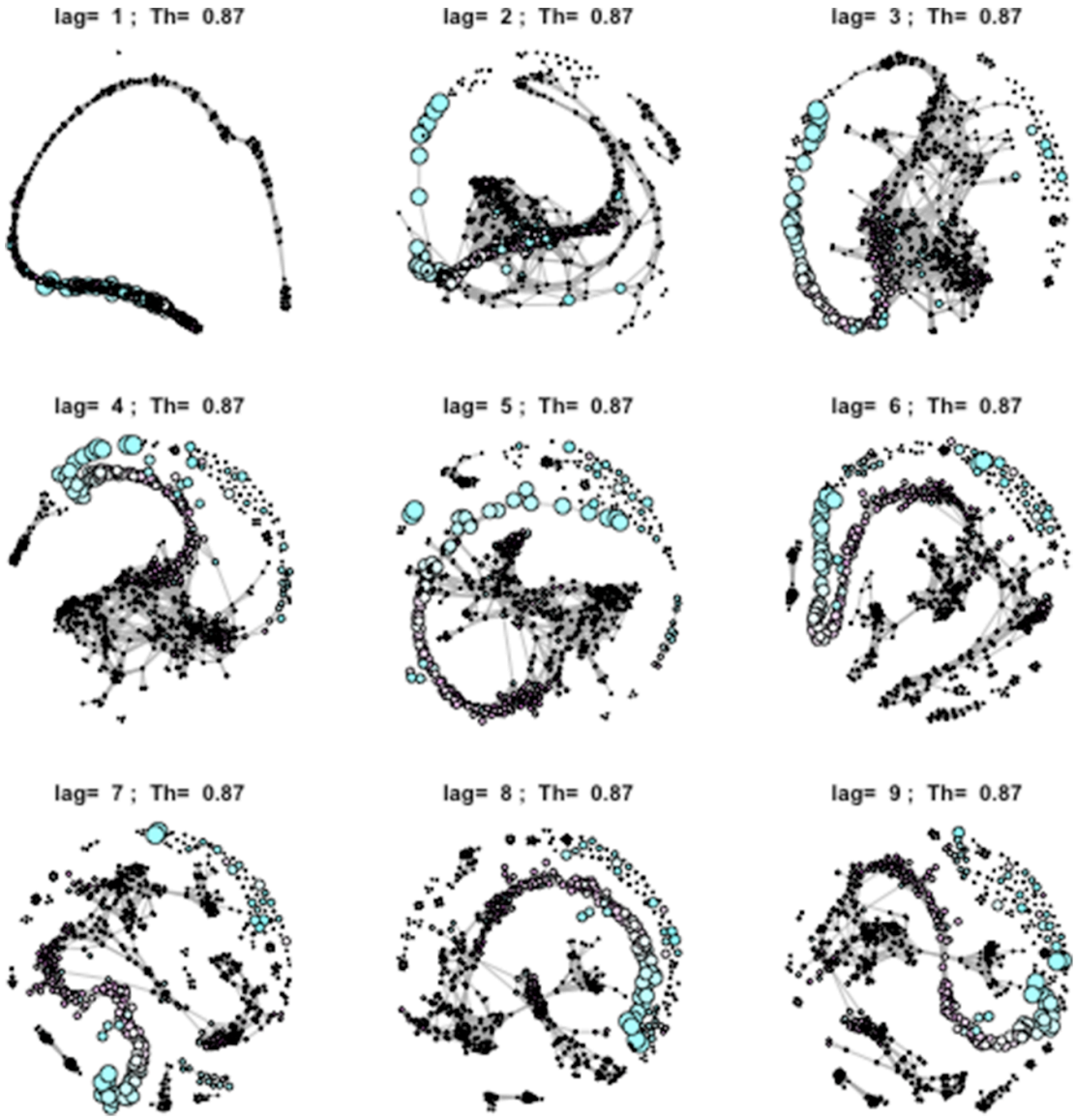


Figure 7. The coincidence networks obtained from the individual autorrelation signatures of the *US airports* network considering successive lags varying from 1 to 9. By supplying additional information about the topology of the original network, networks derived from individual signatures with larger lag values tend to yield progressively more detailed coincidence networks.

To further complement our investigation of the potential of the autorrelation approach to characterizing the individual nodes of complex networks, the obtained individual signatures shown in figure 8 have been organized into three respective clusters for each of the ER, BA and WS networks in figure 1 by employing Ward's hierarchical agglomerative clustering approach (e.g. [65]). Basically, this method successively merge the original patterns into clusters according to a minimum variance principle. The complete individual signatures were considered as features during respective clustering by using the Ward's approach.

Each of the three main clusters identified respectively to each of the three networks in figure 1 are shown as green, blue and salmon in figure 8. The original individual signatures are shown—respectively to the ER (a), BA (d), and WS (g)—in these three colors according to their respective membership in clusters identified by the Ward's methodology. It can be observed that the group marked as green mostly corresponds to the longer individual signatures, while the blue cluster tends to be associated to shorter individual signatures, the blue group corresponding to intermediate-length signatures.

The plates in the intermediate column of figure 8—namely (b), (e), and (h)—show the 2D PCA projections of the respective individual autorrelation signatures. Recall that the identification of the three main clusters was performed while considering the whole signatures, and not their respective 2D PCA projections, which are shown only for visualization purposes.

Figures 8(c), (f), and (i) present visualizations, by using the Fruchterman–Reingold approach (e.g. [66]), of the original ER, BA, and WS networks from which the results in figures 1 and 8 have been obtained, with

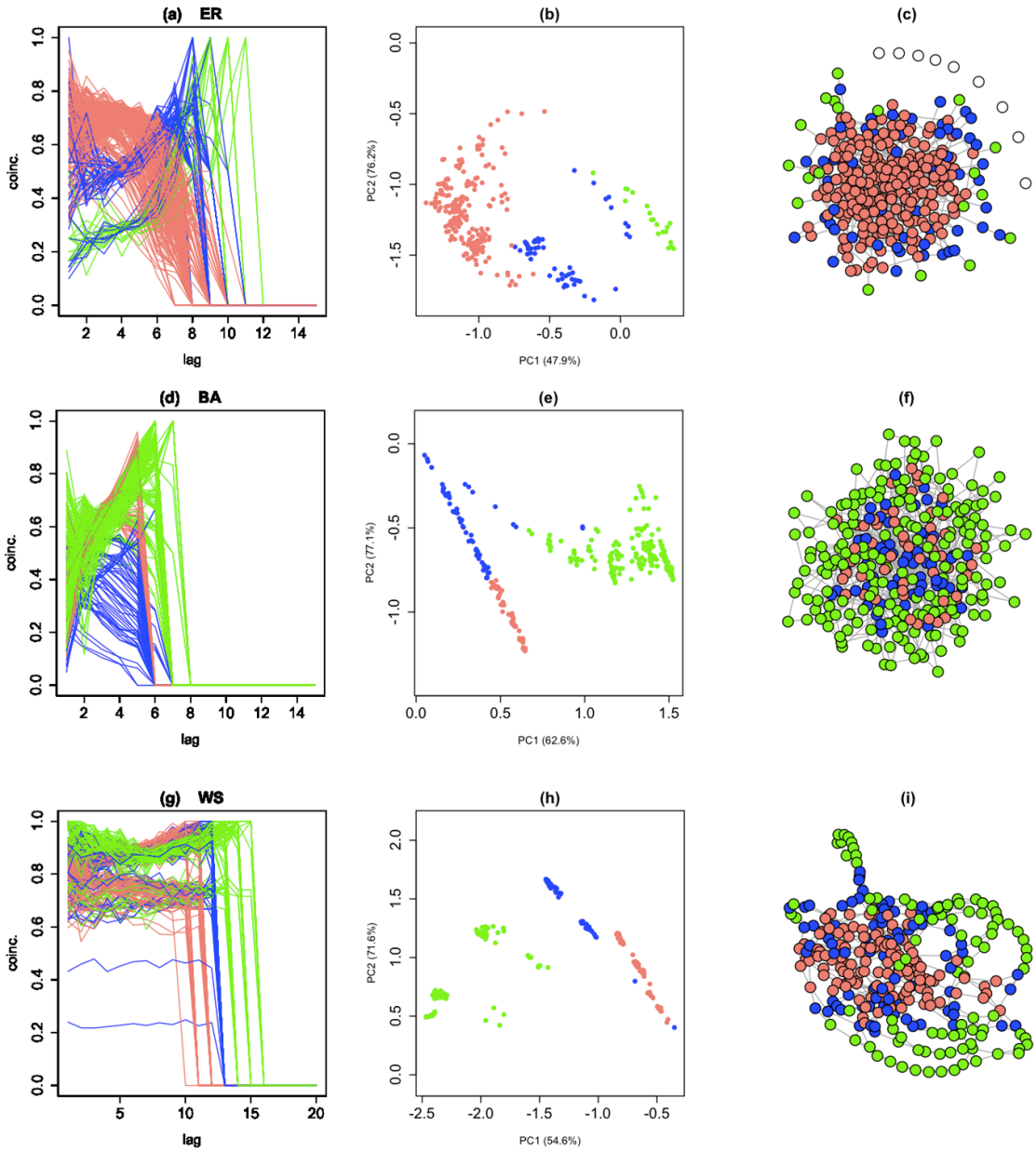


Figure 8. Characterization of nodes associated to the autorrelation curves in figure 1 respectively to the ER (a)–(c), BA (d)–(f), and WS (g)–(i) networks. The original individual signatures are shown in (a), (d) and (g), respectively, for reference. The PCA projections are depicted in (b), (e), and (h). Three main clusters for each of the three groups of signatures have been obtained by applying Ward's hierarchical agglomerative clustering (e.g. [65]) onto the whole original individual signatures, and then used for identifying the nodes in the PCA projections (only for visualization purposes), as well as in the respective networks from which the individual signatures were obtained, shown in (c), (f), and (i). The white nodes in (c) correspond to isolated nodes in the ER network, which have not been taken into account in the analysis. In the three types of networks, a well-defined tendency can be observed of the groups having longer signatures (shown in green) to correspond mostly to the boundary of the original respective networks. At the same time, the nodes associated to shorter individual signatures (shown in salmon) tended to be found at the core of the respective networks, especially in the case of the ER and WS cases. These two observed tendencies are less noticeable in the BA case because of its smaller diameter (7), as well as the more widespread interconnections implemented by the respective hubs, which tend to blur the separation between core and boundary.

the node colors indicating the respective membership relatively to the three main identified clusters. These interesting results indicate that the longer individual autorrelation signatures (green) tended to correspond to the nodes at the boundary of the networks (e.g. [67]), while the shorter signatures (salmon) are mostly related to the respective cores or centers of the networks. The remainder signatures (blue) tended to be associated to intermediate nodes to be found between the respective boundaries and cores. These results indicate, at least for the three considered networks, that the specific characteristics of the individual autorrelation signatures are directly related to the topological situation of the respective nodes within the original network.

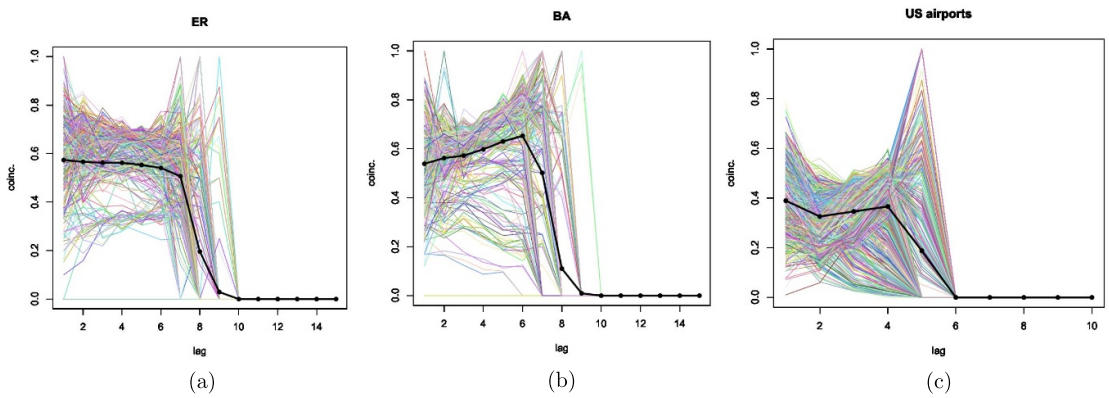


Figure 9. The average (in black) and individual (random colors) cross-relation signatures obtained between the ER network in figure 1(b) and: itself after 500 edge rewirings (a), the original BA network and respective version after 4000 rewirings (b), as well as the *US airports* network and respective network obtained for 3000 rewirings (c).

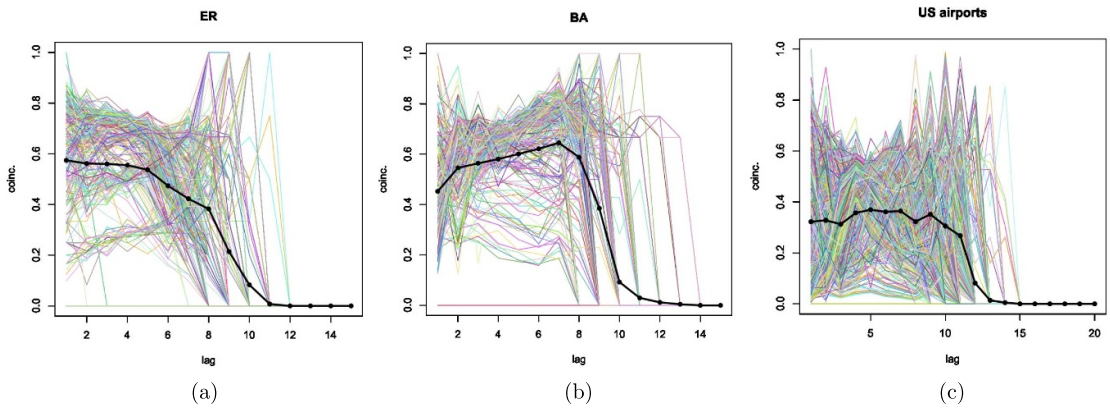


Figure 10. The cross-relation functions between the original ER, BA, and *US airports* networks and their respective versions after removal of the 5% of nodes with the highest degree (hubs).

The remainder of the present section is aimed at providing some results related to the *cross-relation* between two aligned networks respectively to three modifications of the original ER, BA, and WS networks in figure 1.

Examples of *cross-relations* between original and rewired versions of a same network are depicted in figure 9. First, we have the ER network in figure 1(a) before and after 500 rewirings. Interestingly, a noticeable effect of the rewirings on the average signature can be observed, which is possibly a consequence of random fluctuations given the relatively small size of this network. The individual and average autorrelation signatures obtained for the BA network are shown in figure 9(b). The obtained average cross-relation signature is characterized by a less accentuated initial increase, which is a direct consequence of the partial loss of the scale free property, as a consequence of the hubs being rewired into more uniform configurations, as indicated by the associated reduction of the number of outlier (lower) signatures. The rewiring effects on the *US airports* network, shown in figure 9(c), included making the intermediate peak less salient, reducing the extent of the signatures at the same time as they became less dispersed, as well as leading to a sharper final transition.

The cross-relation between networks and their rewired versions is strongly affected by the necessarily implied reduction of the average shortest path length, which substantially narrows the extent of the signatures along the lag δ.

Another interesting study of cross-relations between two graphs consists of considering the original graph and a respective version with a given percentage of hubs removed from the network. Figure 10 illustrates this type of analysis respectively to the removal of 5% of the nodes with the highest degrees respectively to the ER, BA and *US airports* networks.

As could be expected, this modification had almost no effect in the case of the ER network, which does not tend to have nodes with substantially higher degrees (hubs). However, in the case of the BA network, several alterations can be verified as a consequence of the removal of just a few hubs. We can observe a

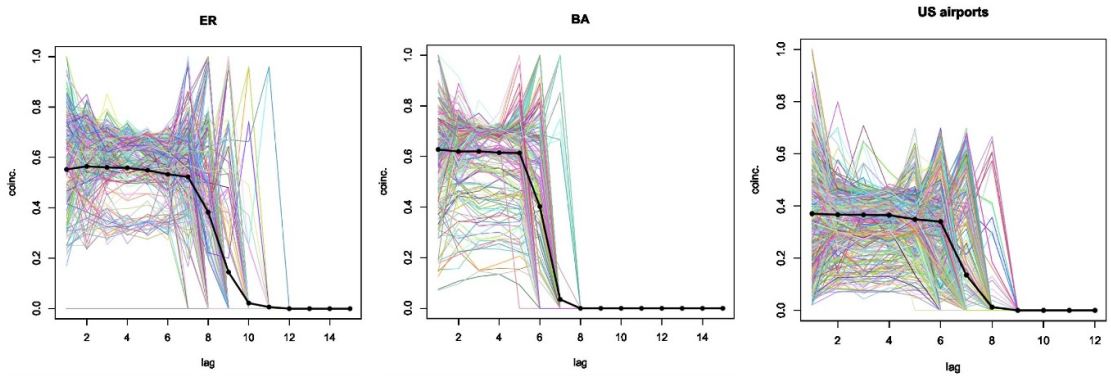


Figure 11. The cross-relation functions between the original ER, BA, and *US airports* networks and their respective versions after permutation of nodes.

reduction of the extent of the signatures, as well as a substantial change in the shape of the average signature, which now corresponds mostly to a plateau instead of a peak. The effects on the *US airports* network are also strong, implying longer signatures as well as a reduction of the peak and change of the shape of most signatures. This result indicates that the hubs in this network tend to play an important role in defining its respective topological properties.

Figure 11 presents the cross-relation analysis of networks under another type of modification, now involving complete permutation of the labels of all nodes respectively to the original network. The effect of this alteration was similar in the case of the ER, BA, and *US airports* networks, consisting in making the signatures more similar.

The above cross-relation examples illustrate the potential of this concept for comparing between two aligned networks while considering respective displacements along the distance lags. The obtained results indicate that the cross-relation average and individual signatures has potential for providing interesting information about the topological properties of the analyzed structures, especially regarding the presence of hubs and the uniformity of the topology around the nodes at successive topological scales.

5. Concluding remarks

An important class of associations between a given pair of mathematical structures regards how much they are interrelated in terms of a relative displacements, controlled by respective lag values. The importance of this type of relationship is reflected in the key role that the related operations of autocorrelation and cross-correlation play in areas including signal and image processing and analysis, control systems, computer vision, neuronal networks and AI.

In the present work, we aimed at developing operations that would be analogue to the autocorrelation and cross-correlation between functions, fields, vectors and matrices, but adapted to the specific nature of graphs and networks. In order to do so, we resourced to the multiset coincidence similarity, which implements a strict and stable quantification of the similarity between any two mathematical structures.

More specifically, we defined a sequence of hierarchical layers around each network node as being analogous to the lag parameter in the classic correlation. Therefore, it became possible to compare the topological features each reference node with those of nodes along the nodes at increasing topological distances, which has been done in terms of the respective average similarity. This operation yields a function of the topological distance from the successive layers, which acts as the lag parameter. The possibility to choose among several possible topological, or other types of features accounts for additional flexibility in the implemented analysis, which can be customized to the aspects of specific interest in each research situation. Once these measurements have been obtained, the estimation of the autorrelation and cross-relation involves relatively small computational cost. In the case of the present work, which adopted the node degree and transitivity as features, the overall cost resulted particularly low given that these measurements are local and can be effectively calculated.

As with the traditional autocorrelation and cross-correlation between two functions, the autorrelation function has its maximum value observed for lag zero. As it has been indicated by several examples involving several model-theoretical and real-world networks, the average and individual autorrelation and cross-relation signatures can reflect several important aspects of the analyzed networks, including the

presence of hubs, overall regularity, scale-free behavior, presence of outliers, as well as the extent of topological scales along which the network is more or less strongly mutually similar. Of particular interest is the intricate coincidence similarity structure obtained from the individual autorrelation signatures which, in the case of the *US airports* network, revealed complex interrelationships between involved signatures, confirming the particularly complex organization of this network. The relationship between the shape of the individual autorrelation signatures obtained for each node and the topological characteristics of these nodes was also studied by using a clustering methodology, with interesting results.

Supplementary material related to the present work can be found at https://github.com/ldafcosta/compl_autorr.

The reported concepts and methods, which are respective to the adopted models, assumptions, and parametric configurations, pave the way to several interesting further complementary developments, including the consideration of other node measurements, alternative model-theoretical networks (especially heterogeneous modular networks), as well as the investigation of other parameter configurations (e.g. average degree and number of nodes). Other promising possibilities concern the application to additional real-world networks. In particular, it would be interesting to study the interesting subject of the scale free property (e.g. [68]) from the perspective of the concepts and methods suggested in the present work.

Data availability statement

No new data were created or analyzed in this study.

Acknowledgment

Luciano da F Costa thanks CNPq (Grant No. 307085/2018-0) and FAPESP (Grant 15/22308-2).

ORCID iD

Luciano da Fontoura Costa https://orcid.org/0000-0001-5203-4366

References

[1] West D B *et al* 2001 *Introduction to Graph Theory* vol 2 (Upper Saddle River, NJ: Prentice Hall)
[2] Bollobás B 1998 *Modern Graph Theory* vol 184 (New York: Springer Science & Business Media)
[3] Bondy J A and Murty U S R 1976 *Graph Theory with Applications* (London: McMillan)
[4] Barabási A L and Pósfai M 2016 *Network Science* (Cambridge: Cambridge University Press)
[5] Newman M 2010 *Networks: An Introduction* (Oxford: Oxford University Press)
[6] da Fontoura Costa L, Rodrigues F A, Travieso G and Villas Boas P R 2007 Characterization of complex networks: a survey of measurements *Adv. Phys.* **56** 167–242
[7] da Fontoura Costa L, Oliveira Jr O N, Travieso G, Rodrigues F A, Villas Boas P R, 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
[8] da Fontoura Costa L 2018 What is a complex network? (available at: www.researchgate.net/publication/324312765_What_is_a_Complex_Network_CDT-2) (Accessed 5 May 2019)
[9] da Fontoura Costa L and Domingues G 2022 Cost-based approach to complexity: a common denominator? *Rev. Bras. Ensino Fís.* **44** 1–14
[10] Knuth D E 1998 *The Art of Computing* (Reading, MA: Addison Wesley)
[11] Wirth N 1985 *Algorithms & Data Structures* (Englewood Cliffs, NJ: Prentice-Hall)
[12] Cormen T H 2009 *Introduction to Algorithms* (Cambridge, MA: MIT Press)
[13] Brigham E O 1988 *Fast Fourier Transform and Its Applications* (New York: Pearson)
[14] da Fontoura Costa L 2009 *Shape Classification and Analysis: Theory and Practice* 2nd edn (Boca Raton, FL: CRC Press)
[15] Vetterli M, Kovačević J and Goyal V K 2014 *Foundations of Signal Processing* (Cambridge: Cambridge University Press)
[16] Haykin S 2013 *Neural Networks and Learning Machines* 9th edn (New York: McGraw-Hill Education)
[17] Gallager R G 2014 *Stochastic Processes: Theory for Applications* (Cambridge: Cambridge University Press)
[18] Gonzalez R C and Woods R E 2018 *Digital Image Processing* (New York: Pearson)
[19] Horn B K P 1986 *Robot Vision* (Cambridge: McGraw Hill)
[20] Davies E R 2005 *Machine Vision* (Amsterdam: Morgan Kaufmann)
[21] Shuman D I, Ricaud B and Vandergheynst P 2012 A windowed graph Fourier transform *2012 IEEE Statistical Signal Processing Workshop (SSP)* (IEEE) pp 133–6
[22] Sardellitti S, Barbarossa S and Di Lorenzo P 2017 On the graph Fourier transform for directed graphs *IEEE J. Sel. Top. Signal Process.* **11** 796–811
[23] Ricaud B, Borgnat P, Tremblay N, Gonçalves P and Vandergheynst P 2019 Fourier could be a data scientist: from graph Fourier transform to signal processing on graphs *C. R. Physique* **20** 474–88
[24] Demaine E D, Emanuel D, Fiat A and Immorlica N 2006 Correlation clustering in general weighted graphs *Theor. Comput. Sci.* **361** 172–87
[25] Hero A and Rajaratnam B 2012 Hub discovery in partial correlation graphs *IEEE Trans. Inf. Theory* **58** 6064–78
[26] Bansal N, Blum A and Chawla S 2004 Correlation clustering *Mach. Learn.* **56** 89–113
[27] Berg J and Lässig M 2002 Correlated random networks *Phys. Rev. Lett.* **89** 228701

[28] Fujita A, Takahashi D Y, Balardin J B, Vidal M C and Sato J R 2017 Correlation between graphs with an application to brain network analysis *Comput. Stat. Data Anal.* **109** 76–92
[29] da Fontoura Costa L 2021 Further generalizations of the Jaccard index (available at: www.researchgate.net/publication/355381945_Further_Generalizations_of_the_Jaccard_Index) (Accessed 21 August 2021)
[30] da Fontoura Costa L 2022 On similarity *Physica* A **599** 127456
[31] da Fontoura Costa L 2022 Coincidence complex networks *J. Phys. Complex.* **3** 015012
[32] da Fontoura Costa L 2021 Comparing cross correlation-based similarities (available at: www.researchgate.net/publication/355546016_Comparing_Cross_Correlation-Based_Similarities) (Accessed 21 October 2021)
[33] da Fontoura Costa L 2021 Multiset neurons (available at: www.researchgate.net/publication/356042155_Multiset_Neurons)
[34] Akbas C E, Bozkurt A, Arslan M T, Aslanoglu H and Cetin A E 2014 L1 Norm Based Multiplication-Free Cosine Similarity Measures for Big Data Analysis *IEEE Computational Intelligence for Multimedia Understanding (IWCIM)* (*Paris, France, November 2014*) pp 1–4
[35] Mirkin B 1996 *Mathematical Classification and Clustering* (Dordrecht: Kluwer Academic Publishers)
[36] Travençolo B A N and da Fontoura Costa L 2008 Hierarchical spatial organization of geographical networks *J. Phys. A: Math. Theor.* **41** 224004
[37] da Fontoura Costa L and Silva F N 2006 Hierarchical characterization of complex networks *J. Stat. Phys.* **125** 841–72
[38] da Fontoura Costa L and Andrade R F S 2007 What are the best concentric descriptors for complex networks? *New J. Phys.* **9** 311
[39] Wasserman S and Faust K 1994 *Social Network Analysis: Methods and Applications* (Cambridge: Cambridge University Press)
[40] Newman M E J 2003 Ego-centered networks and the ripple effect *Soc. Netw.* **25** 83–95
[41] Dehmer M 2008 Information processing in complex networks: graph entropy and information functionals *Appl. Math. Comput.* **201** 82–94
[42] Jaccard P 1901 Distribution de la flore alpine dans le bassin des dranses et dans quelques régions voisines *Bull. Soc. Vaud. Sci. Nat.* **37** 241–72
[43] Wikipedia 2021 Jaccard index (available at: https://en.wikipedia.org/wiki/Jaccard_index) (Accessed 10 October 2021)
[44] Murphy A H 1996 The Finley affair: a signal event in the history of forecast verification *Weather Forecast.* **11** 3–20
[45] da Fontoura Costa L 2022 On the self-coincidence structure of networks (available at: www.researchgate.net/publication/361182049_On_the_Self-Coincidence_Structure_of_Networks)
[46] Hein J 2003 *Discrete Mathematics* (Massachusetts: Jones & Bartlett Pub.)
[47] Blizard W D 1988 Multiset theory *Notre Dame J. Form. Log.* **30** 36–66
[48] Blizard W D 1991 The development of multiset theory *Mod. Log.* **4** 319–52
[49] Mahalakshmi P M and Thangavelu P 2019 Properties of multisets *Int. J. Innov. Technol. Explor. Eng.* **8** 1–4
[50] Singh D, Ibrahim M, Yohana T and Singh J N 2011 Complementation in multiset theory *Int. Math. Forum* **38** 1877–84
[51] Jaccard P 1901 Étude comparative de la distribution florale dans une portion des alpes et des jura *Bull. Soc. Vaud. Sci. Nat.* **37** 547–9
[52] Samanthula B K and Jiang W 1989 Secure multiset intersection cardinality and its application to Jaccard coefficient *IEEE Trans. Dependable Secure Comput.* **13** 591–604
[53] Schubert A and Telcs A 2014 A note on the Jaccardized Czekanowski similarity index *Scientometrics* **98** 1397–9
[54] Vijaymeena M K and Kavitha K 2016 A survey on similarity measures in text mining *Mach. Learn. Appl.* **3** 19–28
[55] da Fontoura Costa L and Tokuda E K 2022 A similarity approach to cities and features *Eur. Phys. J.* B **95** 155
[56] Domingues G S, Tokuda E K and da Fontoura Costa L 2022 Identification of city motifs: a method based on modularity and similarity between hierarchical features of urban networks *J. Phys. Complex.* **3** 045003
[57] Lovász L 1993 Random walks on graphs *Combinatorics, Paul Erdos is Eighty* vol 2 (Hungary: Bolyai Society Mathematical Studies) p 4
[58] Noh J D and Rieger H 2004 Random walks on complex networks *Phys. Rev. Lett.* **92** 118701
[59] Kullback S and Leibler R A 1951 On information and sufficiency *Ann. Math. Stat.* **22** 79–86
[60] Johnson R A and Wichern D W 2002 *Applied Multivariate Analysis* (Englewood Cliffs, NJ: Prentice-Hall)
[61] Gewers F, Ferreira G R, Arruda H F, Silva F N, Comin C H, Amancio D R and da Fontoura Costa L 2020 Principal component analysis: a natural approach to data exploration *ACM Comput. Surv.* **54** 200–19
[62] Zachary W W 1977 An information flow model for conflict and fission in small groups *J. Anthropol. Res.* **33** 452–73
[63] Csardi G and Nepusz T 2006 The igraph software package for complex network research *InterJ. Complex Syst.* **1695** 1–9
[64] Négyessy L, Nepusz T, Kocsis L and Bazsó F 2006 Prediction of the main cortical areas and connections involved in the tactile function of the visual cortex by network analysis *Eur. J. Neurosci.* **23** 1919–30
[65] Duda R O, Hart P E and Stork D G 2000 *Pattern Classification* (New York: Wiley Interscience)
[66] Fruchterman T M, Thomas M J and Reingold E M 1991 Graph drawing by force-directed placement *Softw. Pract. Exp.* **21** 1129–64
[67] Travençolo B A N, Viana M P and da Fontoura Costa L 2009 Border detection in complex networks *New J. Phys.* **11** 063019
[68] Broido A D and Clauset A 2019 Scale-free networks are rare *Nat. Commun.* **10** 1–10