A graph complexity measure based on the spectral analysis of the Laplace operator

Diego M. Mateos [1,2,3*], Federico Morana[3], and Hugo Aimar[1,3]

[1]Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Argentina.
[2]Facultad de Ciencia y Tecnología. Universidad Autónoma de Entre Ríos (UADER). Oro Verde, Entre Ríos, Argentina.
[3]Instituto de Matemática Aplicada del Litoral-UNL-CONICET (IMAL), CCT CONICET, Santa Fé, Argentina.
[*]Corresponding author: Diego M. Mateos, mateosdiego@gmail.com.

October 25, 2022

Graph, Complexity, Laplacian, Spectral analysis

1 Introduction

In this work we introduce a concept of complexity for undirected graphs in terms of the spectral analysis of the Laplacian operator defined by the incidence matrix of the graph. Precisely, we compute the norm of the vector of eigenvalues of both the graph and its complement and take their product. Doing so, we obtain a quantity that satisfies two basic properties that are the expected for a measure of complexity. First, complexity of fully connected and fully disconnected graphs vanish. Second, complexity of complementary graphs coincide. This notion of complexity allows us to distinguish different kinds of graphs by placing them in a "croissant-shaped" region of the plane link density - complexity, highlighting some features like connectivity, concentration, uniformity or regularity and existence of clique-like clusters. Indeed, considering graphs with a fixed number of nodes, by plotting the link density versus the complexity we find that graphs generated by different methods take place at different regions of the plane.

2 The mathematical model

Let $G = (\mathcal{V}, \mathcal{E}, \mathcal{W})$ be a simple undirected graph, where $\mathcal{V} = \{1, \dots, n\}$ is the set of vertices or nodes, $\mathcal{E} = \{e_1, \dots, e_m\} \subset \{\{i, j\} \colon i, j \in \mathcal{V}\}$ is the set of edges and $\mathcal{W} \colon \mathcal{V} \times \mathcal{V} \to \{0, 1\}$ is the adjacency matrix of G with $w_{ij} = 1$ whenever $\{i, j\} \in \mathcal{E}$ and zero otherwise. Since the graph is undirected and simple the matrix $\mathcal{W}$ is symmetric with null diagonal. We will denote $i \sim j$ when $\{i, j\} \in \mathcal{E}$.

The degree of a vertex j is defined by $\delta(j) = \sum_{i \in \mathcal{V}} w_{ij}$. The degree matrix is defined as the diagonal $n \times n$ matrix containing the degrees of the nodes and denoted by $D = diag(\delta(1), \dots, \delta(n))$. The Laplacian of the graph is the lineal operator acting on real or complex functions defined on the nodes, with matrix given by

$$\Delta = \mathcal{W} - D. \tag{1}$$

This operator is symmetric and negative semi-definite. Therefore we can apply the spectral theorem to obtain an orthonormal basis of $\ell^2(\mathcal{V}) \sim \mathbb{R}^n$ of eigenvectors $\{\psi_1, \dots, \psi_n\}$ of Δ. It is usually called the Fourier basis of G. The associated eigenvalues $\{\lambda_1, \dots, \lambda_n\}$ satisfy $0 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. In the following we will refer to the vector $\overline{\lambda} = (\lambda_1, \dots, \lambda_n)$ as the spectrum of the graph G. The trace of the Laplacian is a feature of interest for our further analysis and is given by $\sum_{i=1}^{n} \lambda_i = -2m$, where m is the number of edges of G. For a general reference regarding the spectral theory of the Laplacian on graphs see [1] and references therein.

We may consider equivalent two graphs G and G' that share the spectrum $\bar{\lambda}$. Hence, for G and H two graphs with the same number n of vertices, the function $d_s(G, H) = \left\| \bar{\lambda}_G - \bar{\lambda}_H \right\|$, with $\bar{\lambda}_G$ and $\bar{\lambda}_H$ the spectrum of G and H respectively and $\|\cdot\|$ any norm in $\mathbb{R}^n$, is a distance (metric) between the classes of co-spectrality of G and H. We shall take the usual (euclidean) norm $\left\| \bar{\lambda} \right\| = \left(\sum_{i=1}^{n} |\lambda_i|^2 \right)^{1/2}$. We shall refer to d_s as the spectral distance. Notice that since the first eigenvalue λ_1 of each graph vanishes, we actually have that $d_s(G, H) = |\bar{\Lambda}_G - \bar{\Lambda}_H|$,

where $\bar{\Lambda} = (\lambda_2, ..., \lambda_n)$ and $|\cdot|$ is the euclidean norm in $\mathbb{R}^{n-1}$. The spectral distance on graphs was considered before in [2], see also [3].

In order to introduce our definition of spectral complexity of a graph, let us set Z to denote the null graph, i.e. $w_{ij} = 0$ for every $i, j \in \mathcal{V}$, and F to denote the complete graph, i.e. $w_{ij} = 1$ for every $i \neq j$. Now we can define the **spectral complexity** of a graph G with n vertices as

$$\begin{aligned} \mathcal{C}_s(G) &= d_s(G, Z) \cdot d_s(G, F) \\ &= \left\|\bar{\lambda}_G - \bar{\lambda}_Z\right\| \ \left\|\bar{\lambda}_G - \bar{\lambda}_F\right\| \\ &= |\bar{\Lambda}_G - \bar{\Lambda}_Z| \ |\bar{\Lambda}_G - \bar{\Lambda}_F|. \end{aligned} \tag{2}$$

Two basic premises are behind this definition. The first one is that both, the null graph and the full graph, are the less complex graphs that can be defined on the vertices set $\mathcal{V} = \{1, ..., n\}$. The second is that complementary graphs should have the same complexity.

A second quantity associated to a graph that we shall take into account in our analysis is its link density. The link density ρ of a simple unidirected graph is the number of actual edges divided by the number of all possible edges. With our notation

$$\rho(G) = \frac{2m}{n(n-1)}. \tag{3}$$

Given a positive integer n we shall display all the possible graphs G built on $\mathcal{V} = \{1, 2, ..., n\}$ in the plane of the variables $\rho(G)$ and $\mathcal{C}_s(G)$. Since the density of a graph G and the density of its complement can be quite different, actually $\rho(G) + \rho(G^c) = 1$, it is clear that the link density is not a function of the spectral complexity. It is also simple to show that graphs with the same density may have different spectral complexity. So neither ρ is a function of $\mathcal{C}_s$ nor $\mathcal{C}_s$ is a function of ρ. As could be expected. Nevertheless ρ and $\mathcal{C}_s$ are not completely independent. In fact we empirically determine the region in the region in the plane $(\rho, \mathcal{C}_s)$ spanned by all possible graphs.

The delimitation of the region in the representation plane link *density - complexity* where all the variety of graphs take place is not a trivial task to perform theoretically. Here we obtain an empirical approximation of the upper and lower boundaries, derived after placing a wide variety of graphs generated by random and deterministic methods. The plane obtained is a croissant-shaped. Figure 1 depicts the croissant-shaped and the placement of some paradigmatic graphs of 15 vertices. This notion of plane highlight some features like connectivity, concentration, uniformity or regularity and existence of clique-like clusters

3 Result

Using the plane ρ *vs* $\mathcal{C}_s$, we analysed three well known stochastic models, the Erdös-Rényi model [4], the Watts-Strogatz model [5], and the Barabási-Albert model [6] for the different parameter each one. As we shall see, each of them draws some characteristic pattern contained in the basic croissant shape (see Figure 1b). Finally, as an application to graphs generated by real measurements, we consider the brain connectivity graphs from two epileptic patients obtained from magnetoencephalography (MEG) recording, both in a baseline period and in ictal periods (epileptic seizures). In this case, our definition of complexity could be used as a tool for discerning between states, by the analysis of differences at distinct frequencies of the MEG recording.

References

[1] M. M. Bronstein, J. Bruna, Y. LeCun, A. Szlam, and P. Vandergheynst. Geometric deep learning: Going beyond euclidean data. *IEEE Signal Processing Magazine*, 34(4):18–42, 2017.

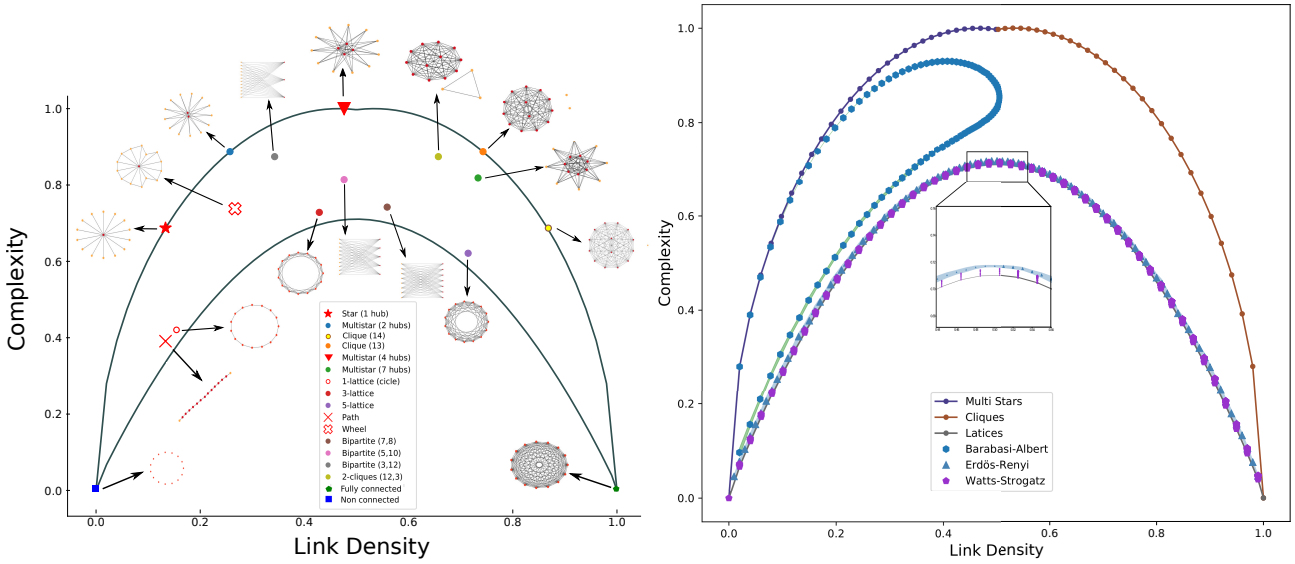
[2] Jiao Gu, Bobo Hua, and Shiping Liu. Spectral distances on graphs. *Discrete Applied Mathematics*, 190:56–74, 2015.

[3] M.M. Deza and E. Deza. *Encyclopedia of Distances*. Springer-Verlag Berlin Heidelberg, 4th edition, 2016.

[4] E. N. Gilbert. Random graphs. *The Annals of Mathematical Statistics*, 30(4):1141–1144, 1959.

[5] D. J. Watts and S. H. Strogatz. Collective dynamics of 'small-world'networks. *nature*, 393(6684):440–442, 1998.

[6] A-L Barabási and R Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.



(a) (b)

Figure 1: Spectral complexity vs link density plane. The full line represent the lower and upper limit. (a) Schematic distribution of the different types of networks for n = 15 on the "croissant-shaped" region. (b) Overview and comparison of the results obtained for all the network models analysed in this work. In this case we use for all models n = 100