

Neural Network Spectral Robustness under Perturbations of the Underlying Graph

Anca Rădulescu

radulesa@newpaltz.edu

*Department of Mathematics, State University of New York at New Paltz,
New Paltz, NY 12561, U.S.A.*

Recent studies have been using graph-theoretical approaches to model complex networks (such as social, infrastructural, or biological networks) and how their hardwired circuitry relates to their dynamic evolution in time. Understanding how configuration reflects on the coupled behavior in a system of dynamic nodes can be of great importance, for example, in the context of how the brain connectome is affecting brain function. However, the effect of connectivity patterns on network dynamics is far from being fully understood. We study the connections between edge configuration and dynamics in a simple oriented network composed of two interconnected cliques (representative of brain feedback regulatory circuitry). In this article our main goal is to study the spectra of the graph adjacency and Laplacian matrices, with a focus on three aspects in particular: (1) the sensitivity and robustness of the spectrum in response to varying the intra- and intermodular edge density, (2) the effects on the spectrum of perturbing the edge configuration while keeping the densities fixed, and (3) the effects of increasing the network size. We study some tractable aspects analytically, then simulate more general results numerically, thus aiming to motivate and explain our further work on the effect of these patterns on the network temporal dynamics and phase transitions. We discuss the implications of such results to modeling brain connectomics. We suggest potential applications to understanding synaptic restructuring in learning networks and the effects of network configuration on function of regulatory neural circuits.

1 Introduction ---

1.1 Network Architecture and Dynamics. The study of networks has been the subject of great interest in recent research. Many natural systems are organized as networks, in which the nodes (be they cells, individuals, or web servers) interact in a time-dependent fashion.

One of the particular points of interest has been the question of how the hardwired structure of a network (its underlying graph) affects its function, for example, in the context of optimal information storage or

transmission between nodes along time (Bullmore & Sporns, 2009). It has been hypothesized that there are two key conditions for optimal function in such networks: a well-balanced adjacency matrix (the underlying graph should appropriately combine robust features and random edges) and well-balanced connection strengths, driving optimal dynamics in the system. A subsequent line of study is to understand the effects of connectivity patterns on the temporal behavior of the network, seen as a dynamical system in which the node variables are coupled according to a connectivity scheme that obeys certain deterministic constraints but also incorporates random aspects. One can then investigate how the phase space dynamics (and the phase transitions that the system undergoes under perturbation) are affected when perturbing the underlying adjacency graph.

While the general aim of our work is to study the relationship between a network's hardwired circuitry and its dynamics, this article focuses primarily on understanding the robustness of certain graph-theoretical features as the network is perturbed or as its size increases—and, to a lesser extent, discusses their potential to further affect the vulnerability or robustness of the system's dynamics (the subject of a related paper: Rădulescu & Verduzco-Flores, 2015). To fix out ideas, we investigate here a bimodular, oriented graph, in which the nodes of two modules connect through fixed numbers of random edges within each module, as well as across modules. Using analytical and numerical computations, we aim to establish whether and why, when fixing the number of both intra- and intermodular edges, the adjacency spectrum of the network remains in general sufficiently robust under particular edge configurations (geometries). This is important, since it would suggest that certain dynamic algorithms in this type of networks may also remain unaffected by such constrained geometry changes.

A very large body of work addresses properties of random matrices (Tao, 2012), that is, matrices whose entries are drawn independently out of a given (typically normal) probability distribution. If, in addition, the matrix represents the adjacency of a random graph, so that each entry equals 1 with a given probability, there are classical methods used when looking for properties of the spectrum (e.g., spectral radius or spectral density). For example, a popular and well-known generative model developed in the early 1980s (Holland, Laskey, & Leinhardt, 1983), whose extensions for directed and weighted edges are still widely used, is the stochastic block model (SBM; Aicher, Jacobs, & Clauset, 2013; Larremore, Clauset, & Jacobs, 2014). SBMs are a powerful way of encoding specific assumptions about the way unknown parameters interact to create edges, and offer many advantageous features. We argue, however, that generative models do not account (mathematically or realistically) for all potential network connectivity schemes that arise in natural systems.

A lot remains to be clarified in terms of how individual (rather than stochastic) changes in simple parameters affect the properties of a graph. Such deterministic constraints are important, since they are likely to appear

in natural systems in conjunction with the stochastic aspects controlling the particular edge distribution. Our model differs from most of the generative approaches in that it conserves the number of edges within and between modules rather than fixing independently the probability of having an edge that connects two given nodes in the same or different modules. We chose the edge density as our parameter of interest, since fixing the total number of node-to-node connections represents a typical normalization condition in networked systems (e.g., a constant sum of weighted edges is a popular normalization scheme in the context of synaptic updating in a neural network (von der Malsburg, 1973)).

While classical results (such as Wigner's semicircle law) apply in the stochastic case with independent and identically distributed edges, various extensions have been worked out for models that do not necessarily abide by these properties. Consider, for example, the configuration model (Farkas, Derényi, Barabási, & Vicsek, 2001), whose spectral properties have been addressed by numerous studies. Since its edges are not statistically independent, a direct analytical approach is very difficult; existing results range from approximating the full spectrum (Dorogovtsev, Goltsev, Mendes, & Samukhin, 2003) to formally deriving the expected values of the leading eigenvalue, but only in the large N limit (Chung, Lu, & Vu, 2003). In a recent paper, Nadakuditi and Newman (2013) took an indirect approach: they considered a model with the same degree sequence as the configuration model, but in which the number of edges between any two nodes was drawn independently from a Poisson distribution. Then the spectra of the two models was shown to agree in the large N limit. In section 4, we apply this idea in our case by carrying out a large N limit comparison between our model and its probabilistic counterpart, with independent, stochastic edges, as Nadakuditi and Newman (2012) considered.

1.2 Brain Function from Graph Theory to the Connectome. In the context of the brain, a network of nodes connected by oriented weighted edges may constitute a valid representation of neural architecture at more than one spatiotemporal scale (the brain "fractal" possibly reusing similar organizational and optimization principles at multiple complexity levels). For example, one may think of the nodes as individual neurons connected by synapses, whose activity (e.g., measured as variations in membrane potential of single cells) is determined by the external input together with the pattern and strengths of synaptic coupling. At a coarser level, one may view a node as a synchronized population of neurons, whose activity (e.g., measured as mean field firing rate) is determined by the external input, as well as its mean field connections from other—excitatory or inhibitory—populations. At the macroscopic level, consistent with imaging techniques such as MRI or EEG, one may think of the nodes as anatomical brain regions (e.g., amygdala, prefrontal cortex) whose activity (e.g., measured as event-related potentials or blood-oxygen-level dependent signals) is determined

by the external stimulus and the interregional connectivity patterns and connection strengths. The way in which various parts of the brain (from the microscale of neurons to the macroscale of functional regions) are wired together is one of the great scientific challenges of this century, currently being addressed by large-scale research collaborations, such as the Human Connectome Project (Toga, Clark, Thompson, Shattuck, & Van Horn, 2012; Craddock et al., 2013).

Recent studies have used graph-theoretical approaches to investigate brain networks, not only in the context of learning, memory formation, and cognitive performance, but also to understand more general organizational and functional principles used by the brain and interpret the effects of different connectivity patterns between the brain's coupled components (Bullmore & Sporns, 2009; Sporns, 2002, 2011). With nodes and edges defined at various scales, according to different empirical modalities (Sporns, 2010), these studies support certain generic topological properties of brain architecture, such as modularity, small worldness, the existence of hubs, and other connectivity density patterns (He & Evans, 2010). These properties, if proven consistent with physiological, behavioral, or genetic factors, may provide us with a better understanding of neural processes and may be effective as biomarkers for behavioral traits or neuropsychiatric conditions.

One thought of potential importance to us is that while the brain itself is a gigantic and relatively densely connected network of billions of neuron nodes (each receiving and providing input to tens of thousands of other nodes), it may be both realistic and computationally advantageous to view the brain as a highly hierarchic network in which the behavior of each "node" at a certain complexity level integrates the behavior of a collection of lower-level nodes. Hence, at each complexity level, the size of the networks we need to study experimentally, represent theoretically, or simulate numerically may be in fact relatively small (a few hundred nodes). For example, at the macroscopic level, compatible with imaging techniques in humans, a small region such as the amygdala is (within typical fMRI acquisitions parameters) as large as 100 to 200 voxel-nodes. For relatively small networks, the traditional large size limit results obtained in random graph theory may no longer apply directly, and new approaches need to be created to extend the results (see section 4.1 for a more detailed discussion).

A recent trend in human imaging research has been oriented toward developing and using graph-theoretical network measures in conjunction with empirical data in order to identify the effects of abnormal connectivity patterns on the efficiency of brain function. By inferring graph-theoretical information from tractography data and temporal time series, various studies have been investigating the sensitivity of systems to removing or adding nodes or edges to different places in the network structure. Data sets have been mined for global (e.g., characteristic path length, clustering coefficient, small-world ratio parameter) and local (e.g., nodal-betweenness centrality, nodal path length, nodal clustering coefficient) network measures

that would optimally differentiate between subject behavioral profiles (GadElkarim et al., 2014; Ajilore et al., 2013). For example, both resting-state tractography-derived measures have been used to understand behavioral impairments in subjects with compromised connectivity due to existing lesions (Corbetta, 2012) or group differences between healthy controls and patients with mental illnesses associated with abnormal feedback circuitry (Fekete et al., 2013; Leow et al., 2013).

However, no matter how well designed or statistically powerful, purely empirically based analyses cannot explain in and of themselves the mechanisms by which connectivity patterns actually act to change the system's dynamics, and thus the observed behavior. Substantial research effort is being directed toward constructing an underlying network model that is tractable theoretically or numerically and could therefore be used in conjunction with the data toward interpreting the empirical results and making further predictions. To this aim, the theoretical dependence of dynamics on connectivity (e.g., in the context of stability and synchronization in networks of coupled neural populations) has been investigated both analytically and numerically, in a variety of contexts—from biophysical models (Gray & Robinson, 2009) to simplified systems (Siri, Quoy, Delord, Cessac, & Berry, 2007). These analyses revealed a rich range of potential dynamic regimes and transitions (Brunel, 2000), shown to depend as much on the coupling parameters of the network as on the arrangement of the excitatory and inhibitory connections (Gray & Robinson, 2009).

The successful construction of a useful (biophysical) computational model seems therefore contingent on our understanding of the control that a network's architecture exercises on its functional dynamic regime. In our work, we are trying to address this question starting with simpler networks and investigating which properties are preserved or emerge in increasingly complex systems. A relatively simple yet general example that we have been considering in previous work is that of an oriented network of two interconnected cliques. For this type of graph geometry and for nodes acting as nonlinear oscillators, we studied how the two intermodular edge densities affect dynamics (Rădulescu & Verduzco-Flores, 2015).

We used this architecture because it is representative of two interacting excitatory and inhibitory densely packed populations, a feedback scheme that provides the underlying control for many brain regulatory loops. For example, this setup successfully informed our human imaging results in the amygdala-prefrontal circuit regulating human emotion (Rădulescu & Mujica-Parodi, 2013). Also, while there are clearly better measures of architecture complexity in a network than edge density, our work was directly motivated by existing hypotheses that relate network functional efficiency precisely to the density of projections between subsets of network nodes. Our analysis is an attempt to provide a formal framework for existing empirical studies, with the potential to reconcile results that may otherwise seem counterintuitive, even mutually contradictory. For example, some studies

found that a lack of adequate amygdalar projections to prefrontal regions may be responsible for trait anxiety (Kim & Whalen, 2009; Kim et al., 2011), while other studies correlated the same phenomenon with major depression (Dannlowski et al., 2009). A formal model investigating the effects of density on dynamics seemed therefore an appropriate starting point for addressing these ambiguities in a quantifiable framework.

While our dynamic results have been promising and clinically informative, one important step (and the centerstone of this work) is to better understand their source. In our two related papers (one studying the graph properties for fixed edge densities, and the other studying their relationship with network dynamics; Rădulescu & Verduzco-Flores, 2015), we investigate the underpinnings of the observed robustness of coupled dynamics to certain changes in the network architecture and its vulnerability to others, as well as the differences between updating connection strengths versus perturbing connection density or geometry. We aim to clarify that this robustness is not a parameter-dependent property or a numerical artifact but rather an intrinsic feature based on network hardwiring. It is the robustness of certain network architectural features (in this case, the narrow distribution of the adjacency spectrum) that reflects the robustness of the temporal systemic dynamics.

Although there are clearly more complex generative models discussed in the literature on random graphs, random matrices, and complex systems, we found our model interesting to analyze and discuss for a few reasons. While its simple structure may lead to understating its mathematical behavior, we found that understanding spectral properties of configurations with fixed-edge density poses, even in this simple case, an interesting analytical problem and numerical challenges that increase with the size of the network. The question has surfaced before in the literature in a very similar form (Juhasz, 1990) but has been dismissed with partial results that do not completely explain our observations, presented here with systematic simulations and additional analytic justifications. In light of natural systems' predilection for choosing simple schemes to produce complex behavior, we hypothesize that such a setup may contribute, possibly in conjunction with stochastic components, to computational algorithms performed in complex networks such as human brain circuits. Especially in light of our incomplete knowledge of how plastic brain networks perform normalization, it seems as likely to have a global normalization scheme (choosing one network configuration from a distribution of possible options with a certain property) as it is to have local normalization (e.g., assigning a certain probability for each pair of nodes to be joined by an edge).

Even in its simplest form, the model opens questions on the dependence of dynamics on coupling parameters in a network with variable architecture (addressed primarily in Rădulescu & Verduzco-Flores, 2015) and the properties of a bimodular graph with variable edge geometry (e.g, the distribution of adjacency and Laplacian spectra, addressed in this article). In our

work, we continue to study more complex architectures (connected hubs and strong components) and other types of coupled dynamic schemes (e.g., discrete map iterations).

1.3 Network Dynamics from Spectral Measures

1.3.1 The Adjacency Spectrum. A variety of studies have examined random graphs with a general given expected degree distribution and have established bounds or other descriptions of their adjacency spectra. While it is well known that the largest eigenvalue of a graph's adjacency matrix is determined by its maximum degree m together with the weighted average \bar{d} of the squares of the expected degrees (Chung & Lu, 2006), work on random matrices has delivered more accurate estimates. For example, Chung et al. (2003) investigated an ensemble of random uncorrelated, nonoriented networks and found that in the large N limit, the expected largest eigenvalue is determined by the ratio of the second to first moment of the average degree distribution $\langle d^2 \rangle / \langle d \rangle$, together with the expected largest degree d_{\max} . More generally, for directed (oriented) networks without edge degree correlations, a first-order approximation to the leading eigenvalue is given by $\langle d^{\text{in}} d^{\text{out}} \rangle / \langle d \rangle$, where d^{in} and d^{out} are, respectively, the in- and out-degrees of the graph, and $\langle d^{\text{in}} \rangle = \langle d^{\text{out}} \rangle = \langle d \rangle$ (Restrepo, Ott, & Hunt, 2007).

It is therefore clear that the in- and out-degrees, as well as their correlations, have crucial effects on the leading eigenvalue. In general, a graph's defining feature is its distribution of edges. Among other properties, edge density, edge clustering, and presence of hubs have been intensely studied. Detecting and interpreting the modularity of a network (i.e., the presence of community structures within the graph, defined as densely connected groups of nodes, with sparser inter-group connections) has been of particular interest recently (Mucha, Richardson, Macon, Porter, & Onnela, 2010; Chauhan, Girvan, & Ott, 2009; Nadakuditi & Newman, 2012, 2013; Sarkar, Henderson, & Robinson, 2013). Whether the graph represents the architecture of a social (Gilbert, Simonetto, Zaidi, Jourdan, & Bourqui, 2011), climate (Donges, Zou, Marwan, & Kurths, 2009), transportation (Zanin & Lillo, 2013) or disease (Barabási, Gulbahce, & Loscalzo, 2011; Van Mieghem, 2011; Supekar, Menon, Rubin, Musen, & Greicius, 2008) network, modularity reflects into adjacency properties of the network, controlling the structural and functional properties, and implicitly the temporal behavior of the system.

1.3.2 The Graph Laplacian Spectrum. The Laplacian matrix L of a graph is defined as the difference between the node degree matrix and the adjacency matrix. In the case of directed graphs, either the in-degree or out-degree can be used, depending on the application. Laplacian dynamics is perhaps the most studied representation of networked systems and is also known

as the consensus protocol (Olfati-Saber, Fax, & Murray, 2007), in which the network aims to reach agreement on a certain quantity of interest. Although this model has been explored in more elaborate contexts (Olfati-Saber & Murray, 2004; Rahmani, Ji, Mesbahi, & Egerstedt, 2009), in its simplest form, the dynamics of each node is driven by the sum of differences between its own state and its neighbors' states, as defined by the adjacency graph. Then the dynamic evolution of the entire system can be appropriately captured by the linear equation: $\dot{x}(t) = -Lx(t)$.

While the consensus protocol has attracted a lot of attention and effort (Wu, 2013), it is not a complete representation of all the recent work on networked dynamic systems. For example, relative sensing networks are an important class of systems whose control has been described using both their incidence matrix (Smith & Hadaegh, 2007), as well as more completely in terms of spanning trees in the connection topology (Sandhu, Mesbahi, & Tsukamaki, 2005). In fact, the dynamical stability of certain networks seems to remain most successfully defined in terms of quantities derived from the eigenspectrum of the adjacency matrix (Small, Judd, & Stemler, 2012).

In our own work, we considered a bimodal oriented network of coupled nodes, each acting as a Wilson-Cowan-type nonlinear oscillator (Rădulescu & Verduzco-Flores, 2015). Even for such a network, one cannot expect either adjacency or the Laplacian spectrum to be fully predictive of the system's dynamics. Indeed, both cospectral graphs and Laplacian cospectral graphs may produce different phase and parameters-space behavior in the corresponding system (examples of this correspondence are shown in appendix A). A stronger requirement for the graphs to be isomorphic would most likely lead to identical coupled dynamics, but while isomorphic graphs are cospectral and Laplacian cospectral respectively, the converse is not true in either case (Barghi & Ponomarenko, 2009; Zelazo, 2008). These being said, however, both adjacency and graph Laplacian matrices have properties that reflect into the network dynamic behavior.

Throughout this article, we will be working with oriented graphs composed of two interconnected modules X and Y , each composed in turn of N nodes. Within both X and Y , the edge density is fixed to the same fraction γ (out of the possible maximum of N^2). The density of the X -to- Y edges is fixed to a fraction α of the N^2 possible X -to- Y connections and the density of the Y -to- X edges is fixed to β of the N^2 possible Y -to- X connections. The parameters α , β , and γ can take any values of the form $\frac{k}{N^2} \in [0, 1]$, where k is an integer between zero and N^2 (not necessarily requiring that $\gamma > \alpha, \beta$). In this setup, when $\gamma = 0$, the modules are totally disconnected, and when $\gamma = 1$, the modules are fully connected (cliques). Most of this article is dedicated to studying interconnected cliques.

As discussed in section 1.2, this graph structure was used in previous work as a schematic architectural representation of a neural circuit, in which X and Y represent the excitatory (resp. inhibitory) modules of a neural feedback loop, so that X projects to Y through a fraction α of excitatory

connections, and Y in turn modulates X through a fraction β of feedback, inhibitory connections. In such a circuit, the overall connectivity density may remain constant during a cognitive process such as learning, even though the network may exhibit high plasticity and constantly inspect a variety of edge geometry combinations. Throughout the process, the connectivity profile is constantly remodeled, with existing connections being silenced or disappearing, while new connections are being created or activated.

The adjacency matrix of such an oriented graph is a $2N \times 2N$ binary block matrix of the form $\mathbf{T} = \begin{bmatrix} \mathbf{P} & \mathbf{A} \\ \mathbf{B} & \mathbf{Q} \end{bmatrix}$, where the blocks \mathbf{P} and \mathbf{Q} have a fixed fraction γ of 1 versus 0 entries (i.e., edge density), while \mathbf{A} and \mathbf{B} have densities α and β , respectively. Here, we study the sensitivity and robustness properties of the adjacency and Laplacian spectra for our specific class of oriented graphs. We focus in particular on understanding, for increasing size N , how the eigenvalues are perturbed when changing the density profile (α, β, γ) and when changing only the edge distribution while keeping densities fixed. We use a combination of analytical and numerical methods to understand the distribution (mean and standard deviation) of each eigenvalue in the adjacency and Laplacian eigenspectrum. In a separate paper (briefly previewed in section 4.2), we investigate the connections between graph properties and the dynamics of a corresponding system of coupled node oscillators.

1.4 Organization of the Article. Our work is organized as follows. In the following two sections, we study properties of the adjacency and the Laplacian matrix of the graph. In section 2, we focus on the behavior and robustness of the adjacency spectrum when changing the edge density and configuration. (In the text, we restrict ourselves to the case of two interconnected cliques, $\gamma = 1$. However, in appendix B, we relax the full connectedness requirement to $\gamma \leq 1$ and analyze how the properties of the spectrum change with the trimming of intramodular edges.) In section 3, we investigate numerically, by looking at increasing network sizes and variable edge densities, whether the same robustness is characteristic to the spectrum of the graph Laplacian. In section 4, we put our results in the context of the existing work on eigenspectra of random graphs. As a preview to our subsequent work in Rădulescu and Verduzco-Flores (2015), we briefly explore connections with the temporal behavior of a coupled dynamical system and discuss the feasibility of dynamic classification based on classes of adjacency or Laplacian spectra. Finally, we discuss the significance of our results in light of neural connectivity and learning plasticity.

2 Dependence of Adjacency Spectrum of Edge Density and Network Size

When considering the case of fully connected cliques X and Y (see Figure 1), the diagonal blocks of the adjacency matrix \mathbf{T} are $\mathbf{P} = \mathbf{Q} = \mathbf{M}$ (where \mathbf{M} is

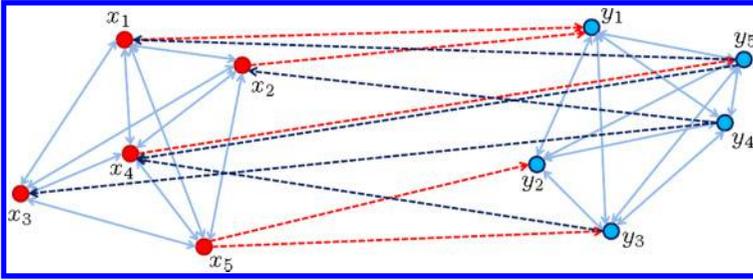


Figure 1: Schematic representation of the network for $N = 5$ nodes per module, as used in our application. Module X is shown on the left; module Y is shown on the right. Both are fully connected, local subgraphs of the full network. The dotted red arrows represent the X -to- Y connections, and the dotted blue arrows represent the Y -to- X connections, generated randomly for low-connectivity densities $\alpha = \beta = 5/25 = 0.2$, to maintain the clarity of the illustration (Rădulescu & Verduzco-Flores, 2015).

the appropriate size matrix with all entries equal to one). Note that this scenario includes self-loops at all nodes; eliminating loops is equivalent to subtracting the identity from the adjacency matrix, with the only effect of shifting all the eigenvalues, and preserving the eigenvectors. The off-diagonal $N \times N$ blocks \mathbf{A} and \mathbf{B} are binary matrices, with fractions α and, respectively, β of ones.

By discussing the effects of edge density we mean analyzing how the spectrum of \mathbf{T} changes when the values of α and β are varied; we will represent these changes in the form of surface plots with respect to pairs $(\alpha, \beta) \in [0, 1]^2$. By discussing the effects of geometry, we mean understanding the effects on the spectrum of the edge configuration, under the constraint of fixed densities α and β . We measure these effects by estimating the mean and standard deviation of the eigenvalues of \mathbf{T} over the edge geometries admissible by any fixed density pair (α, β) . We call λ_j , $j = \overline{1, 2N}$ the eigenvalues of \mathbf{T} , ordered in decreasing order of their magnitudes: $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_{2N}|$.

Let us notice here that λ_1 is guaranteed to be real by an extended version of the Perron-Frobenius theorem, for matrices with nonnegative entries. It is easy to see that the graph with adjacency \mathbf{T} is strongly connected, with every two nodes connected by at least one oriented path (unless $\alpha = 0$ or $\beta = 0$, i.e., if there are no X -to- Y or Y -to- X connections). Hence the matrix \mathbf{T} is irreducible anywhere except on the boundary of the density square (α, β) . Moreover, since the length of the shortest path between two nodes is 1, it follows that the matrix is also aperiodic. An extension of Perron-Frobenius for irreducible matrices guarantees that there is a unique real positive eigenvalue equal to the spectral radius and that the other

eigenvalues are complex. We will be referring to λ_1 and λ_2 as the two leading eigenvalues of \mathbf{T} .

Definition 1. For fixed $0 \leq \alpha, \beta \leq 1$, we call $\mathcal{D}^{\alpha, \beta}$ the distribution of $2N \times 2N$ adjacency matrices \mathbf{T} with off-diagonal blocks \mathbf{A} and \mathbf{B} having densities α and β , respectively. We call $\mathcal{L}_j^{\alpha, \beta}$ the corresponding distribution of each of the eigenvalue real parts $\text{Re}(\lambda_j)$ (with $j = \overline{1, 2N}$).

It is easy to see that the cardinality $|\mathcal{D}^{\alpha, \beta}| = C_{N^2}^{\alpha N^2} C_{N^2}^{\beta N^2}$. While in general the exact eigenvalues of \mathbf{T} depend on the representative $\mathbf{T} \in \mathcal{D}^{\alpha, \beta}$ (i.e., on the actual exact positions of the 1s within the blocks \mathbf{A} and \mathbf{B}), all $\mathcal{L}_j^{\alpha, \beta}$ are trivial on the boundary (i.e., for α or β in $\{0, 1\}$).

Lemma 1. Fixing $\alpha = 1$ fixes the eigenvalues of \mathbf{T} , so that $|\mathcal{L}_j^{1, \beta}| = 1$, for all $j = \overline{1, 2N}$. More precisely, the eigenvalues of any $\mathbf{T} \in \mathcal{D}^{1, \beta}$ are given by (from largest to smallest in absolute value) $\lambda_1 = N + N\sqrt{\beta}$, $\lambda_2 = N - N\sqrt{\beta}$, and $\lambda_3 = \dots = \lambda_{2N} = 0$. Similarly, for $\beta = 1$, the eigenvalues of any $\mathbf{T} \in \mathcal{D}^{\alpha, 1}$ are given by $\lambda_1 = N + N\sqrt{\alpha}$, $\lambda_2 = N - N\sqrt{\alpha}$, and $\lambda_3 = \dots = \lambda_{2N} = 0$.

Proof. We calculate directly, for $\mathbf{T} \in \mathcal{D}^{1, \beta}$, the eigenvalues λ and eigenvectors $\begin{bmatrix} \mathbf{V} \\ \mathbf{W} \end{bmatrix}$ (where $\mathbf{V} = [v_1, \dots, v_N]^t$ and $\mathbf{W} = [w_1, \dots, w_N]^t$):

$$\begin{bmatrix} \mathbf{M} & \mathbf{M} \\ \mathbf{B} & \mathbf{M} \end{bmatrix} \begin{bmatrix} \mathbf{V} \\ \mathbf{W} \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{V} \\ \mathbf{W} \end{bmatrix}.$$

Call $\Sigma_v = \sum_{j=1}^N v_j$ and $\Sigma_w = \sum_{j=1}^N w_j$, and \mathbf{B}_j = the j th row of the block matrix \mathbf{B} , with $\varphi_j(\mathbf{B})$ being the number of 1s in that row. We then have that

$$\Sigma_v + \Sigma_w = \lambda v_j, \text{ for all } j = \overline{1, N},$$

$$\mathbf{B}_j \mathbf{V} + \Sigma_w = \lambda w_j, \text{ for all } j = \overline{1, N}.$$

If $\lambda \neq 0$, then $v_1 = v_2 = \dots = v_N = v$, implying that $\Sigma_w = (\lambda - N)v$. It follows that $\varphi_j(\mathbf{B})v + \Sigma_w = \lambda w_j$ for all $j = \overline{1, N}$. By summing up and using the fact that $\sum_{j=1}^N \varphi_j(\mathbf{B}) = N^2\beta$, we get

$$N^2\beta v + N(\lambda - N)v = \lambda(\lambda - N)v.$$

Clearly $v \neq 0$; otherwise, $w_j = 0$ for all j , and $\begin{bmatrix} \mathbf{V} \\ \mathbf{W} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}$. We then have that $(\lambda - N)^2 = N^2\beta$, hence $\lambda = N \pm N\sqrt{\beta}$.

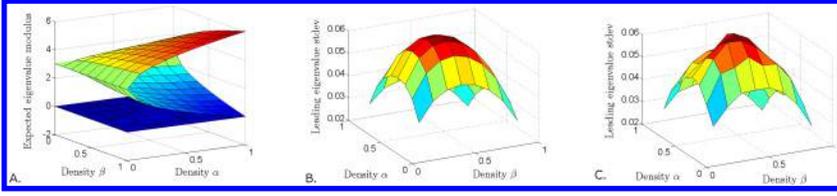


Figure 2: Mean and standard deviation for the leading eigenvalue of \mathbf{T} for $N = 3$, as functions of the densities α and β . (A) For each pair (α, β) , the mean of the leading eigenvalue real part was calculated over all $\mathbf{T} \in \mathcal{D}^{\alpha, \beta}$ (i.e., over all possible combinatorial configurations with the given densities). (B) For each pair (α, β) , the corresponding standard deviation was calculated over all combinatorial configurations in each $\mathcal{D}^{\alpha, \beta}$. (C) For each pair (α, β) , the standard deviation of the leading eigenvalue was also calculated using a sample of the distribution, obtained by choosing randomly 2500 configurations for \mathbf{T} .

In conclusion, any matrix $\mathbf{T} \in \mathcal{D}^{1, \beta}$ has one largest eigenvalue $\lambda_1 = N + N\sqrt{\beta}$, with eigenvector given by $v_j = v = N + N\sqrt{\beta}$, $w_j = \varphi_j(\mathbf{B}) + \sqrt{\beta}$, and a second largest eigenvalue $\lambda_2 = N - N\sqrt{\beta}$, with eigenvector given by $v_j = v = N - N\sqrt{\beta}$, $w_j = \varphi_j(\mathbf{B}) - \sqrt{\beta}$. The rest of $2N - 2$ eigenvalues are zero. Note that in the case of $\beta = 1$, then $\lambda_1 = 2N$ and $\lambda_2 = 0$ as well.

Lemma 2. Fixing $\alpha = 0$ fixes the eigenvalues of \mathbf{T} so that $|\mathcal{L}_j^{0, \beta}| = 1$ for all j . The eigenvalues of any $\mathbf{T} \in \mathcal{D}^{0, \beta}$ are given by $\lambda_1 = N$, $\lambda_2 = \dots = \lambda_{2N} = 0$. Similarly, for $\beta = 0$, the eigenvalues of any $\mathbf{T} \in \mathcal{D}^{\alpha, 0}$ are given by $\lambda_1 = N$, $\lambda_2 = \dots = \lambda_{2N} = 0$.

Proof. The proof is similar to that of lemma 1.

Clearly, the distributions $\mathcal{L}_j^{\alpha, \beta}$ are not trivial in general. If we restricted our interest to finding only the leading eigenvalue of the matrix $\mathbf{T} = \begin{bmatrix} \mathbf{M} & \mathbf{A} \\ \mathbf{B} & \mathbf{M} \end{bmatrix}$, a variety of existing tools can assist us. However, even the computations involved in a task such as expanding the powers \mathbf{T}^k (equivalent to finding all paths of length exactly k in the graph) or in approximating the leading eigenvalue using perturbation theory become very complex quite fast (see section 2.2 and appendix B). It is in this light that we now proceed numerically to support a few conjectures.

Our goal is to obtain descriptions of $\mathcal{L}_j^{\alpha, \beta}$ for all values of $\alpha, \beta \in (0, 1)$; in particular, we want to estimate their means and standard deviations and observe how these depend on the values of α and β and on the size N of the network. For small network sizes ($N \leq 4$), the mean and standard deviation of the entire distribution $\mathcal{L}_j^{\alpha, \beta}$, for each α, β , and j , can be computed directly quite efficiently (see Figures 2A and 2B). However, for larger values of N , the factorial increase in the distribution size makes inspecting all

configurations computationally very expensive (e.g., for $N = 5$ and $\alpha = \beta = 12/25$, we have $|\mathcal{D}^{\alpha,\beta}| = (C_{25}^{12})^2 \sim 10^{13}$ configurations, although some will produce identical spectra). So for larger N s, we estimated the means and standard deviations based on a sample $\mathcal{S} \subset \mathcal{D}^{\alpha,\beta}$ of the distribution. Figures 2B and 2C show a comparison between the whole-distribution and sample-based computations of the standard deviation for $\mathcal{L}_1^{\alpha,\beta}$ for $N = 3$. Even for larger values of N , considering samples of size $|\mathcal{S}| = 500$, or $|\mathcal{S}| = 2500$ produced numerically consistent results (as explained later in this section).

2.1 Numerical Estimates of Eigenvalue Distributions. A few contexts in the literature on eigenspectra of random graphs relate to our problem. The eigenspectrum of the adjacency matrix of a network with communities is known to have leading eigenvalues that are well separated from the rest of the spectrum (Chauhan et al., 2009).

A result more qualitatively related to our question is due to Juhász (1990). Viewed in Juhász's general framework, the adjacency matrix \mathbf{T} is a block matrix with (weighted) density matrix $\mathbf{D} = \begin{bmatrix} 1 & \alpha \\ \beta & 1 \end{bmatrix}$, whose eigenvalues are $\mu_{1,2} = 1 \pm \sqrt{\alpha\beta}$. According to the main theorem in Juhász's, \mathbf{T} has two eigenvalues $\lambda_{1,2}$ that are large (of order N) in magnitude; the other eigenvalues are close to zero. More precisely, $\lambda_{1,2} = N \pm N\sqrt{\alpha\beta} + o(N^{1/2+\varepsilon})$ in probability, while the other eigenvalues are of order $o(N^{1/2+\varepsilon})$ in probability (for any $\varepsilon > 0$).

A first thought is that $N \pm N\sqrt{\alpha\beta}$ may provide in our case the exact formal expressions for the means $E(|\lambda_{1,2}|)$ in terms of the densities α and β . The formulas look particularly promising, since they seem to naturally extend the boundary expressions obtained in the two lemmas (for $\alpha \in \{0, 1\}$ or $\beta \in \{0, 1\}$) and since they match tightly our numerical results (as shown in Figures 2A and 3A). Simple direct computations of the spectra for $N = 3, 4$ immediately reveal, however, that the formulas $N \pm N\sqrt{\alpha\beta}$ do not give the exact means for the leading eigenvalue magnitudes, although this may be the case only for finite sizes N and the estimates may be in fact improving with increasing size and may become exact in the limit $N \rightarrow \infty$. An interesting question to address is that of understanding not only the shape of the leading eigenvalue distributions but also the source of the error terms in their means compared to $N \pm N\sqrt{\alpha\beta}$, and their own behavior with respect to the size N .

For the rest of the section, we gain a numerical insight, for size up to $N = 20$, and provide a few numerically based conjectures on the behavior of the spectrum as the size increases. In section 2.2, we back up analytically some of the conjectures speculated in this section, based on our simulations.

Figure 2 illustrates, for size $N = 3$, the standard deviation of $\mathcal{L}_1^{\alpha,\beta}$ as a function of the densities. For each pair (α, β) , we computed the standard

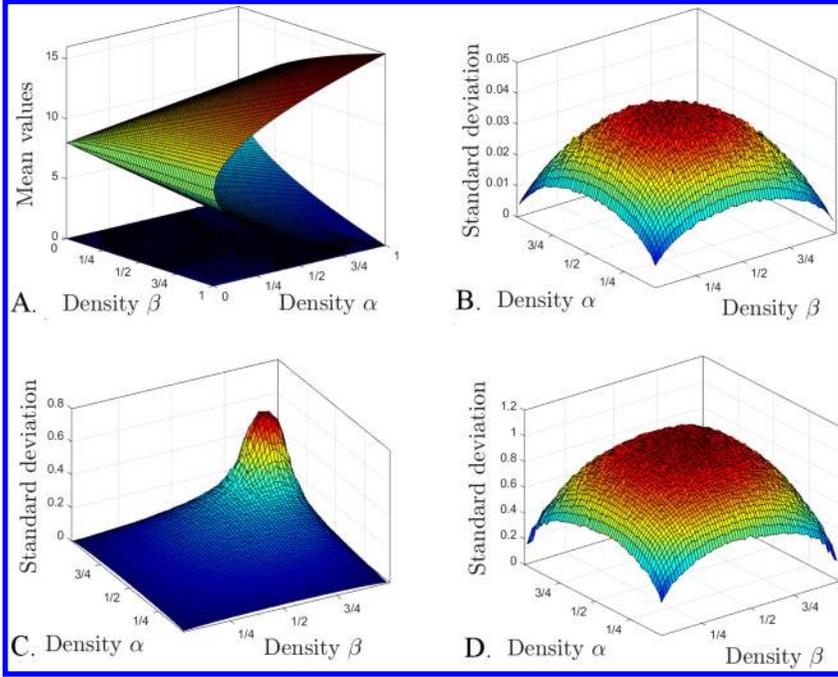


Figure 3: Mean and standard deviation of eigenvalue magnitudes for $N = 8$, estimated numerically for each pair of densities (α, β) by considering a random sample of 2500 matrices \mathbf{T} . (A) The mean eigenvalue real parts are represented as surfaces with respect to (α, β) . The two top surfaces fit very closely the expressions $N \pm N\sqrt{\alpha\beta}$; the other surfaces are all close to zero. For each of the first (panel B), second (panel C), and third (panel D) leading eigenvalues (in magnitude), we represent the corresponding standard deviation as a surface with respect to (α, β) .

deviation of $\mathcal{L}_1^{\alpha,\beta}$ over all configurations in $\mathcal{D}^{\alpha,\beta}$ (see Figure 2B), as well as over a random sample of 2500 representatives for \mathbf{T} . Figures 3B to 3D show similar results for $N = 8$; for each pair (α, β) , we used 2500 samples for \mathbf{T} to estimate numerically the standard deviations of $\mathcal{L}_1^{\alpha,\beta}$, $\mathcal{L}_2^{\alpha,\beta}$ and $\mathcal{L}_3^{\alpha,\beta}$. In all cases, the surfaces decrease toward the edges, illustrating the narrowing of the corresponding distributions when (α, β) gets closer to the boundary of the unit square. We point out the possible confound that the numerical scheme may be introducing by considering the same cardinality (2500) for sampling the larger distributions in the center, as well as the slimmer distributions near the boundary (i.e., the underestimation due to sampling may be more pronounced around the center of the surface than towards the boundary).

For a fixed N , the distribution $\mathcal{L}_j^{\alpha,\beta}$ for each eigenvalue λ_j is clearly largest at intermediate values of α and β . Following the same logic (“higher cardinality likely produces higher variance”), one would expect standard deviations to increase when the size N is increased (recall that $|\mathcal{D}^{\alpha,\beta}| = C_{N^2}^{\alpha N^2} C_{N^2}^{\beta N^2}$, which increases factorially with N). Juhász’s estimate goes along the same lines, claiming an almost everywhere correction term of magnitude $o(N^{1/2+\varepsilon})$, which increases with N . This means that there are almost no outliers out of the Juhász range, even though the spread of each $\mathcal{L}_j^{\alpha,\beta}$ remains quite large, of order $o(N)$, as discussed in section 2.2.

In Figure 4, we illustrate specifically the outcome of our numerical simulations of how the standard deviations behave with increasing N (with approximation algorithms based on sample distributions). In Figures 4A to 4C, we show, for $2 \leq N \leq 20$, the standard deviations for the three leading eigenvalues, each represented as a surface with respect to density pairs (α, β) . Figure 4D tracks the behavior of the maximum of the surface corresponding to each of the first four eigenvalues over the unit (α, β) square. Our estimates suggest that for $j = 3, 4$, the standard deviations of $\mathcal{L}_j^{\alpha,\beta}$ increase as a power function of N (with the power $\sim 1/2$). This is not surprising in light of the existing results already described. However, interestingly, the simulations suggest a decreasing power rule $\sim N^{-1/2}$ for the standard deviation of $\mathcal{L}_1^{\alpha,\beta}$ and a logarithmic increase for the standard deviation of $\mathcal{L}_2^{\alpha,\beta}$, implying that for the two large eigenvalues, Juhász’s result can be greatly refined in terms of standard deviations. This is a useful fact to investigate, since narrowness of the distributions $\mathcal{L}_{1,2}^{\alpha,\beta}$ with N ensures better separation between the leading eigenvalues and the rest of the spectrum, and subsequently more “recognizable” modularity properties (as discussed in section 4.1). This feature can become quite important when the graph operates as a functional network (e.g., as a brain feedback circuit).

We summarize our initial theoretical and numerical observations in the case of two connected cliques in the form of a conjecture, which remains open to a more rigorous investigation:

Conjecture 1. In the case of fully connected modules $\gamma = 1$ (i.e., $\mathbf{S} = \mathbf{R} = \mathbf{M}$), the spectrum of the matrix \mathbf{T} varies with respect to the intermodular densities α and β of the blocks \mathbf{A} and \mathbf{B} as follows:

1. For $(\alpha, \beta) \neq (1, 1)$, the spectrum has two eigenvalues λ_1 and λ_2 whose mean magnitudes are large, while the other $2N - 2$ have small mean magnitudes (close to zero). As $(\alpha, \beta) \rightarrow (1, 1)$, the second largest eigenvalue $\lambda_2 \rightarrow 0$ as well.
2. For each size N and each density pair (α, β) , the mean real parts of the two leading eigenvalues (over all adjacency configurations corresponding to (α, β)), are given approximately by $N \pm N\sqrt{\alpha\beta}$, with error terms approaching zero as $N \rightarrow \infty$.

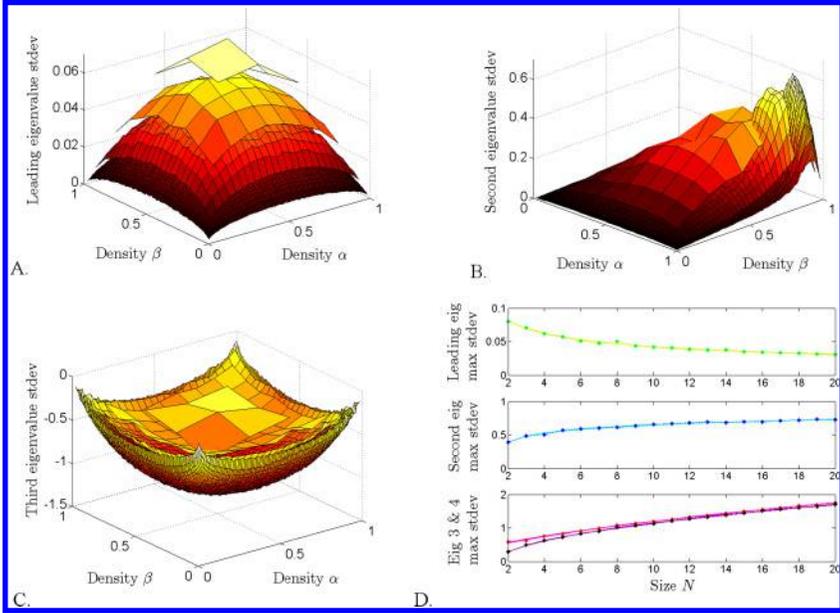


Figure 4: Illustration of the evolution of the standard deviation of $\mathcal{L}_j^{\alpha,\beta}$ when increasing the network size N . (A) Each surface represents the standard deviation of $\mathcal{L}_1^{\alpha,\beta}$ with respect to (α, β) , for a different size N ; from top to bottom: $N = 2$, $N = 3$, $N = 5$, and $N = 10$. (B) Each surface represents the standard deviation of $\mathcal{L}_2^{\alpha,\beta}$ for a different size N ; from lowest to highest: $N = 4$, $N = 6$, and $N = 10$. (C) Each surface represents the standard deviation of $\mathcal{L}_3^{\alpha,\beta}$, for a different size N ; from top to bottom: $N = 2$, $N = 3$, $N = 5$, and $N = 10$. In order to make all surfaces visible, the figure is vertically flipped (we show minus the standard deviation). (D) The dotted plots show how the global maximum value of each surface evolves when increasing the size up to $N = 20$. For each curve, we used a Levenberg-Marquardt algorithm to determine the best functional fit, shown as a solid line (in some cases, the solid line is hard to see because of its almost perfect overlap with the simulation data). Top: The maximum of $\mathcal{L}_1^{\alpha,\beta}$ decreases with N (dotted green curve), as $\sim N^{-0.47}$ (yellow solid curve), with residuals norm $\varepsilon = 0.0048$. Middle: The maximum of $\mathcal{L}_2^{\alpha,\beta}$ increases with N (dotted blue curve), as $\sim \log(N)$ (cyan solid curve), with residuals norm $\varepsilon = 0.0013$. Bottom: The maxima of $\mathcal{L}_3^{\alpha,\beta}$ (dotted red curve) and $\mathcal{L}_4^{\alpha,\beta}$ (dotted black curve) increase as $\sim N^{0.59}$ (solid pink) and $\sim N^{0.51}$ (solid purple), with residual norms $\varepsilon = 0.0034$ and 0.0014 , respectively. The estimates for panels A, B, and C are based on samples of size 2500. The estimates for panel D are based on samples of size 400.

3. For any size N , the standard deviation of each eigenvalue's real part is a "unimodal" surface, with a point of maximum in the open square $(0, 1)^2$, and which is zero when $\alpha \in \{0, 1\}$ or $\beta \in \{0, 1\}$.

4. For the leading eigenvalue λ_1 , the standard deviations for all $(\alpha, \beta) \in [0, 1]$ are very small. Moreover, the standard deviation of $\mathcal{L}_1^{\alpha, \beta}$ decreases monotonically with N for each fixed pair (α, β) . The maximum attainable standard deviation of $\mathcal{L}_1^{\alpha, \beta}$ over $(\alpha, \beta) \in [0, 1]$ decreases approximately as $N^{-1/2}$. (This transcends qualitatively the corresponding Juhász estimate.)
5. For the second eigenvalue λ_2 , the maximum attainable standard deviation of $\mathcal{L}_2^{\alpha, \beta}$ over $(\alpha, \beta) \in [0, 1]$ increases logarithmically with N . (This transcends quantitatively the corresponding Juhász estimate.)
6. For the rest of the eigenvalues $\lambda_j, j \geq 3$, the maximum attainable standard deviation of $\mathcal{L}_j^{\alpha, \beta}$ over $(\alpha, \beta) \in [0, 1]$ increases approximately as $N^{1/2}$. (This is the same as the rate of the almost everywhere error term previously obtained by Juhász.)

Remark 1. We are in particular interested in understanding the robustness of the leading eigenvalues to changes in configuration once the densities have been fixed. First, one might suspect that this robustness is due to a large extent to the existence of the two fully connected cliques in our graph. In appendix B, we investigate how results change when we relax the fully connectedness condition. Second, recall that we are ultimately interested in whether robust features in the adjacency spectrum translate into robustness in dynamics (if we consider the corresponding network of coupled oscillators). In our follow-up paper (briefly previewed in section 4.2 and in appendix A), we discuss this aspect further, and the potential connections between adjacency and dynamics classes.

2.2 Estimating the Maximal Eigenvalue. One of our goals is to understand the source of the standard deviation of the maximal eigenvalue when considered over the matrix distribution $\mathcal{D}^{\alpha, \beta}$. There are a few ways in which one typically proceeds to approximate the maximal eigenvalue of a matrix. In this technical section, we present what two such methods entail in our case and justify our preference for a numerical approach to computing the mean and standard deviation. First, we use Taylor series to expand the eigenvalue and corresponding eigenvector around their values for a matrix with simple known spectrum. Second, we use the power method, initiated at the vector $\mathbf{v}_0 = \mathbf{1}$, to obtain the leading eigenvector and then compute the corresponding eigenvalue.

Throughout this section, \mathbf{M} denotes the $N \times N$ matrix with all entries equal to 1, $\mathbf{1}$ denotes the $N \times 1$ column vector with all entries 1, and φ denotes the function that computes the sum of all entries, for any arbitrary size matrix.

2.2.1 *Taylor Expansion Approach.* The adjacency matrix \mathbf{T} for our graph is of the form $\left[\begin{array}{c|c} \mathbf{M} & \mathbf{A} \\ \mathbf{B} & \mathbf{M} \end{array} \right]$, where $\varphi(\mathbf{A}) = \alpha N^2$ and $\varphi(\mathbf{B}) = \beta N^2$. At the start of section 2, we found the spectrum of \mathbf{T} when (α, β) is on the boundary of the unit square. The spectrum is also easy to find for the matrix $\mathbf{C} = \left[\begin{array}{c|c} \mathbf{M} & \alpha \mathbf{M} \\ \beta \mathbf{M} & \mathbf{M} \end{array} \right]$, a nonbinary matrix that averages out all configurations \mathbf{T} for a fixed pair (α, β) .

Lemma 3. *The matrix $\mathbf{C} = \left[\begin{array}{c|c} \mathbf{M} & \alpha \mathbf{M} \\ \beta \mathbf{M} & \mathbf{M} \end{array} \right]$ has eigenvalues:*

- $\lambda_1 = N + N\sqrt{\alpha\beta}$, with corresponding eigenvector $\mathbf{u}_1 = \begin{bmatrix} \sqrt{a}\mathbf{1} \\ \sqrt{b}\mathbf{1} \end{bmatrix}$
- $\lambda_2 = N - N\sqrt{\alpha\beta}$, with corresponding eigenvector $\mathbf{u}_2 = \begin{bmatrix} \sqrt{a}\mathbf{1} \\ -\sqrt{b}\mathbf{1} \end{bmatrix}$
- $\lambda_3 = \dots = \lambda_{2N} = 0$, with corresponding eigenspace spanned by the vectors $\mathbf{u}_k = \begin{bmatrix} \mathbf{t}_k \\ \mathbf{s}_k \end{bmatrix}$, where \mathbf{t}_k and \mathbf{s}_k are $N \times 1$ column vectors with $\varphi(\mathbf{t}_k) = \varphi(\mathbf{s}_k) = 0$, for $k \geq 3$.

Proof. The proof is direct, and will be omitted.

We are interested in the spectrum of the matrix \mathbf{T} , which we write in the form $\mathbf{T} = \mathbf{C} + \mathbf{Z}$, where the difference $\mathbf{Z} = \mathbf{T} - \mathbf{C} = \left[\begin{array}{c|c} \mathbf{0} & \mathbf{U} \\ \mathbf{V} & \mathbf{0} \end{array} \right]$ has $\varphi(\mathbf{U}) = \varphi(\mathbf{V}) = 0$. More generally, we consider the matrix family $\mathbf{T}_\varepsilon = \mathbf{C} + \varepsilon \mathbf{Z}$ (so that $\mathbf{T} = \mathbf{T}_\varepsilon$, for $\varepsilon = 1$). The leading eigenvalue $\lambda_1(\varepsilon)$ (and its corresponding eigenvector) of \mathbf{T}_ε can then be expanded as a Taylor series around the original leading eigenvalue $\lambda_1 = \lambda_1(0)$ of \mathbf{C} (with corresponding original eigenvector \mathbf{u}_1):

$$\mathbf{T}_\varepsilon(\mathbf{u}_1 + \varepsilon \mathbf{x} + \varepsilon^2 \mathbf{y} + \dots) = (\lambda_1 + \varepsilon \mu + \varepsilon^2 \nu + \dots)(\mathbf{u}_1 + \varepsilon \mathbf{x} + \varepsilon^2 \mathbf{y} + \dots) \quad (2.1)$$

(we can assume, without loss of generality, that \mathbf{x} is perpendicular to \mathbf{u}_1 , \mathbf{y} is perpendicular to both \mathbf{x} and \mathbf{u}_1 , and so on). One can easily compute the first term (of order ε) in the expansion of $\lambda_1(\varepsilon)$ by considering the terms $O(\varepsilon)$:

$$\mathbf{T}_\varepsilon(\mathbf{u}_1 + \varepsilon \mathbf{x}) = (\lambda_1 + \varepsilon \mu)(\mathbf{u}_1 + \varepsilon \mathbf{x}). \quad (2.2)$$

Expanding with respect to ε and identifying the coefficients of ε , we get

$$\mathbf{Z}\mathbf{u}_1 + \mathbf{C}\mathbf{x} = \lambda_1 \mathbf{x} + \mu \mathbf{u}_1.$$

In the basis $(\mathbf{u}_k)_{k=1,2N}$ of eigenvectors of \mathbf{C} , one can write $\mathbf{x} = \sum x_k \mathbf{u}_k$ and the matrix \mathbf{Z} as $(z_{jk})_{i,j=1,2N}$, so that $\mathbf{Z}\mathbf{u}_j = \sum z_{jk} \mathbf{u}_k$. Then our equation becomes

$$\sum x_k \lambda_k \mathbf{u}_k + \sum z_{1k} \mathbf{u}_k = \lambda_1 \sum x_k \mathbf{u}_k + \mu \mathbf{u}_1.$$

Solving in components, this gives us:

- $\mu = z_{11}$, for $k = 1$
- $x_k = \frac{z_{1k}}{\lambda_1 - \lambda_k}$, for $k \geq 2$

Given the form of the eigenvectors in the basis $(\mathbf{u}_k)_{k=1,2N}$, we easily can calculate some of the z_{jk} s that are most useful to continue our computation. For example, on one hand,

$$\mathbf{Z}\mathbf{u}_1 = \begin{bmatrix} \mathbf{0} & \mathbf{U} \\ \mathbf{V} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \sqrt{a}\mathbf{1} \\ \sqrt{b}\mathbf{1} \end{bmatrix} = \begin{bmatrix} \sqrt{b}\mathbf{U}\mathbf{1} \\ \sqrt{a}\mathbf{V}\mathbf{1} \end{bmatrix},$$

and on the other hand, in components,

$$\mathbf{Z}\mathbf{u}_1 = z_{11} \begin{bmatrix} \sqrt{a}\mathbf{1} \\ \sqrt{b}\mathbf{1} \end{bmatrix} + z_{12} \begin{bmatrix} \sqrt{a}\mathbf{1} \\ -\sqrt{b}\mathbf{1} \end{bmatrix} + \sum_{k \geq 3} z_{1k} \begin{bmatrix} \mathbf{t}_k \\ \mathbf{s}_k \end{bmatrix}.$$

Recall that $\varphi(\mathbf{t}_k) = \varphi(\mathbf{s}_k) = 0$, for all $k \geq 3$. Applying the operator φ separately over the first the top and bottom N entries, we get, respectively,

$$N\sqrt{a}(z_{11} + z_{12}) = \sum_{k \geq 3} z_{1k} \varphi(\mathbf{t}_k) = 0,$$

$$N\sqrt{b}(z_{11} - z_{12}) = \sum_{k \geq 3} z_{1k} \varphi(\mathbf{s}_k) = 0.$$

This implies that $z_{11} = z_{12} = 0$, and subsequently $\mu = 0$. Hence the order one Taylor polynomial in ε of the leading eigenvalue of \mathbf{T}_ε is $P_1(\varepsilon) = N + N\sqrt{\alpha\beta}$ (the same as the leading eigenvalue of \mathbf{C}). One can continue computing higher-order terms. For example, we can include terms of order ε^2 in the expansions of both eigenvalue and eigenvector from equation 2.1, and aim to calculate v :

$$\mathbf{T}_\varepsilon(\mathbf{u}_1 + \varepsilon\mathbf{x} + \varepsilon^2\mathbf{y}) = (\lambda_1 + \varepsilon\mu + \varepsilon^2\nu)(\mathbf{u}_1 + \varepsilon\mathbf{x} + \varepsilon^2\mathbf{y}). \quad (2.3)$$

Identifying the coefficients of ε^3 and ε^4 and using the fact that $\mu = 0$, we get two more equations, which can be used to completely determine ν and \mathbf{y} :

$$\mathbf{C}\mathbf{y} + \mathbf{Z}\mathbf{x} = \lambda_1\mathbf{y} + \nu\mathbf{u}_1, \quad (2.4)$$

$$\mathbf{Z}\mathbf{y} = \nu\mathbf{x}.$$

Writing equation 2.4 in components, we have

$$\sum_k \lambda_k \mathbf{y}_k \mathbf{u}_k + \sum_{k,l} \mathbf{x}_k z_{kl} \mathbf{u}_l = \lambda_1 \sum_k \mathbf{y}_k \mathbf{u}_k + \nu \mathbf{u}_1.$$

Projecting in the direction of \mathbf{u}_1 , replacing $x_k = \frac{z_{1k}}{\lambda_1 - \lambda_k}$, for all $k \geq 2$, and also recalling that $z_{11} = 0$, we have that

$$\nu = \sum_k \mathbf{x}_k z_{k1} = \sum_{k \geq 2} \frac{z_{1k} z_{k1}}{\lambda_1 - \lambda_k}. \quad (2.5)$$

We additionally know that $z_{12} = 0$ and that $\lambda_k = 0$, for $k \geq 3$. Hence:

$$\nu = \sum_{k \geq 3} \frac{z_{1k} z_{k1}}{\lambda_1} = \frac{1}{N + N\sqrt{\alpha\beta}} \sum_{k \geq 3} z_{1k} z_{k1}. \quad (2.6)$$

But $\sum_{k \geq 3} z_{1k} z_{k1} = \sum_{k \geq 1} z_{1k} z_{k1}$ is in fact nothing but the first component of the matrix \mathbf{Z}^2 , written in the basis $(\mathbf{u}_k)_{k=1,2,\overline{N}}$. In other words, if we write in components $\mathbf{Z}^2 \mathbf{u}_1 = A_1 \mathbf{u}_1 + A_2 \mathbf{u}_2 + \sum_{k \geq 3} A_k \mathbf{u}_k$, then $\sum_{k \geq 1} z_{1k} z_{k1} = A_1$. To calculate A_1 , we can use the fact that $\langle \mathbf{u}_1, \mathbf{u}_k \rangle = \langle \mathbf{u}_2, \mathbf{u}_k \rangle = 0$, for all $k \geq 3$ and calculate:

$$\langle \mathbf{u}_1, \mathbf{Z}^2 \mathbf{u}_1 \rangle = A_1 \|\mathbf{u}_1\| + A_2 \langle \mathbf{u}_1, \mathbf{u}_2 \rangle = N(a+b)A_1 + N(a-b)A_2,$$

$$\langle \mathbf{u}_1, \mathbf{Z}^2 \mathbf{u}_1 \rangle = A_1 \langle \mathbf{u}_1, \mathbf{u}_2 \rangle + A_2 \|\mathbf{u}_2\| = N(a-b)A_1 + N(a+b)A_2.$$

On the other hand, $\mathbf{Z}^2 = \left[\begin{array}{c|c} \mathbf{UV} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{VU} \end{array} \right]$, so that

$$\langle \mathbf{u}_1, \mathbf{Z}^2 \mathbf{u}_1 \rangle = \alpha \mathbf{1}^T \mathbf{UV} \mathbf{1} + \beta \mathbf{1}^T \mathbf{VU} \mathbf{1},$$

$$\langle \mathbf{u}_2, \mathbf{Z}^2 \mathbf{u}_1 \rangle = \alpha \mathbf{1}^T \mathbf{UV} \mathbf{1} - \beta \mathbf{1}^T \mathbf{VU} \mathbf{1}.$$

Combining the two, we get

$$N(A_1 + A_2) = \mathbf{1}^T \mathbf{UV} \mathbf{1},$$

$$N(A_1 - A_2) = \mathbf{1}^T \mathbf{VU} \mathbf{1},$$

hence $A_1 = \frac{1}{2N} (\mathbf{1}^T \mathbf{UV} \mathbf{1} + \mathbf{1}^T \mathbf{VU} \mathbf{1})$.

In conclusion, we obtained the quadratic Taylor polynomial in ε for the eigenvalue $\lambda_1(\varepsilon)$:

$$P_2(\varepsilon) = N + N\sqrt{\alpha\beta} + \varepsilon^2 \frac{1}{2N} \frac{1}{N + N\sqrt{\alpha\beta}} (\mathbf{1}^T \mathbf{U} \mathbf{V} \mathbf{1} + \mathbf{1}^T \mathbf{V} \mathbf{U} \mathbf{1}). \quad (2.7)$$

Keep in mind, however, that the difference between $P_2(\varepsilon)$ and $\lambda_1(1)$ (the exact leading eigenvalue of \mathbf{T}) is the second-order Taylor remainder evaluated at $\varepsilon = 1$, which we have not yet reason to believe is small (that, assuming the Taylor series converges at $\varepsilon = 1$). While obtaining additional higher-order terms could provide increasingly confident estimates $P_n(\varepsilon)$, the mean and standard deviation of $P_n(\varepsilon)$ become harder to calculate (involving more complex, mixed products of the matrices \mathbf{U} and \mathbf{V}). This would make an exact analytic expression of the mean and standard deviation intractable using this particular method. To illustrate, we study the statistics of $P_2(\varepsilon)$ over the distribution $\mathcal{D}^{\alpha,\beta}$ and estimate how much of the mean and standard deviation of $\lambda_1(\varepsilon)$ are captured by its quadratic ‘‘approximation’’ $P_2(\varepsilon)$:

Proposition 1. *The mean of $\varphi(\mathbf{U}\mathbf{V} + \mathbf{V}\mathbf{U}) = \mathbf{1}^T \mathbf{U} \mathbf{V} \mathbf{1} + \mathbf{1}^T \mathbf{V} \mathbf{U} \mathbf{1}$ over all matrix configurations in $\mathcal{D}^{\alpha,\beta}$ is zero.*

Proof. Consider two $N \times N$ binary matrices \mathbf{A} and \mathbf{B} with densities α and, respectively, β . Then $E[\varphi(\mathbf{A}\mathbf{B})] = N^3\alpha\beta$, where E represents the mean over all configurations in $\mathcal{D}^{\alpha,\beta}$. Indeed, we have $E(\mathbf{A}^i) = \alpha N$ and $E(\mathbf{B}_i) = \beta N$, for all $1 \leq i \leq N$. Since the matrices \mathbf{A} and \mathbf{B} are independent, we can easily compute $E[\varphi(\mathbf{A}\mathbf{B})] = N \cdot \alpha N \cdot \beta N = \alpha\beta N^3$.

Since the second correction term in the eigenvalue expansion has zero mean, $P_2(\varepsilon)$ does not account for any of the difference between the mean of $\lambda_1(1)$ and the value $N + N\sqrt{\alpha\beta}$, which we observed numerically and have been investigating. This difference must be therefore entirely contained within the higher-order terms in the expansion. This is not true for the standard deviation $\sigma(P_2(\varepsilon))$, which in fact seems to explain much of the standard deviation of $\lambda_1(1)$. Indeed:

$$\sigma(P_2(\varepsilon)) = \frac{\varepsilon^2 \sigma(\mathbf{1}^T \mathbf{U} \mathbf{V} \mathbf{1} + \mathbf{1}^T \mathbf{V} \mathbf{U} \mathbf{1})}{2N(N + N\sqrt{\alpha\beta})}.$$

We can express $\sigma^2(\mathbf{1}^T \mathbf{U} \mathbf{V} \mathbf{1}) = E[\varphi(\mathbf{U}\mathbf{V}\mathbf{U}\mathbf{V})] + E[\varphi(\mathbf{U}\mathbf{V}\mathbf{V}\mathbf{U})] + E[\varphi(\mathbf{V}\mathbf{U}\mathbf{V}\mathbf{U})]$. The mean for a product of the type $\mathbf{U}\mathbf{V}\mathbf{U}\mathbf{V}$ is nontrivial to express analytically (unlike $E[\varphi(\mathbf{U}\mathbf{V})]$ and $E[\varphi(\mathbf{V}\mathbf{U})]$ and $E[\varphi(\mathbf{U}\mathbf{V}\mathbf{V}\mathbf{U})]$, which are easy to determine). We compute the standard deviation of $P_2(1)$ for $N = 3$ and all pairs (α, β) . In Figure 5, we compare it with the standard deviation of the leading eigenvalue $\lambda_1(1)$ for all values of (α, β) .

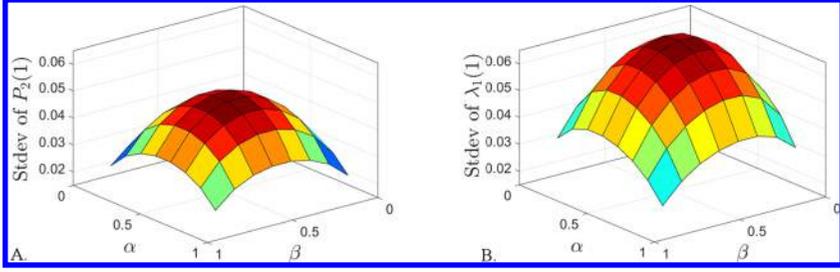


Figure 5: Comparison between the standard deviations of the leading eigenvalue of \mathbf{T} (left) and its quadratic Taylor polynomial (right). The computations were done for $N = 3$, for all pairs $(\alpha, \beta) \in (0, 1)^2$.

While the computation of the standard deviation can get complicated fast, one can more easily calculate bounds for the leading eigenvalue around its mean, by noticing that:

Proposition 2. $-2N^3\alpha\beta \leq \varphi(\mathbf{UV}) \leq 2N^3\sqrt{\alpha\beta}(1 - \sqrt{\alpha\beta})$.

Proof. Consider two $N \times N$ binary matrices \mathbf{A} and \mathbf{B} with densities of ones α and, respectively, β , that is, $\varphi(\mathbf{A}) = \alpha N^2$ and $\varphi(\mathbf{B}) = \beta N^2$. We have that

$$\varphi(\mathbf{AB}) = [\mathbf{A}^1 \dots \mathbf{A}^N] \cdot \begin{bmatrix} \mathbf{B}_1 \\ \vdots \\ \mathbf{B}_N \end{bmatrix},$$

where \mathbf{A}^i is the sum of the elements in the i th column of \mathbf{A} and \mathbf{B}_i is the sum of the elements in the i th row of \mathbf{B} , hence $\sum \mathbf{A}^i = \alpha N^2$ and $\sum \mathbf{B}_i = \beta N^2$. Using the Cauchy-Schwartz inequality, we can see that

$$\varphi(\mathbf{AB}) = \sum \mathbf{A}^i \mathbf{B}_i \leq \sqrt{\sum (\mathbf{A}^i)^2 \cdot \sum (\mathbf{B}_i)^2}.$$

Furthermore, each $(\mathbf{A}^i)^2 = (\sum a_{i1})^2 \leq \sum a_{i1}^2 \cdot \sum 1 = N \sum a_{i1}$. Similarly, each $\mathbf{B}_i \leq N \sum b_{1i}$, hence $\sqrt{\sum (\mathbf{A}^i)^2 \cdot \sum (\mathbf{B}_i)^2} \leq \sqrt{N\varphi(\mathbf{A}) \cdot N\varphi(\mathbf{B})} = N^3\sqrt{\alpha\beta}$.

In conclusion, $\varphi(\mathbf{AB}) \leq N^3\sqrt{\alpha\beta}$. We use finally use the fact that $\mathbf{U} = \mathbf{A} - \alpha\mathbf{M}$ and $\mathbf{V} = \mathbf{B} - \beta\mathbf{M}$ to compute $\varphi(\mathbf{UV}) = \varphi(\mathbf{AB} - \alpha\mathbf{BM} - \beta\mathbf{AM} + \alpha\beta\mathbf{M}^2) = \varphi(\mathbf{AB}) - \alpha\beta N^3 - \alpha\beta N^3 + \alpha\beta N^3 = \varphi(\mathbf{AB}) - \alpha\beta N^3$.

The inequality in proposition 2 is not helpful, however, since it estimates a maximal variability $\frac{1}{2N} \frac{1}{N+N\sqrt{\alpha\beta}} \cdot N^3$ for the eigenvalue $\lambda_1(1)$ around $N + N\sqrt{\alpha\beta}$. While we desire to narrow the estimates as N increases, this bound increases like $O(N)$ with size (a similar problem as with Juhasz's

estimate) and fails to explain our numerical observations. However, this bound cannot be improved to a lower order of N since, for all N , one can always find outliers in the distribution $\mathcal{L}_1^{\alpha,\beta}$ at a distance $\sim N$ from $N + N\sqrt{\alpha\beta}$. An explanation that reconciles both observations, as well as Juhász's almost everywhere bounds, is that these outliers are less representative as N increases, causing the distribution to remain narrow, with a small standard deviation that decreases with N .

2.2.2 Power Iteration Approach. It can be shown that the recurrent sequence of vectors obtained from the power algorithm $\mathbf{v}_k = \mathbf{T}^k \mathbf{v}_0$, initiated at $\mathbf{v}_0 = \mathbf{1}$, converges in our case to the leading eigenvector \mathbf{v} of \mathbf{T} . Then the leading eigenvalue can be computed as the Rayleigh quotient $\frac{\mathbf{v}^t \mathbf{T} \mathbf{v}}{\mathbf{v}^t \mathbf{v}}$. To estimate a term like $\mathbf{v}^t \mathbf{v}$, we calculate powers of the matrix $\mathbf{T} = \begin{bmatrix} \mathbf{M} & \mathbf{A} \\ \mathbf{B} & \mathbf{M} \end{bmatrix}$ and then compute the limit of $\mathbf{v}_0 (\mathbf{T}^k)^t \mathbf{T}^k \mathbf{v}_0 = \varphi[(\mathbf{T}^k)^t \mathbf{T}^k]$. As with the Taylor method, this computation quickly becomes intractable when facing the combinatorial problem of evaluating the function φ for mixed products of the blocks \mathbf{A} , \mathbf{B} and their transposes.

2.2.3 Numerical Approach. Due to these apparent difficulties to support the conjectured statements with analytic results for the standard deviation, we resorted to numerical illustrations. These present different but surmountable difficulties, such as the increased potential for inaccuracy in computing the standard deviation based on fixed-size sample distributions. Indeed, recall that the size of $\mathcal{L}_1^{\alpha,\beta}$ increases factorially with N , making it unrealistic to explore all configurations in this distribution. Hence any computationally tractable approach based on sample distributions can only increase the sample sizes with N at a much slower rate than the rate at which the actual size of $\mathcal{L}_1^{\alpha,\beta}$ increases, making these samples potentially less and less reliable with larger sizes. For a brief illustration of the appropriateness of our sample-based computations, we compare in Figure 6 the histogram of $\mathcal{L}_1^{\alpha,\beta}$ (containing, for $N = 4$ and $\alpha = \beta = 8$, a total of 12,870² configurations) with that produced by a sample of 100² configurations.

3 Dependence of Laplacian Spectrum of Network Size and Edge Densities

For our oriented graph with adjacency matrix $\mathbf{T} = \begin{bmatrix} \mathbf{M} & \mathbf{A} \\ \mathbf{B} & \mathbf{M} \end{bmatrix}$, we consider the in-node degree diagonal matrix $\mathbf{\Delta}$, with

$$\Delta_{jj} = \varphi_j(\mathbf{T}) \text{ for all } j = \overline{1, 2N}$$

so that the corresponding Laplacian matrix is given by $\mathbf{L} = \mathbf{\Delta} - \mathbf{T}$.

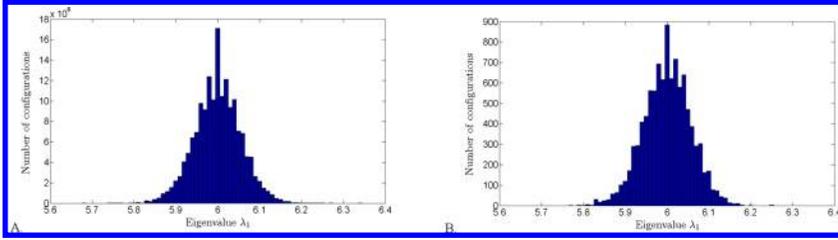


Figure 6: Comparison between the distribution $\mathcal{L}_1^{\alpha, \beta}$ and a sample-based distribution. (A) Histogram of the distribution $\mathcal{L}_1^{\alpha, \beta}$, for $N = 4$, $\alpha = 8$, $\beta = 8$. (B) Histogram for a random subset of values in $\mathcal{L}_1^{\alpha, \beta}$, for $N = 4$, $\alpha = 8$, $\beta = 8$, computed based on a sample of size 10^4 configurations out of the total of $12,870^2$ configurations.

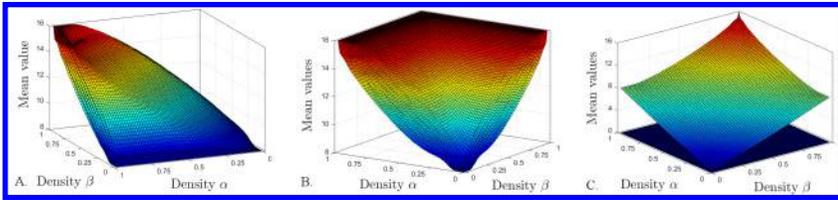


Figure 7: Mean of Laplacian eigenvalue real parts for $N = 8$, estimated numerically for each pair of densities (α, β) by considering the same random sample of 2500 adjacency configurations as in Figure 3A. Ordered by their magnitude: (A) The leading $N - 1$ eigenvalues (1–6), (B) The following $N - 1$ eigenvalues (7–14). (C) The two smallest eigenvalue (15 and 16). The smallest eigenvalue is zero (the Laplacian matrix is always rank degenerate).

The Laplacian eigenvalue spectrum has been used as a measure of system dynamics. For example, the algebraic connectivity, defined as the second smallest eigenvalue μ_{N-1} of the discrete Laplacian matrix, is known to play an important role on synchronization dynamics and network robustness. In an effort to study the effect of interdependent topologies on the mutual synchronization of networks, Martin-Hernandez, Wang, Van Mieghem, and D’Agostino (2013) focused on computing and approximating the algebraic connectivity of two interdependent networks and on showing that it experiences a phase transition with the addition of a sufficient number of links among two interdependent networks. Here, we study the dependence of the Laplacian eigenvalues on the densities (α, β) .

Following the same numerical scheme as in section 2, we computed the Laplacian eigenvalues for a sample of configurations, chosen randomly from the large distribution of all configurations corresponding to any fixed density pair (α, β) . Based on this sample, we estimated for each (α, β) the

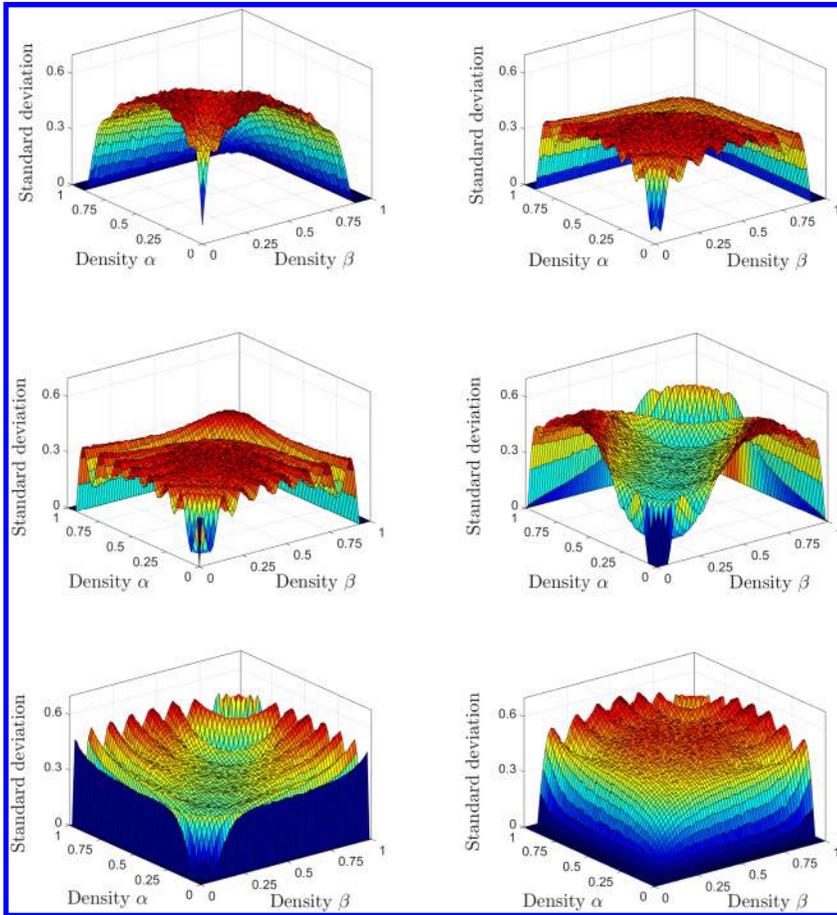


Figure 8: Standard deviations of Laplacian eigenvalue real parts for $N = 8$. The panels represent, from top to bottom and left to right, the standard deviations for the eigenvalues 1, 2, 3, 7, 10, and 14 (ordered by magnitudes).

mean and standard deviation of the real part of the spectrum, as illustrated in Figures 7 and 8 for $N = 8$.

The behavior of the standard deviations for the real parts of the Laplacian eigenvalues with respect to the density pair (α, β) is very different from that of the standard deviations for the adjacency spectrum. While the adjacency standard deviation surfaces were unimodal on the domain $[0, 1]^2$, decreasing from a central peak toward the boundary, in the case of the Laplacian, the surfaces are rippled (Figure 8), with the amplitude and distribution of the ripples depending on a variety of factors (as illustrated in Figure 10 and discussed below).

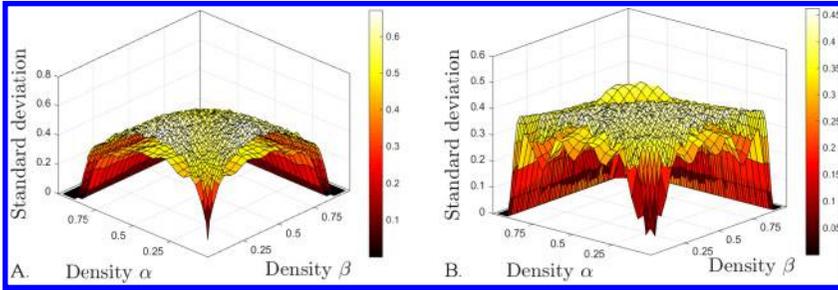


Figure 9: Illustration of the evolution of the standard deviation of the Laplacian eigenvalue real parts, when increasing network size N . (A) Surfaces for the first eigenvalue, computed for $N = 5$ (lower surface) and $N = 8$ (higher surface). (B) Surfaces for the second eigenvalue, for $N = 5$ (lower surface) and $N = 8$ (higher surface).

Such variability in the standard deviation values makes it easier for the system to switch from robust regimes (with a narrow distribution of eigenvalues) to more scattered regimes (with a wider distribution of potential eigenvalues) by introducing a small change in the density (α, β) . Scattered regimes are more sensitive to configuration, since wide changes in the Laplacian spectrum (and implicitly in Laplacian-driven dynamics) are accessible even under the same density pair by slightly altering the configuration. This could be in principle viewed as an adaptability feature that makes Laplacian driven a desirable type of dynamics.

However, the emergent robustness observed in the case of the adjacency leading eigenvalue (standard deviation of the real part decreasing with the size N) does not hold in the case of the leading Laplacian eigenvalue. In fact, the maximum standard deviations over the (α, β) domain seem to increase as powers of N for all the eigenvalues in the Laplacian spectrum after an initial transient phase for very small N (see Figures 9, 10, and 11). As N increases, the central regions of the surface, which raise with N , smoothen out and in the process push the ripples toward the borders.

If comparing the behavior of the two (adjacency and Laplacian) spectra when changing (α, β) and increasing N , one could say that the desirable feature of the adjacency model is robustness of the leading eigenvalue, which increases with size, while the feature of the Laplacian model is swiftness between robust and loose regimes, which degrades with increasing size.

4 Discussion

4.1 Comparison with Random Graphs Approaches to Modularity. Nadakuditi and Newman (2012) considered a stochastic, nonoriented

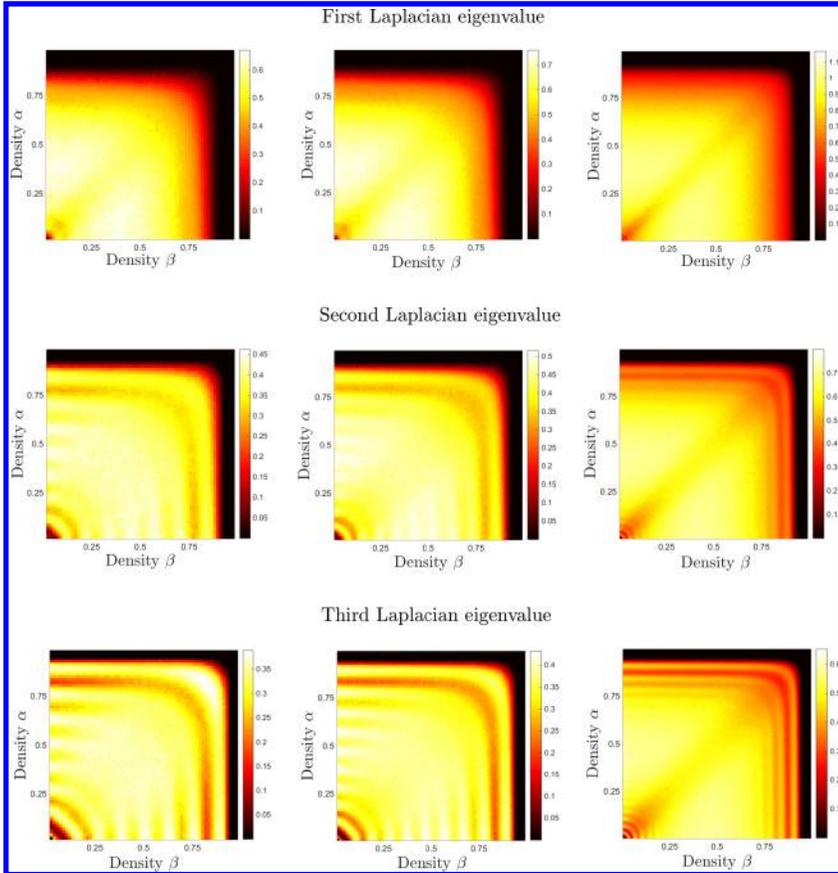


Figure 10: Illustration of the evolution of the standard deviation of the Laplacian eigenvalue real parts when increasing the network size N . The surfaces for the first second and third eigenvalues are shown top to bottom as pcolor plots for $N = 5$ (left), $N = 10$ (center), and $N = 20$ (right).

network with two communities and computed the ensemble means for the two large eigenvalues of its symmetric adjacency matrix in the large N limit. The method involved first finding the eigenvalues of the modularity matrix, then showing that these are identical in the large N limit to the eigenvalues of the adjacency matrix. Their asymptotic expressions z_1 and z_2 were computed in terms of $c_{\text{in}} = np_{\text{in}}$ and $c_{\text{out}} = np_{\text{out}}$ (where the notations in the original text are n for the matrix size, p_{in} for the probability of two nodes within a module to be directly connected, and p_{out} for the probability of two nodes that are not in the same module to be directly connected). More

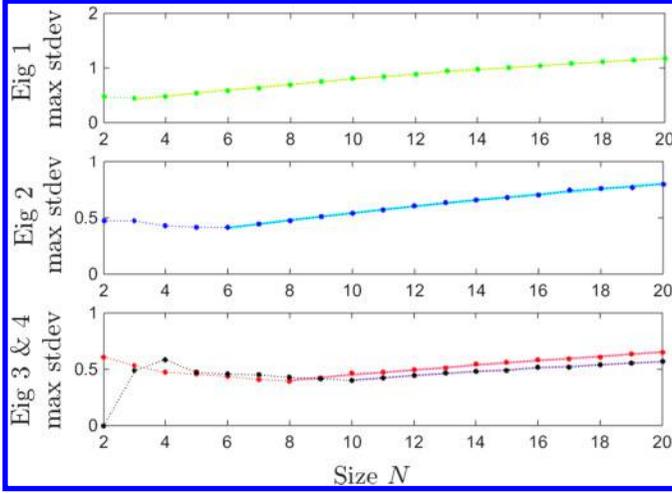


Figure 11: Curve fitting that illustrates the evolution of the Laplacian eigenvalue standard deviation as the size increases. The dotted plots show how the global maximum value of each surface evolves when increasing the size up to $N = 20$. For each curve, we used a Levenberg-Marquardt algorithm to determine the best functional fit, shown as a solid line. Top: The maximum mean real part for the first eigenvalue increases with $N \geq 3$ (dotted green curve), as $\sim N^{0.42}$ (yellow solid curve), with residuals norm $\varepsilon \sim 10^{-3}$. Middle: The maximum for the second eigenvalue increases with $N \geq 6$ (dotted blue curve), as $\sim N^{0.52}$ (cyan solid curve), with residuals norm $\varepsilon \sim 10^{-4}$. Bottom: The maxima for the third (dotted red curve) and fourth eigenvalues (dotted black curve) increase as $\sim N^{0.69}$ (solid pink) and $\sim N^{0.33}$ (solid purple) for $N \geq 8$ and $N \geq 10$ respectively, with residual norms $\varepsilon \sim 10^{-4}$. The estimates are based on samples of size 500.

precisely:

$$z_1 = \frac{1}{2}(c_{\text{in}} + c_{\text{out}}) + 1,$$

$$z_2 = \frac{1}{2}(c_{\text{in}} - c_{\text{out}}) + \frac{c_{\text{in}} + c_{\text{out}}}{c_{\text{in}} - c_{\text{out}}}.$$

With our notation, $c_{\text{in}} = 2N$, $c_{\text{out}} = 2\alpha N$ and the adjacency matrix is symmetric ($\beta = \alpha$). Accounting for the presence of loops for all the nodes in our network (which were excluded in the Nadakuditi-Newman model), we get:

$$z_1 = N + \alpha N,$$

$$z_2 = N(1 - \alpha) + \frac{1 + \alpha}{1 - \alpha} - 1,$$

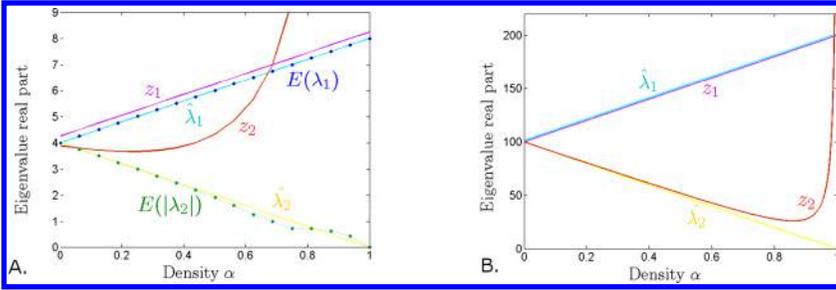


Figure 12: Comparison between our results and those of Nadakuditi and Newman (2012) in the case of a bimodular, nonoriented graph. We compare the values of $z_{1,2}$ computed by Nadakuditi and Newman (solid curves in purple and brown, respectively) with the formal means $E(\lambda_{1,2})$ (dotted curves in blue and green) and their close approximations $N \pm \alpha N$ (solid curves in yellow and cyan). (A) Comparison for $N = 4$. (B) Comparison for $N = 100$. Here, we used only the approximations $\hat{\lambda}_{1,2} = N \pm \alpha N$, since the formal means are computationally too expensive.

so that $z_1 > z_2$ if $\alpha < 1 - \frac{1}{N}$. In Figure 12, we show a comparison between our results and those of Nadakuditi and Newman (2012) when applied to a non-oriented graph with two fully connected communities by illustrating on the same axes as $z_{1,2}$ the formal means $E(\lambda_{1,2})$, and their close approximations obtained earlier as $\hat{\lambda}_{1,2} = N \pm \alpha N$. The approximations approach exactness in the large N limit, at least for values of $\alpha < 1 - \sqrt{\frac{2}{N}}$ (this is the density where z_1 has its global minimum, after which it shoots up, detaching from the graph of $E(|\lambda_2|)$).

Finally, the reference (Nadakuditi & Newman, 2012) investigated the spectral distribution of the modularity matrix (i.e., whose spectral radius is, in the large N limit, also the spectral radius for the adjacency matrix of the original nonoriented graph). The spectrum consisted of a continuous semicircular band of eigenvalues and an additional, unique leading eigenvalue. As long as the leading eigenvalue is well separated from the semicircular band, there is evidence of community structure in the network. When the leading eigenvalue passes the edge of the band ($z = \sqrt{c_{\text{in}} + c_{\text{out}}}$), the community structure is no longer detectable. As we will suggest in section 4.3, the property appears to extend to the case of the oriented random graph that constitutes our study case. In Figure 13 we show, for $N = 4$, a sample (100 configurations) of the spectrum, observing the separation between eigenvalues, as the modularity of the network changes. Since the adjacency matrix is no longer symmetric, the eigenvalues are plotted in the complex plane. All eigenvalues are distributed within the unit disc except the first two largest in absolute value, which for $\gamma = 1$ are real and

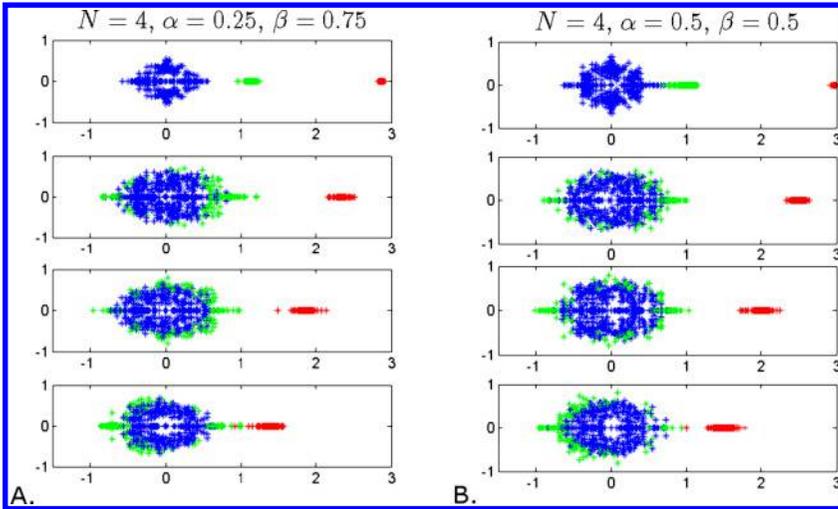


Figure 13: Illustration of separation of eigenvalues, when changing the community structure. The eigenvalues are plotted in the complex plane. The leading eigenvalue in absolute value is shown in red, the second largest in green, the rest in blue. All plots are for $N = 4$ and are based on samples of 100 matrix configurations under the following restrictions: (A) $\alpha = 1/4, \beta = 3/4, \gamma = 1, 3/4, 1/2, 1/4$ (from top to bottom). (B) $\alpha = 1/2, \beta = 1/2, \gamma = 1, 3/4, 1/2, 1/4$ (from top to bottom).

significantly larger than 1. When beginning to decrease the “community structure,” (i.e., γ decreases), the second leading eigenvalue collides into the unit disc and starts diffusing around its boundary. If we continue decreasing γ , the first leading eigenvalue will also become indistinguishable from the pool distribution.

4.2 Adjacency to Dynamics: Strengthening versus Restructuring. The main interest of our current work lies in studying the consequences of spectral robustness under fixed-edge densities on the network’s coupled dynamics. We focus primarily on finding the measures of architecture and dynamics that are optimal for quantifying their relationship. For our node-wise dynamics, we have been using both discrete iterated maps and continuous time nonlinear oscillators. For our first coupled system (Rădulescu & Verduzco-Flores, 2015), we chose one of the most studied historical models in theoretical neuroscience, which has inspired many other analytical and modeling efforts (Av-Ron, Parnas, & Segel, 1993; Borisyuk, Borisyuk, Khibnik, & Roose, 1995): the Wilson-Cowan model (Wilson & Cowan, 1972). This is a variation of the two-dimensional Fitzhugh-Nagumo system in which the coupled variables represent the fraction of neurons active at any specific time in a pair of interacting excitatory and inhibitory neural populations.

The model was shown to exhibit hysteresis and Hopf bifurcations with bistability windows (in which the system has both an attracting equilibrium and an attracting limit cycle, separated by an unstable cycle). It was later shown (Borisyyuk et al., 1995) that by varying the strength of the symmetric weak coupling between two Wilson-Cowan excitatory and inhibitory units, one can produce very rich four-dimensional phase-space transitions (bifurcations between symmetric, antisymmetric, and nonsymmetric attractors like equilibria, cycles, and invariant tori).

One interesting direction we investigate is to compare how dynamic behavior depends on architecture (viewed as a system parameter) versus how it depends on other parameters (see Rădulescu & Verduzco-Flores, 2015, as well as appendix A). For example, consider two alternative ways to increase information diffusion between the two modules of our case study network: one by increasing the intermodular edge weights and the other by increasing their density. Both actions lead to “increasing connectivity” between X and Y and to similar effects on the spectrum of the connectivity matrix, so one may suspect that they also lead to similar changes in the temporal behavior of the corresponding dynamic network. However, our work suggests that this is not the case and that the effects obtained when perturbing these two different aspects of the network connectivity can be very similar in some instances but qualitatively different in others.

Choosing the appropriate interplay between perturbing the configuration of the network and changing the coupling strengths seems to be an important part of the continuous choices a complex system like the brain needs to make to maintain optimal function. Under some circumstances, local configuration perturbations to the network may have more substantial dynamic effects than those obtained by a global change in the system’s weights. In the context of optimal dynamics in a functional network, this may be seen as a vulnerability (simple addition of a few edges may drastically affect the function) but also as an adaptability feature (the system can more easily obtain the optimal flexibility that triggers efficient responses to the outside world).

4.3 Edge Updating and Learning Algorithms. Many different models describe, qualitatively or quantitatively, the synaptic adjustments that may take place in a network of neurons during learning. Since the oriented graph in this article may be viewed as a model network of coupled neurons (with each edge representing a weighted synapse), we discuss briefly how the results reflect on our knowledge of synaptic update mechanisms and their consequence to learning.

In general, the process of synaptic updating is assumed to involve not only weight changes of existing synapses but also activation of “silent” sites (thus creating new connections), and silencing, or pruning of active sites (thus deleting existing synapses). In terms of our model, this means that not only the edge weights but also the edge distribution is likely to change during learning. A clear biological restriction on synaptic updating has to

be that the connections are somehow prevented from increasing without bound, which is why most models incorporate a normalization scheme. Although it is important to understand the different consequences of using different normalization mechanisms when modeling synaptic updating and rewiring, the manner in which a normalization step may actually be implemented by the brain is not at all clear and has been the subject of scientific controversy. Some rules assume the process to be local—subtractive normalization rules (Oja, 1982; Goodhill, 1993; Willshaw & von der Malsburg, 1976; Miller & MacKay, 1994), weight-dependent rules (Elliott & Shadbolt, 2002), or BCM rules (Cooper, 2004)), but one can imagine various other ways of ensuring stability, possibly involving “homeostasis” or “synaptic scaling” (Turrigiano, Leslie, Desai, Rutherford, & Nelson, 1998; Turrigiano & Nelson, 2004). Many models support a global normalization, for which the state of the whole network is assessed at each updating step and a specific norm is imposed at each weight update.

Hence, while most models of learning introduce the updates into the weights themselves, another (additional) theoretically acceptable scheme could be to “normalize” (at least in the short term) by simply maintaining the overall number of active network connections approximately constant, so that in the updating process, in the long-term average, one synapse will turn off whenever a new site is activated. One would then want to understand how these architectural dynamics may promote/influence learning, and how the effects of geometry updating complement or compare with the effects of weight updating.

We finally note that the local mechanism of adding or deleting edges based on a probabilistic process (as described in section 4.1), even though equivalent to our model in the large N limit, produces substantially different spectra from our alternative normalization scheme for finite N (see Figure 12B for $N = 100$). Since many brain networks appear to operate with hundreds of nodes, it is important to understand the apparent distinctions between the two models for relatively large but finite values of N .

Knowledge of the geometry of the network is very important when determining which connectivity schemes are plausible to use for models of learning. The choices currently used in modeling range from considering fully connected to fully disconnected interacting modules or layers (O’Reilly & Frank, 2006). Our results suggest that convergence (learning) is not a priori prevented in either case. In developing future iterations of this model, it will also be important to explore how the learning process itself shapes the connectivity scheme. Siri et al. (2007) suggest that the structure emerging during learning breaks down into different numbers of hublike subnetworks, which is very likely to affect the spectral robustness demonstrated in our modular network. Understanding the source and limits of this robustness is an instrument that could be used to investigate which architectures favor convergence under particular learning algorithms, and which do not.

Appendix A: Correspondence between Adjacency Spectra

A.1 Laplacian Spectra and Network Dynamics. In Rădulescu and Verduzco-Flores (2015), we considered the following $2N$ -dimensional system of coupled nonlinear oscillators:

$$\begin{aligned}\dot{x}_k &= -x_k + (1 - x_k) \cdot \mathcal{S}_{b_{x_x}, \theta_x} \left(- \sum_{p=1}^N g_{yx} a_{kp} y_p + \sum_{p=1}^N g_{xx} x_p + P \right), \\ \dot{y}_k &= -y_k + (1 - y_k) \cdot \mathcal{S}_{b_{y_y}, \theta_y} \left(\sum_{p=1}^N g_{xy} b_{kp} x_p + \sum_{p=1}^N g_{yy} y_p + Q \right)\end{aligned}\quad (\text{A.1})$$

with $1 \leq k \leq N$. Each node is driven by external sources (P for the nodes x_k in the module X and Q for the nodes y_k in the module Y). In addition, each node receives input from all other nodes that are connected to it through incoming edges, with weights g (with the corresponding subscripts). The coefficients $a_{kp}, b_{kp} \in \{0, 1\}$ are the binary entries of the adjacency blocks A and B . The effective input to each node is the sum of all such external and internal sources, modulated by the sigmoidal:

$$\mathcal{S}_{b, \theta}[Z] = \frac{1}{1 + \exp(-b[Z - \theta])} - \frac{1}{1 + \exp(b\theta)}, \quad (\text{A.2})$$

with parameters in the range used in the original Wilson-Cowan model (Wilson & Cowan, 1972), as well as in subsequent papers (Borisyyuk et al., 1995).

In Rădulescu and Verduzco-Flores (2015), we considered as an application networks of size 4 (i.e., $N = 2$) and inspected the dynamic behavior of the system for every possible theoretical configuration of the adjacency matrix corresponding to a fixed pair of edge densities (α, β) . To quantify the changes in dynamics produced by varying system parameters (such as the intermodular connectivity weights g_{xy} and g_{yx}), we used bifurcation diagrams in the (g_{xy}, g_{yx}) parameter plane. Then we observed how these diagrams changed when perturbing the underlying adjacency graph. We constructed all possible (g_{xy}, g_{yx}) parameter planes that can be obtained for $N = 2$ for each of two density pairs: $(\alpha, \beta) = (3/4, 3/4)$ and $(\alpha, \beta) = (1/2, 3/4)$, respectively. All 16 combinatorial configurations in $\mathcal{D}^{3/4, 3/4}$ produced only four distinct dynamic parameter planes (which we labeled i through iv). Similarly, all 24 combinatorial configurations in $\mathcal{D}^{1/2, 3/4}$ produced only six dynamic classes (which we labeled i through vi).

In Tables 1 and 2 we illustrate for these two examples to what extent cospectral and Laplacian cospectral graphs lead to the same dynamics. In the case of $(\alpha, \beta) = (3/4, 3/4)$, there are three classes of adjacency

Table 1: Classes of Adjacency and Laplacian Spectra in Correspondence with Dynamic Classes, for $N = 2$, Density Type $(\alpha, \beta) = (3/4, 3/4)$.

$\left[\begin{array}{c cc} & 1 & 1 \\ \hline 1 & 1 & \\ 1 & 0 & \end{array} \right] (\mathcal{A}, \text{I})_{iii}$	$\left[\begin{array}{c cc} & 1 & 1 \\ \hline 1 & 1 & 0 \\ 1 & 0 & \end{array} \right] (\mathcal{B}, \text{II})_{iv}$	$\left[\begin{array}{c cc} & 1 & 0 \\ \hline 1 & 1 & 1 \\ 1 & 0 & \end{array} \right] (\mathcal{B}, \text{II})_{ii}$	$\left[\begin{array}{c cc} & 0 & 1 \\ \hline 1 & 1 & 1 \\ 1 & 0 & \end{array} \right] (\mathcal{C}, \text{III})_i$
$\left[\begin{array}{c cc} & 1 & 1 \\ \hline 1 & 1 & \\ 0 & 1 & \end{array} \right] (\mathcal{B}, \text{II})_{ii}$	$\left[\begin{array}{c cc} & 1 & 1 \\ \hline 1 & 1 & 0 \\ 0 & 1 & \end{array} \right] (\mathcal{C}, \text{III})_i$	$\left[\begin{array}{c cc} & 1 & 0 \\ \hline 1 & 1 & 1 \\ 0 & 1 & \end{array} \right] (\mathcal{A}, \text{I})_{iii}$	$\left[\begin{array}{c cc} & 0 & 1 \\ \hline 1 & 1 & 1 \\ 0 & 1 & \end{array} \right] (\mathcal{B}, \text{II})_{iv}$
$\left[\begin{array}{c cc} & 1 & 1 \\ \hline 1 & 0 & \\ 1 & 1 & \end{array} \right] (\mathcal{B}, \text{II})_{iv}$	$\left[\begin{array}{c cc} & 1 & 1 \\ \hline 1 & 0 & 0 \\ 1 & 1 & \end{array} \right] (\mathcal{A}, \text{I})_{iii}$	$\left[\begin{array}{c cc} & 1 & 0 \\ \hline 1 & 0 & 1 \\ 1 & 1 & \end{array} \right] (\mathcal{C}, \text{III})_i$	$\left[\begin{array}{c cc} & 0 & 1 \\ \hline 1 & 0 & 1 \\ 1 & 1 & \end{array} \right] (\mathcal{B}, \text{II})_{ii}$
$\left[\begin{array}{c cc} & 1 & 1 \\ \hline 0 & 1 & \\ 1 & 1 & \end{array} \right] (\mathcal{C}, \text{III})_i$	$\left[\begin{array}{c cc} & 1 & 1 \\ \hline 0 & 1 & 0 \\ 1 & 1 & \end{array} \right] (\mathcal{B}, \text{II})_{ii}$	$\left[\begin{array}{c cc} & 1 & 0 \\ \hline 0 & 1 & 1 \\ 1 & 1 & \end{array} \right] (\mathcal{B}, \text{II})_{iv}$	$\left[\begin{array}{c cc} & 0 & 1 \\ \hline 0 & 1 & 1 \\ 1 & 1 & \end{array} \right] (\mathcal{A}, \text{I})_{iii}$

Note: Adjacency classes are designated as \mathcal{A} – \mathcal{C} , Laplacian classes as I–III, and dynamics classes by subscripts i – iv .

eigenspectra (designated by the letters \mathcal{A} through \mathcal{C}), which in this case are also the three classes for Laplacian eigenspectra (designated I to III). The four distinct dynamics classes (designated by indices i through iv) are mapped to the spectral classes in a well-defined but not surjective way: that is, no dynamics can be obtained from multiple adjacency classes but some adjacency classes can lead to multiple dynamics.

Similarly, Table 2 shows how the six dynamic classes accessible to $\mathcal{D}^{1/2, 3/4}$ are mapped to the adjacency and Laplacian spectral classes. In this case, the adjacency spectral classes (\mathcal{A} through \mathcal{D}) do not coincide with the Laplacian classes (I–III). Dynamics is once again well mapped to both adjacency and Laplacian spectral classes, although not surjectively (in fact, the many-to-one convergence is higher for Laplacian classes).

This suggests that while the adjacency and Laplacian spectra, together with the density type, clearly have a contribution to dynamics, neither cannot be directly used to predict these dynamics. In fact, in this case, it is likely that the spectrum of the Laplacian gives less information on the dynamics than the spectrum of the adjacency matrix.

Table 2: Spectral and Dynamics Classes for $N = 2$, Density Type $(\alpha, \beta) = (1/2, 3/4)$.

$\left[\begin{array}{c cc} & 1 & 1 \\ \hline 1 & 1 & 0 \\ 1 & 0 & \end{array} \right]$	$(\mathcal{A}, \text{I})_v$	$\left[\begin{array}{c cc} & 1 & 1 \\ \hline 1 & 1 & 0 \\ 0 & 1 & \end{array} \right]$	$(\mathcal{B}, \text{III})_{vi}$	$\left[\begin{array}{c cc} & 1 & 1 \\ \hline 1 & 0 & 0 \\ 1 & 1 & \end{array} \right]$	$(\mathcal{A}, \text{I})_v$	$\left[\begin{array}{c cc} & 1 & 1 \\ \hline 0 & 1 & 0 \\ 1 & 1 & \end{array} \right]$	$(\mathcal{B}, \text{III})_{vi}$
$\left[\begin{array}{c cc} & 1 & 0 \\ \hline 1 & 1 & 0 \\ 1 & 0 & \end{array} \right]$	$(\mathcal{A}, \text{II})_i$	$\left[\begin{array}{c cc} & 1 & 0 \\ \hline 1 & 1 & 0 \\ 0 & 1 & \end{array} \right]$	$(\mathcal{A}, \text{II})_i$	$\left[\begin{array}{c cc} & 1 & 0 \\ \hline 1 & 0 & 0 \\ 1 & 1 & \end{array} \right]$	$(\mathcal{B}, \text{I})_{ii}$	$\left[\begin{array}{c cc} & 1 & 0 \\ \hline 0 & 1 & 0 \\ 1 & 1 & \end{array} \right]$	$(\mathcal{B}, \text{I})_{ii}$
$\left[\begin{array}{c cc} & 1 & 0 \\ \hline 1 & 1 & 0 \\ 1 & 0 & \end{array} \right]$	$(\mathcal{C}, \text{III})_{iv}$	$\left[\begin{array}{c cc} & 1 & 0 \\ \hline 1 & 1 & 0 \\ 0 & 1 & \end{array} \right]$	$(\mathcal{D}, \text{I})_{iii}$	$\left[\begin{array}{c cc} & 1 & 0 \\ \hline 1 & 0 & 0 \\ 1 & 1 & \end{array} \right]$	$(\mathcal{D}, \text{I})_{iii}$	$\left[\begin{array}{c cc} & 1 & 0 \\ \hline 0 & 1 & 0 \\ 1 & 1 & \end{array} \right]$	$(\mathcal{C}, \text{III})_{iv}$
$\left[\begin{array}{c cc} & 0 & 1 \\ \hline 1 & 1 & 0 \\ 1 & 0 & \end{array} \right]$	$(\mathcal{D}, \text{I})_{iii}$	$\left[\begin{array}{c cc} & 0 & 1 \\ \hline 1 & 1 & 0 \\ 0 & 1 & \end{array} \right]$	$(\mathcal{C}, \text{III})_{iv}$	$\left[\begin{array}{c cc} & 0 & 1 \\ \hline 1 & 0 & 0 \\ 1 & 1 & \end{array} \right]$	$(\mathcal{C}, \text{III})_{iv}$	$\left[\begin{array}{c cc} & 0 & 1 \\ \hline 0 & 1 & 0 \\ 1 & 1 & \end{array} \right]$	$(\mathcal{D}, \text{I})_{iii}$
$\left[\begin{array}{c cc} & 0 & 1 \\ \hline 1 & 1 & 0 \\ 1 & 0 & \end{array} \right]$	$(\mathcal{B}, \text{I})_{ii}$	$\left[\begin{array}{c cc} & 0 & 1 \\ \hline 1 & 1 & 0 \\ 0 & 1 & \end{array} \right]$	$(\mathcal{B}, \text{I})_{ii}$	$\left[\begin{array}{c cc} & 0 & 1 \\ \hline 1 & 0 & 0 \\ 1 & 1 & \end{array} \right]$	$(\mathcal{A}, \text{II})_i$	$\left[\begin{array}{c cc} & 0 & 1 \\ \hline 0 & 1 & 0 \\ 1 & 1 & \end{array} \right]$	$(\mathcal{A}, \text{II})_i$
$\left[\begin{array}{c cc} & 0 & 0 \\ \hline 1 & 1 & 1 \\ 1 & 0 & \end{array} \right]$	$(\mathcal{B}, \text{III})_{vi}$	$\left[\begin{array}{c cc} & 0 & 0 \\ \hline 1 & 1 & 1 \\ 0 & 1 & \end{array} \right]$	$(\mathcal{A}, \text{I})_v$	$\left[\begin{array}{c cc} & 0 & 0 \\ \hline 1 & 0 & 1 \\ 1 & 1 & \end{array} \right]$	$(\mathcal{B}, \text{III})_{vi}$	$\left[\begin{array}{c cc} & 0 & 0 \\ \hline 0 & 1 & 1 \\ 1 & 1 & \end{array} \right]$	$(\mathcal{A}, \text{I})_v$

Note: Adjacency classes are denoted \mathcal{A} - \mathcal{D} , Laplacian classes are denoted I-III, and dynamics classes are denoted as indices as i - vi .

Appendix B: Connecting Sparser Modules

To investigate more general networks, we want to relax the full-connectedness condition of the two modules and explore other intramodular edge configurations, more realistic in the context of brain connectivity. It is well known that the eigenspectrum of the adjacency matrix of a network organized in communities has leading eigenvalues that are well separated from the rest of the eigenvalues (Chauhan et al., 2009). If our bimodular

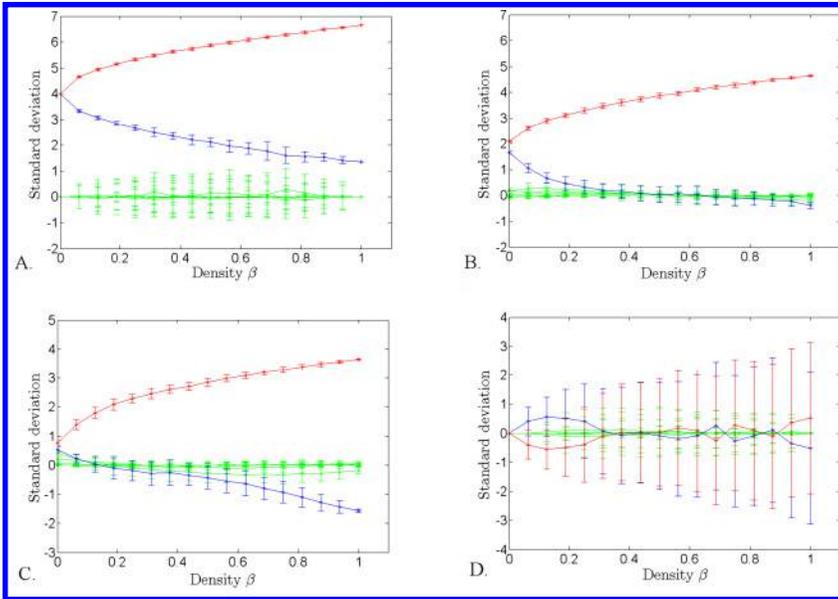


Figure 14: Behavior of eigenvalues of T as γ decreases from 1 to 0. Here, $N = 4$ and $\alpha = 1/2$. In each panel: $\gamma = 1$ (panel A), $\gamma = 1/2$, (panel B), $\gamma = 1/4$ (panel C), and $\gamma = 0$ (panel D). The mean values of each eigenvalue magnitude are represented along each curve together with the corresponding standard deviation (as error bars): the largest eigenvalue in red, the second largest in blue, and the remaining (footnotesize) eigenvalues in green.

graph is thought of as describing the underlying coupling scheme for a dynamical system, the position and overlap of the distributions $\mathcal{L}_j^{\alpha,\beta}$ will automatically reflect in the spectral properties of the network connectivity matrix (see section 4.2) and implicitly in the system's Jacobian matrix, thus affecting local dynamics around its equilibria.

In this section, we illustrate in our specific case how the eigenvalue distributions and the distance between them evolve as the modularity structure is gradually lost (how the leading eigenvalues approach the distribution of the remaining eigenvalues, as γ decreases).

Figure 14 shows the means and standard deviations of $\mathcal{L}_j^{\alpha,\beta}$ for four levels of intramodular connectivity (each panel corresponds to a different value of γ , with α fixed and β varied along the x -axis). When $\gamma = 1$ (see Figure 14A), we recover the fully connected modules studied in section 2: the values of the standard deviations are small, and the first and second eigenvalues (whose means are well approximated in magnitude by $N \pm N\sqrt{\alpha\beta}$) remain to a large extent separated from the other small eigenvalues. As γ

decreases from 1, this situation gradually changes, and the large expected eigenvalues decay in mean approximately as $N\gamma \pm N\sqrt{\alpha\beta}$ respectively (see Figures 14B and 14C), to eventually completely collapse only when $\gamma = 0$ (see Figure 14D).

To understand why the leading eigenvalues are close to $N\gamma \pm N\sqrt{\alpha\beta}$ in mean, one can use a similar Taylor expansion computation to the one carried out in equation 2.7, as follows:

Lemma 4. *The matrix $\mathbf{C} = \begin{bmatrix} \gamma\mathbf{M} & \alpha\mathbf{M} \\ \beta\mathbf{M} & \gamma\mathbf{M} \end{bmatrix}$ has eigenvalues:*

- $\lambda_1 = \gamma N + N\sqrt{\alpha\beta}$, with corresponding eigenvector $\mathbf{u}_1 = \begin{bmatrix} \sqrt{a}\mathbf{1} \\ \sqrt{b}\mathbf{1} \end{bmatrix}$
- $\lambda_2 = \gamma N - N\sqrt{\alpha\beta}$, with corresponding eigenvector $\mathbf{u}_2 = \begin{bmatrix} \sqrt{a}\mathbf{1} \\ -\sqrt{b}\mathbf{1} \end{bmatrix}$
- $\lambda_3 = \dots = \lambda_{2N} = 0$, with corresponding eigenspace spanned by the vectors $\mathbf{u}_k = \begin{bmatrix} \mathbf{t}_k \\ \mathbf{s}_k \end{bmatrix}$, where \mathbf{t}_k and \mathbf{s}_k are $N \times 1$ column vectors with $\varphi(\mathbf{t}_k) = \varphi(\mathbf{s}_k) = 0$, for $k \geq 3$.

Proof. The proof follows directly from lemma 3.

Proposition 3. *The leading (real) eigenvalue of a binary matrix $\mathbf{T} = \begin{bmatrix} \mathbf{P} & \mathbf{A} \\ \mathbf{B} & \mathbf{Q} \end{bmatrix}$, with $\varphi(\mathbf{A}) = \alpha N^2$, $\varphi(\mathbf{B}) = \beta N^2$ and $\varphi(\mathbf{R}) = \varphi(\mathbf{S}) = \gamma N^2$ is of the form:*

$$\gamma N + N\sqrt{\alpha\beta} + \frac{1}{2N} \frac{1}{N + N\sqrt{\alpha\beta}} \mathbf{1}^T \mathbf{W} \mathbf{1} + O(1)$$

with

$$\mathbf{W} = \mathbf{R}^2 + \mathbf{S}^2 + \mathbf{U}\mathbf{V} + \mathbf{V}\mathbf{U} + \frac{\alpha}{\beta}(\mathbf{V}\mathbf{R} + \mathbf{S}\mathbf{V}) + \frac{\beta}{\alpha}(\mathbf{R}\mathbf{U} + \mathbf{U}\mathbf{S}),$$

where $\mathbf{U} = \mathbf{A} - \alpha\mathbf{M}$, $\mathbf{V} = \mathbf{B} - \beta\mathbf{M}$, $\mathbf{R} = \mathbf{P} - \gamma\mathbf{M}$ and $\mathbf{S} = \mathbf{Q} - \gamma\mathbf{M}$ are all matrices with $\varphi(\mathbf{U}) = \varphi(\mathbf{V}) = \varphi(\mathbf{R}) = \varphi(\mathbf{S}) = 0$.

Proof. The proof follows the same steps as equation 2.7. As before, we consider a perturbation $\mathbf{T}_\varepsilon = \mathbf{C} + \varepsilon\mathbf{Z}$ of \mathbf{C} , where $\mathbf{Z} = \begin{bmatrix} \mathbf{R} & \mathbf{U} \\ \mathbf{V} & \mathbf{S} \end{bmatrix}$ has $\varphi(\mathbf{U}) = \varphi(\mathbf{V}) = \varphi(\mathbf{R}) = \varphi(\mathbf{S}) = 0$. We then similarly compute correction terms in the expansion of λ_1 :

$$\mathbf{T}_\varepsilon(\mathbf{u}_1 + \varepsilon\mathbf{x}) = (\lambda_1 + \varepsilon\mu_1)(\mathbf{u}_1 + \varepsilon\mathbf{x}) \quad (\text{B.1})$$

with \mathbf{x} perpendicular to \mathbf{u}_1 . Expanding with respect to ε and identifying the coefficients of ε , we get

$$\mathbf{Z}\mathbf{u}_1 + \mathbf{C}\mathbf{x} = \lambda_1\mathbf{x} + \mu_1\mathbf{u}_1.$$

Expanding $\mathbf{x} = \sum x_k\mathbf{u}_{k'}$, and $\mathbf{Z}\mathbf{u}_j = \sum z_{jk}\mathbf{u}_k$ in the \mathbf{C} eigenvector basis $(\mathbf{u}_k)_{k=1,2N}$ of eigenvectors of \mathbf{C} and rewriting equation B.2 in components, we obtain that $\mu_1 = z_{11}$ and $x_k = \frac{z_{1k}}{\lambda_1 - \lambda_k}$ for $k \geq 2$.

We then calculate

$$\mathbf{Z}\mathbf{u}_1 = \begin{bmatrix} \mathbf{R} & \mathbf{U} \\ \mathbf{V} & \mathbf{S} \end{bmatrix} \begin{bmatrix} \sqrt{a}\mathbf{1} \\ \sqrt{b}\mathbf{1} \end{bmatrix} = \begin{bmatrix} (\sqrt{a}\mathbf{R} + \sqrt{b}\mathbf{U})\mathbf{1} \\ (\sqrt{a}\mathbf{V} + \sqrt{b}\mathbf{S})\mathbf{1} \end{bmatrix}$$

and, in components:

$$\mathbf{Z}\mathbf{u}_1 = z_{11} \begin{bmatrix} \sqrt{a}\mathbf{1} \\ \sqrt{b}\mathbf{1} \end{bmatrix} + z_{12} \begin{bmatrix} \sqrt{a}\mathbf{1} \\ -\sqrt{b}\mathbf{1} \end{bmatrix} + \sum_{k \geq 3} z_{1k} \begin{bmatrix} \mathbf{t}_k \\ \mathbf{s}_k \end{bmatrix}.$$

Summing separately over the top and bottom N entries, we get, respectively,

$$N\sqrt{a}(z_{11} + z_{12}) = \sum_{k \geq 3} z_{1k}\varphi(\mathbf{t}_k) = 0,$$

$$N\sqrt{b}(z_{11} - z_{12}) = \sum_{k \geq 3} z_{1k}\varphi(\mathbf{s}_k) = 0,$$

implying that $z_{11} = z_{12} = 0$ and subsequently $\mu_1 = 0$.

We continue for an $O(\varepsilon^2)$ approximation:

$$\mathbf{T}_\varepsilon(\mathbf{u}_1 + \varepsilon\mathbf{x} + \varepsilon^2\mathbf{y}) = (\lambda_1 + \varepsilon^2\nu_1)(\mathbf{u}_1 + \varepsilon\mathbf{x} + \varepsilon^2\mathbf{y}) \quad (\text{B.2})$$

with \mathbf{y} perpendicular to \mathbf{u}_1 . Identifying the coefficients of ε^3 , we get

$$\mathbf{C}\mathbf{y} + \mathbf{Z}\mathbf{x} = \lambda_1\mathbf{y} + \nu_1\mathbf{u}_1,$$

and, projected along the \mathbf{u}_1 component,

$$\nu_1 = \sum_k \mathbf{x}_k z_{k1} = \sum_{k \geq 2} \frac{z_{1k} z_{k1}}{\lambda_1 - \lambda_k}.$$

Since $z_{12} = 0$ and $\lambda_k = 0$, for $k \geq 3$, this becomes

$$v_1 = \sum_{k \geq 3} \frac{z_{1k} z_{k1}}{\lambda_1} = \frac{1}{N + N\sqrt{\alpha\beta}} \sum_{k \geq 3} z_{1k} z_{k1}. \quad (\text{B.3})$$

If we expand \mathbf{Z}^2 in components as $\mathbf{Z}^2 \mathbf{u}_1 = A_1 \mathbf{u}_1 + A_2 \mathbf{u}_2 + \sum_{k \geq 3} A_k \mathbf{u}_k$, then $\sum_{k \geq 3} z_{1k} z_{k1} = A_1$. To calculate A_1 , we calculate

$$\begin{aligned} \langle \mathbf{u}_1, \mathbf{Z}^2 \mathbf{u}_1 \rangle &= A_1 \|\mathbf{u}_1\| + A_2 \langle \mathbf{u}_1, \mathbf{u}_2 \rangle = N(a+b)A_1 + N(a-b)A_2, \\ \langle \mathbf{u}_1, \mathbf{Z}^2 \mathbf{u}_1 \rangle &= A_1 \langle \mathbf{u}_1, \mathbf{u}_2 \rangle + A_2 \|\mathbf{u}_2\| = N(a-b)A_1 + N(a+b)A_2. \end{aligned} \quad (\text{B.4})$$

Since $\mathbf{Z}^2 = \left[\begin{array}{c|c} \mathbf{R}^2 + \mathbf{UV} & \mathbf{RU} + \mathbf{US} \\ \mathbf{VS} + \mathbf{SV} & \mathbf{VU} + \mathbf{S}^2 \end{array} \right]$, we also have

$$\begin{aligned} \langle \mathbf{u}_1, \mathbf{Z}^2 \mathbf{u}_1 \rangle &= \mathbf{1}^T [\alpha(\mathbf{R}^2 + \mathbf{UV}) + \sqrt{\alpha\beta}(\mathbf{RU} + \mathbf{US} + \mathbf{VR} + \mathbf{SV}) \\ &\quad + \beta(\mathbf{VU} + \mathbf{S}^2)] \mathbf{1}, \\ \langle \mathbf{u}_2, \mathbf{Z}^2 \mathbf{u}_1 \rangle &= \mathbf{1}^T [\alpha(\mathbf{R}^2 + \mathbf{UV}) + \sqrt{\alpha\beta}(\mathbf{RU} + \mathbf{US} - \mathbf{VR} - \mathbf{SV}) \\ &\quad - \beta(\mathbf{VU} + \mathbf{S}^2)] \mathbf{1}. \end{aligned} \quad (\text{B.5})$$

Combining equations B.4 and B.5, we get:

$$A_1 = \frac{1}{2N} \mathbf{1}^T \left[\mathbf{UV} + \mathbf{VU} + \mathbf{R}^2 + \mathbf{S}^2 + \frac{\alpha}{\beta} (\mathbf{VR} + \mathbf{SV}) + \frac{\beta}{\alpha} (\mathbf{RU} + \mathbf{US}) \right] \mathbf{1};$$

hence,

$$v_1 = \frac{1}{2N} \frac{1}{\gamma N + N\sqrt{\alpha\beta}} \mathbf{1}^T \left[\mathbf{UV} + \mathbf{VU} + \mathbf{R}^2 + \mathbf{S}^2 + \frac{\alpha}{\beta} (\mathbf{VR} + \mathbf{SV}) + \frac{\beta}{\alpha} (\mathbf{RU} + \mathbf{US}) \right] \mathbf{1}.$$

As before, however, this computation does not directly estimate the mean or the standard deviation of the eigenvalue distributions, which we instead explore numerically next.

In Figure 15, we illustrate the dependence of the standard deviations simultaneously on the intermodular edge densities α and β (represented on the x - and y -axes) and on the intramodular density γ (different plots in each panel correspond to different values of $\gamma \in [0, 1]$). The figure shows the standard deviation of $\mathcal{L}_1^{\alpha, \beta}$ for $N = 5$, as a function of (α, β) for all discrete values $\gamma \leq 1$. We notice first that although the surfaces do not generally exhibit the same shape and unique ‘‘central’’ maximum as in the particular case of section 2, the unimodality still holds in cross-sections. Moreover, as γ decreases,

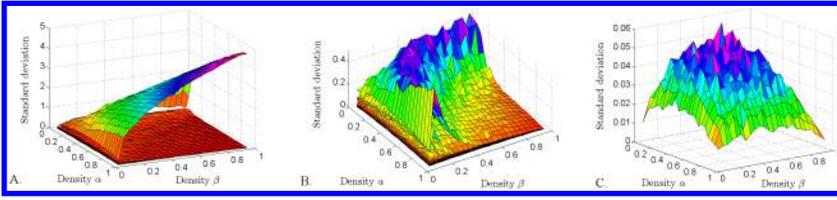


Figure 15: Illustration, for $N = 5$, of the standard deviation of $\mathcal{L}_1^{\alpha,\beta}$ for all values of intramodular connectivity γ . (A) Each surface represents the standard deviation of $\mathcal{L}_1^{\alpha,\beta}$ for one value of $\gamma \in [0, 1]$ (low to high surfaces as γ decreases). For better visualization of the surfaces, we omitted the boundaries $(\alpha, 1)$ and $(1, \beta)$. (B) The panel shows the same surfaces as in panel A, except for the top surface, corresponding to $\gamma = 0$; this was excluded to better illustrate that for all other values of γ , the standard deviations remain small, even with increasing N . (C) The surface corresponding to $\gamma = 1$ (shown in this panel) recovers the results in section 2 (compare with Figure 2C, for $N = 3$; with Figure 3B, for $N = 8$; and with Figure 4A, for multiple N values). The computations were based on sample distributions obtained by considering for each γ a sample of size 100 pairs (\mathbf{P}, \mathbf{Q}) and samples of size 10 for **A** and for **B** for each fixed α and β .

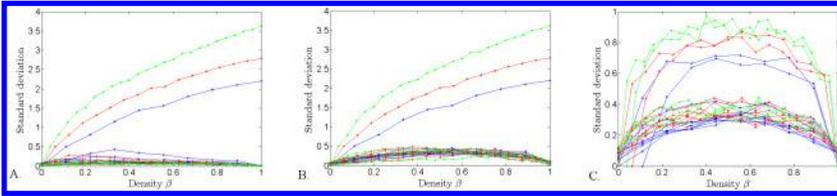


Figure 16: Behavior of the standard deviations when increasing the size N . Shown in blue are the curves for $N = 3$, corresponding to all possible values of $\gamma = k/9$ for $0 \leq k \leq 9$. In red are the curves for $N = 4$, corresponding to $\gamma = 2k/16$, for $0 \leq k \leq 8$. In green are the curves for $N = 5$, corresponding to $\gamma = 5k/25$, for $0 \leq k \leq 5$. For $\mathcal{L}_1^{\alpha,\beta}$ and $\mathcal{L}_2^{\alpha,\beta}$ (panels **A** and **B**), the standard deviations are low, except in the extreme case $\gamma = 0$ (top curve of each color). For $\mathcal{L}_3^{\alpha,\beta}$, the standard deviations remain low for all γ , with a slight increase with N for values of γ close to zero (top two curves of each color). The computations were based on the same sample distributions used for Figure 15.

the standard deviation surfaces rise higher, corresponding to an expectable loss of the system's robustness when decreasing modular cohesion.

However, while the standard deviation values do change with γ , the changes do not appear to be all that significant until γ actually approaches 0. The values are instead bounded by a relatively small upper bound until $\gamma = 0$, when this robustness breaks down. In the case of the leading eigenvalue, the depreciation is monotonous: the standard deviations, very small when

$\gamma = 1$, increase slowly as γ decreases from 1, then faster as the values of γ get close to 0, with a complete crash occurring at $\gamma = 0$ (also see Figure 14). For intermodular connectivity close to saturation (i.e., pairs (α, β) close to the corner $(1, 1)$), the surfaces are barely affected by the intramodular density γ as long as $\gamma > 0$. If one had speculated that the intramodular full connectedness confers robustness to the network eigenvalue spectrum, one would now notice that this robustness is surprisingly well preserved as the full connectedness is gradually loosened by pruning out random edges and thus lowering the intramodular density. The property is completely lost only when the two moduli remain totally disconnected.

The next natural question is to ask, as before, how the robustness of the distributions changes with the size N . In Figures 16B to 16D we show cross-sections of the surfaces introduced in Figure 15 (obtained by fixing one density α), compared for increasing values of N , suggesting that robustness is not substantially affected when the network increases in size, except for values of γ close to zero.

References

- Aicher, C., Jacobs, A. Z., & Clauset, A. (2013). *Adapting the stochastic block model to edge-weighted networks*. arXiv preprint arXiv:1305.5782.
- Ajilore, O., Zhan, L., GadElkarim, J., Zhang, A., Feusner, J. D., Yang, S., . . . Leow, A. (2013). Constructing the resting state structural connectome. *Frontiers in Neuroinformatics*, 7.
- Av-Ron, E., Parnas, H., & Segel, L. A. (1993). A basic biophysical model for bursting neurons. *Biological Cybernetics*, 69(1), 87–95.
- Barabási, A.-L., Gulbahce, N., & Loscalzo, J. (2011). Network medicine: A network-based approach to human disease. *Nature Reviews Genetics*, 12(1), 56–68.
- Barghi, A. R., & Ponomarenko, I. (2009). Non-isomorphic graphs with cospectral symmetric powers. *Electronic Journal of Combinatorics*, 16(R120), 1.
- Borisyuk, G. N., Borisyuk, R. M., Khibnik, A. I., & Roose, D. (1995). Dynamics and bifurcations of two coupled neural oscillators with different connection types. *Bulletin of Mathematical Biology*, 57(6), 809–840.
- Brunel, N. (2000). Dynamics of sparsely connected networks of excitatory and inhibitory spiking neurons. *Journal of Computational Neuroscience*, 8(3), 183–208.
- Bullmore, E., & Sporns, O. (2009). Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience*, 10(3), 186–198.
- Chauhan, S., Girvan, M., & Ott, E. (2009). Spectral properties of networks with community structure. *Physical Review E*, 80(5), 056114.
- Chung, F. R., & Lu, L. (2006). *Complex graphs and networks*. Providence, RI: American Mathematical Society.
- Chung, F., Lu, L., & Vu, V. (2003). Spectra of random graphs with given expected degrees. *PNAS*, 100(11), 6313–6318.
- Cooper, L. (2004). *Theory of cortical plasticity*. Singapore: World Scientific.

- Corbetta, M. (2012). Functional connectivity and neurological recovery. *Developmental Psychobiology*, 54(3), 239–253.
- Craddock, R. C., Jbabdi, S., Yan, C.-G., Vogelstein, J. T., Castellanos, F. X., Di Martino, A., ... Milham, M. P. (2013). Imaging human connectomes at the macroscale. *Nature Methods*, 10(6), 524–539.
- Dannlowski, U., Ohrmann, P., Konrad, C., Domschke, K., Bauer, J., Kugel, H., ... Suslow, T. (2009). Reduced amygdala–prefrontal coupling in major depression: Association with MAOA genotype and illness severity. *International Journal of Neuropsychopharmacology*, 12(1), 11–22.
- Donges, J. F., Zou, Y., Marwan, N., & Kurths, J. (2009). Complex networks in climate dynamics. *European Physical Journal-Special Topics*, 174(1), 157–179.
- Dorogovtsev, S. N., Goltsev, A. V., Mendes, J. F., & Samukhin, A. N. (2003). Spectra of complex networks. *Physical Review E*, 68(4), 046109.
- Elliott, T., & Shadbolt, N. (2002). Multiplicative synaptic normalization and a nonlinear Hebb rule underlie a neurotrophic model of competitive synaptic plasticity. *Neural Computation*, 14, 1311–1322.
- Farkas, I. J., Derényi, I., Barabási, A.-L., & Vicsek, T. (2001). Spectra of “real-world” graphs: Beyond the semicircle law. *Physical Review E*, 64(2), 026704.
- Fekete, T., Wilf, M., Rubin, D., Edelman, S., Malach, R., & Mujica-Parodi, L. R. (2013). Combining classification with fMRI-derived complex network measures for potential neurodiagnostics. *PloS One*, 8(5), e62867.
- GadElkarim, J. J., Ajilore, O., Schonfeld, D., Zhan, L., Thompson, P. M., Feusner, J. D., ... Leow, A. D. (2014). Investigating brain community structure abnormalities in bipolar disorder using path length associated community estimation. *Human Brain Mapping*, 35(5), 2253–2264.
- Gilbert, F., Simonetto, P., Zaidi, F., Jourdan, F., & Bourqui, R. (2011). Communities and hierarchical structures in dynamic social networks: Analysis and visualization. *Social Network Analysis and Mining*, 1(2), 83–95.
- Goodhill, G. (1993). Topography and ocular dominance: A model exploring positive correlations. *Biological Cybernetics*, 69(2), 109–118.
- Gray, R. T., & Robinson, P. A. (2009). Stability and structural constraints of random brain networks with excitatory and inhibitory neural populations. *Journal of Computational Neuroscience*, 27(1), 81–101.
- He, Y., & Evans, A. (2010). Graph theoretical modeling of brain connectivity. *Current Opinion in Neurology*, 23(4), 341.
- Holland, P. W., Laskey, K. B., & Leinhardt, S. (1983). Stochastic blockmodels: First steps. *Social Networks*, 5(2), 109–137.
- Juhasz, F. (1990). On the characteristic values of non-symmetric block random matrices. *Journal of Theoretical Probability*, 3(2), 199–205.
- Kim, M. J., Loucks, R. A., Palmer, A. L., Brown, A. C., Solomon, K. M., Marchante, A. N., & Whalen, P. J. (2011). The structural and functional connectivity of the amygdala: From normal emotion to pathological anxiety. *Behavioural Brain Research*, 223(2), 403–410.
- Kim, M. J., & Whalen, P. J. (2009). The structural integrity of an amygdala—prefrontal pathway predicts trait anxiety. *Journal of Neuroscience*, 29(37), 11614–11618.
- Larremore, D. B., Clauset, A., & Jacobs, A. Z. (2014). Efficiently inferring community structure in bipartite networks. *Physical Review E*, 90(1), 012805.

- Leow, A., Ajilore, O., Zhan, L., Arienzo, D., GadElkarim, J., Zhang, A., . . . Altshuler, L. (2013). Impaired inter-hemispheric integration in bipolar disorder revealed with brain network analyses. *Biological Psychiatry*, *73*(2), 183–193.
- Martin-Hernandez, J., Wang, H., Van Mieghem, P., & D'Agostino, G. (2013). *On synchronization of interdependent networks*. arXiv preprint arXiv:1304.4731.
- Miller, K., & MacKay, D. (1994). The role of constraints in Hebbian learning. *Neural Computation*, *6*(1), 100–126.
- Mucha, P. J., Richardson, T., Macon, K., Porter, M. A., & Onnela, J.-P. (2010). Community structure in time-dependent, multiscale, and multiplex networks. *Science*, *328*(5980), 876–878.
- Nadakuditi, R. R., & Newman, M. (2012). Graph spectra and the detectability of community structure in networks. *Physical Review Letters*, *108*(18), 188701.
- Nadakuditi, R. R., & Newman, M. E. (2013). Spectra of random graphs with arbitrary expected degrees. *Physical Review E*, *87*(1), 012803.
- Oja, E. (1982). Simplified neuron model as a principal component analyzer. *Journal of Mathematical Biology*, *15*(3), 267–273.
- Olfati-Saber, R., Fax, J. A., & Murray, R. M. (2007). Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, *95*(1), 215–233.
- Olfati-Saber, R., & Murray, R. M. (2004). Consensus problems in networks of agents with switching topology and time-delays. *IEEE Transactions on Automatic Control*, *49*(9), 1520–1533.
- O'Reilly, R., & Frank, M. (2006). Making working memory work: A computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation*, *18*(2), 283–328.
- Rădulescu, A., & Mujica-Parodi, L. R. (2013). Network connectivity modulates power spectrum scale invariance. *NeuroImage*, *90*, 436–448.
- Rădulescu, A., & Verduzco-Flores, S. (2015). Nonlinear network dynamics under perturbations of the underlying graph. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *23*(1), 013116.
- Rahmani, A., Ji, M., Mesbahi, M., & Egerstedt, M. (2009). Controllability of multi-agent systems from a graph-theoretic perspective. *SIAM Journal on Control and Optimization*, *48*(1), 162–186.
- Restrepo, J. G., Ott, E., & Hunt, B. R. (2007). Approximating the largest eigenvalue of network adjacency matrices. *Phys. Rev. E*, *76*, 056119.
- Sandhu, J., Mesbahi, M., & Tsukamaki, T. (2005). Relative sensing networks: Observability, estimation, and the control structure. In *Proceedings of Decision and Control, 2005 and 2005 European Control Conference, 44th IEEE Conference* (pp. 6400–6405). Piscataway, NJ: IEEE.
- Sarkar, S., Henderson, J. A., & Robinson, P. A. (2013). Spectral characterization of hierarchical network modularity and limits of modularity detection. *PloS One*, *8*(1), e54383.
- Siri, B., Quoy, M., Delord, B., Cessac, B., & Berry, H. (2007). Effects of Hebbian learning on the dynamics and structure of random networks with inhibitory and excitatory neurons. *Journal of Physiology—Paris*, *101*, 136–148.
- Small, M., Judd, K., & Stemler, T. (2012). *The stability of networks—towards a structural dynamical systems theory*. arXiv preprint arXiv:1206.2145.

- Smith, R., & Hadaegh, F. (2007). Closed-loop dynamics of cooperative vehicle formations with parallel estimators and communication. *IEEE Transactions on Automatic Control*, 52(8), 1404–1414.
- Sporns, O. (2002). Graph theory methods for the analysis of neural connectivity patterns. In R. Kötler (Ed.), *Neuroscience databases: A practical guide* (pp. 171–186).
- Sporns, O. (2010). *Networks of the brain*. Cambridge, MA: MIT Press.
- Sporns, O. (2011). The non-random brain: Efficiency, economy, and complex dynamics. *Frontiers in Computational Neuroscience*, 5.
- Supekar, K., Menon, V., Rubin, D., Musen, M., & Greicius, M. D. (2008). Network analysis of intrinsic functional brain connectivity in Alzheimer's disease. *PLoS Computational Biology*, 4(6), e1000100.
- Tao, T. (2012). *Topics in random matrix theory*. Providence, RI: American Mathematical Society.
- Toga, A. W., Clark, K. A., Thompson, P. M., Shattuck, D. W., & Van Horn, J. D. (2012). Mapping the human connectome. *Neurosurgery*, 71(1), 1.
- Turrigiano, G., Leslie, K., Desai, N., Rutherford, L., & Nelson, S. (1998). Activity-dependent scaling of quantal amplitude in neocortical neurons. *Nature*, 892–895.
- Turrigiano, G., & Nelson, S. (2004). Homeostatic plasticity in the developing nervous system. *Nature Reviews Neuroscience*, 5(2), 97–107.
- Van Mieghem, P. (2011). The n -intertwined SIS epidemic network model. *Computing*, 93(2–4), 147–169.
- von der Malsburg, C. (1973). Self-organization of orientation sensitive cells in the striate cortex. *Kybernetik*, 14, 85–100.
- Willshaw, D., & von der Malsburg, C. (1976). How patterned neural connections can be set up by self-organization. *Proceedings of the Royal Society of London, Series B: Biological Sciences*, 194(1117), 431–445.
- Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal*, 12(1), 1–24.
- Wu, C. W. (2013). Control of networks of coupled dynamical systems. In L. Kocarev (Ed.), *Consensus and synchronization in complex networks* (pp. 23–50). New York: Springer.
- Zanin, M., & Lillo, F. (2013). Modelling the air transport with complex networks: A short review. *European Physical Journal Special Topics*, 215(1), 5–21.
- Zelazo, D. L. (2008). *Analysis and synthesis of networked dynamic systems*. Doctoral dissertation, University of Washington.