Adaptive Rewiring on Logistic Maps with Heterogeneous Parameters

Author Name(s), First M. Last, Omit Titles and Degrees

Institutional Affiliation(s)

Author Note

Include any grant/funding information and a complete correspondence address.

Abstract

The abstract should be one paragraph of between 150 and 250 words. It is not indented. Section titles, such as the word Abstract above, are not considered headings so they don’t use bold heading format. Instead, use the Section Title style. This style automatically starts your section on a new page, so you don’t have to add page breaks. (To see your document with pagination, on the View tab, click Reading View.) Note that all text styles for this template are available on the Home tab of the ribbon, in the Styles gallery.

Keywords: Add keywords here. To replace this (or any) tip text with your own, just select it and then start typing. Don’t include space to the right or left of the characters in your selection.

Adaptive Rewiring on Logistic Maps with Heterogeneous Parameters

# Introduction

Brain network architecture is shaped dynamically through structural plasticity (Butz, Wörgötter, & van Ooyen, 2009). Structural plasticity encompasses a variety of mechanisms, all of which incorporate the *functional connectivity* between network components, i.e. their mutual statistical dependencies in neural activity (Avena-Koenigsberger, Misic, & Sporns, 2018; Rubinov, Sporns, van Leeuwen, & Breakspear, 2009). This common principle has become known as *adaptive rewiring* (Gong & van Leeuwen, 2003, 2004; Papadopoulos, Kim, Kurths, & Bassett, 2017). Adaptive rewiring implements the Hebbian principle of “what fires together, wires together” at the level of network dynamics.

Adaptive rewiring can be modeled in abstract dynamical systems, in which the network components may represent units at the scale of single neurons (Bi & Poo, 2001) or neural masses (Breakspear, Terry, & Friston, 2003). In the latter case, the activity is described as a chaotic oscillator governed by the attractor in Figure S1A in the supplementary materials. Dimensional reduction via Poincaré section yields the relationship in Figure S1B that can approximately be described, minus the noise, by a logistic map Fig. S1C). Thus, the logistic map is the simplest possible abstract representation of neural mass activity. Logistic maps are known to exhibit universal dynamical properties (van Strien, 1987).

The logistic map is of the form shown in Equation 1.

EQ 1 ,

in which *x* is a continuous variable in the range <-1,1> which is updated in discrete time *t*, and is the *turbulence parameter*. For certain regimes of , the behavior of the logistic map converges to one or more limit-cycle attractors, but otherwise it exhibits chaotic behavior. In these regimes, logistic maps produce deterministic chaotic bounded time series that, indeed, qualitatively resemble the oscillations of neural mass activity (see Fig. S1).

Whereas the logistic map could be considered as an abstract representation of neural mass activity, systems of coupled neural mass oscillators may be represented by coupled logistic maps. Because of the universal dynamics of logistic maps, networks of such simple maps may capture generic properties of interacting nonlinear systems (Kaneko, 1992). The logistic maps are coupled according to Equation 2. Through the effect of the neighbors, the map activity of Fig. S1C regains a noisy appearance more in line with Fig. S1B. In matrix notation, for a network with |V| nodes, the activity of nodes at time t+1 is calculated via

EQ 2

The right-hand side of Equation 2 constitutes the vector form of the logistic map, in which is the vector of turbulence parameters and X\_t is the vector of node's activities at time t. The map is Hadamard-multiplied[[1]](#footnote-2) by a coupling term. In the coupling term, is the vector of coupling strengths, is the connectivity matrix at time t, and denotes a vertical unit vector of size |V|. In the coupling term, is Hadamard-divided by , normalizing the former by the sum of the weights of the edges connected to each node. For binary networks, the term in the denominator counts the number of connections for each node. We consider only binary graphs for convenience. For adaptive rewiring in weighted networks, see Hellrigel, Jarman, and van Leeuwen (2019).

Based on the network activity as defined by Equation 2, adaptive rewiring takes the following form: after several updates to the network activity, a rewiring step is made. At each rewiring step, the connections of a random node are optimized as the node is disconnected from the neighbor most dissimilar in activity and is connected to the most similar nodes to which it was not connected before. The dissimilarity of two nodes at a given time is defined as the absolute value of the difference in the value of their activity. Note that although rewiring steps are local, for simplicity, nonlocal information is used for finding the most dissimilar unconnected nodes. For algorithms using local, or rather, regional information for this purpose, see Jarman et al. (2017) and Jarman et al. (2014).

According to the adaptive rewiring principle, network structure evolves over time from random to complex architectures, showing the characteristics of small world, modularity, and the rich club effect (Hellrigel et al., 2019; Gong & van Leeuwen, 2003; 2004; Rubinov et al., 2009). Hence, at least according to these global structural characteristics, adaptively rewiring networks evolve brain-like structures, as small worldness (Sporns & Zwi, 2004), modularity (Meunier, Lambiotte, & Bullmore, 2010) and the rich club effect (van de Heuvel & Sporns, 2011) are characteristics of large-scale brain networks. Adaptive rewiring may thus be considered to capture, in a highly simplified form, the common principle of structural plasticity mechanisms in the brain.

All model studies of adaptive rewiring mentioned so far have, for simplicity, assumed the coupling strength and turbulence to be uniform across the system. This assumption severely reduces the cognitive functionality of such models. In earlier logistic map network studies using fixed architectures, these parameters have been varied locally to represent perceptual and memory functions. In a network model to model perceptual organization (van Leeuwen, Steyvers, & Nooter, 1997), a sensory input function was realized by modulation of the turbulence parameters of the local oscillators. Presence of sensory input brought these parameters down to values imposing a more stable regime on the oscillators. As a result, connected units receiving similar inputs showed a synchronization bias, leading to perceptual grouping preferences. In a memory model (van Leeuwen & Raffone, 2001), connectivity parameter values were locally incremented to represent the presence of a memory trace. This, too, established synchronization biases, leading to spontaneous pattern rehearsals and subsequent relearning of dynamic memory traces.

When networks having cognitive functions are to evolve brain-like structures through adaptive rewiring, adaptive rewiring must be robust to non-uniform parameter values. We will explore the effect of non-uniform turbulence and coupling parameters on adaptive rewiring of coupled logistic maps. In particular, we will compare networks of coupled logistic maps with uniform parameters with those that have a subset of the turbulence and coupling parameters deviate from the majority value. We study the effect of non-uniform parameters on network evolution and final network structure.

(Transitioning sentences/paragraph to portray the layout of the paper?) In what follows, the method section describes details of the composition and initialization of the models; the rewiring algorithm; and the qualitative and quantitative measures of network structures used to describe, characterize, and compare models. In the results section, we describe our findings, mainly that ….. A discussion ends the paper.

# Method

### Description of networks

An undirected, binary graph (or network) *G* is a set of 3-tuples of vertices (or nodes) and and an edge (or connection) between them. The connection can assume values of {0,1}.This set is called adjacency list, wherein, conventionally, zero valued elements are omitted.. We assume no self-connection, i.e. *eii* = 0. The set of edges and vertices of *G* are represented by E and V, respectively. *G* can be graphically depicted by circles representing V connected by line segments representing E. The adjacency matrix *M* of *G* is a square matrix of the size |V| with elements . Since G is undirected, M is a symmetrical around the main diagonal.

**Dynamics on the graph**

To each , an activation value is assigned according to Equation 2. The corresponding parameters values, i.e., coupling strength and turbulence remain fixed in our models. Models with identical parameter sets are called families. Five families of models are simulated, each with ten model instantiations, comprising a total of 50 model instantiations. Each network was ran for 20 million iterations. All the simulations and analyses are conducted in R programming language version 3.5.0 (R Core Team, 2018).

### Parameter setting and Initialization

In our models, all *G* have |V| = 300 and . Model structure is initialized by randomly assigning values "1" symmetrically to non-diagonal entries of *M*, and zeros to the remaining entries. Each node in the network is randomly assigned an initial value, uniformly distributed between 0 and 1, i.e., . The 10 model instantiations within each family were ran with different initializations, which were identical between families to allow matched comparison.

Previous research (citation) has used coupled logistic maps with and in the ranges of [1.7-1.9] and [0.3-05], respectively. Here the midpoints of the ranges are assumed the baseline values for the parameters. Nodes with lower and higher values of and are called under- and over-turbulent and under- and over-coupled, respectively.

The nodes were assigned to two subsets: minority (the first 50 nodes) and majority (the remaining 250 nodes). While keeping the parameters of the majority at the baseline level () five different combinations of parameters were assigned to the minorities, each combination called a "family": The baseline family (BL; ), and the families with under-turbulent minority (UT; ), over- turbulent family (OT; ), under-coupled minorities (UC; ), and over-coupled minority (OC; ). As shown in Figure S2, the logistic maps show chaotic behavior for the three levels chosen for turbulence parameter, as such higher values of yield more chaotic activity. The nodes within each subset and the edges between them form subgraphs within G, henceforth called minority and majority partitions. A third subgraph is assumed as interpartition, which comprises of all nodes but with only the edges connecting minority nodes to the majority nodes. In the Results section, we identify model instantiations by the two capitals indicating their family, together with a serial number [1-10], e.g. BL7, OT10.

### Adaptive rewiring algorithm

A rewiring attempt takes place after every 20 updates of the logistic maps, meaning that over the 20 million updates of the model, one million rewiring attempts are performed. At each rewiring attempt, at time t, a node i is selected randomly from V, a vector[[2]](#footnote-3) of its distance from other nodes is calculated as , and another vector of similarities is defined as .

The most dissimilar neighbor and the most similar non-neighbor of node i, respectively denoted as and , respectively, are marked by finding the index of the maxima of the following vectors:

The matrix multiplication of M and (1-M) with , respectively, ensures that the search for the edges subject to rewiring happens in the right subset of edges. The rewiring is then changing the corresponding elements of the adjacency matrix:

## Characterizing and comparing models

The state of each model at any given time *t* is described by adjacency matrix ( (henceforth, "anatomical connectivity") which is subject to adaptive rewiring, and the vector of activation values, (). Model's "functional connectivity" at *t*, (with adjacency matrix ) is defined by the momentary pairwise differences of node activation values.

### Qualitative description of network structures.

Network structure can be qualitatively assessed by means of visual inspection of the graph diagram or the adjacency matrix. The adjacency matrices are serialized using … algorithm, implemented in the package `seriation` (Hahsler et al., 2008), which orders the rows and columns of the matrix to maximize visual identifiability of modules within the network (better/more precise phrasing?).

### Quantitative measures of the structure.

In network science, a wide range of structural measures of connectivity, also known as network statistics, have been proposed (Costa, Rodrigues, Travieso, & Villas Boas, 2007). After each rewiring attempt, we calculate six network connectivity measures: clustering coefficient, average path length, small-worldness, modularity, assortativity, and edge density Furthermore, we calculate an additional measure, namely, rich club coefficient, after the last rewiring attempt. All measures are calculated separately for the whole and the three subgraphs (viz.,minority and majority partition and the interpartition subgraph). Unless mentioned otherwise, the `igraph` package (citation) is used for calculating the measures.

#### Clustering coefficient.

This measure can be defined either locally or globally and gives an indication for the tendency of nodes to form clusters. We use the global clustering coefficient, which is defined as the number of closed triplets of nodes (the triplets of nodes that are all connected) divided by the number of connected triplets, either open (i.e., paths of length 2) or closed (i.e., triangles). The numerator is equal to three times the number of triangles in the graph. Using linear algebra, one can show that the global clustering coefficient can be calculated formally from the adjacency matrix as shown in Equation 3.

EQ 3

#### Average path length.

Average path length is the mean value of lengths of shortest path between all pairs of nodes. This measure, calculated using the `igraph` package, gives an indication of how closely the nodes of a network are located from each other.

#### Small-worldness.

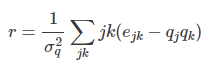
Small-worldness is a measure of the degree to which the graph shows properties akin to the structures known as small world (citation). It is defined as the multiplication of normalized clustering coefficient and efficiency of the network, i.e.. Efficiency of a network quantifies the efficiency of information exchange within the network. It is defined as sum of inverses of the distances between nodes, normalized by network size. More formally, for a network with N edges, efficiency is calculated as . in the denominator of are the expected clustering coefficient and efficiency of a random graph of the same as the graph in question. Since all networks modeled in this study start off with random networks of the same size and density, for computational reasons, a non-normalized version of small-worldness coefficient (i.e., ) is calculated and reported.

#### Modularity.

Modularity of a graph, as proposed by Newman (2006) and denoted by Q, is a measure of how (and to what degree), for a certain labeling of nodes, the nodes tend to form communities with the nodes of the same label and tend to not connect to other nodes of the graph. This measure requires a priori labeling of nodes that defines the communities to which they are believed to belong. Algorithmic discovery of modules enables the measures of modularity is maximized (for a review of the proposed methods, cf. Zhang, Ma, Zhang, Sun, & Yan, 2018). The communities discovered by these algorithms can thus be used as labels for calculating modularity of the network. In line with (Clauset, Newman, & Moore, 2004), the fast greedy algorithm used to optimally detect communities and calculate the modularity based on community membership of the nodes.

#### Assortativity.

Assortativity coefficient is a measure of homophily in networks that indicates the preferences of nodes to connect to "similar" nodes by summarizing the probability of connections between the similar nodes. The similarity can be imposed externally, e.g., by assigning categories to the nodes using labels (known as nominal assortativity), or by internal criteria like the node degrees (degree assortativity). Degree assortativity, measures the probability that nodes of similar degree (i.e., number of connections) are connected. (Newman, 2003) Degree assortativity of undirected networks can be calculated as shown in Equation 4, where is the degree of node j and …

EQ 4 

#### Edge density

For a subset of nodes, this coefficient is the proportion of existing edges in a (sub)graph to the maximum number of edges possible in the same (sub)graph. Since the total number of edges remains the same during the adaptive rewiring, this coefficient gives an indication of how strongly each partition has attracted new nodes at every rewiring step.

#### Rich Club coefficient.

This coefficient quantifies the tendency of nodes with higher than a certain degree to connect to each other. More formally, as Equation 5 shows, it is equivalent to the edge density of the subgraph of the network where the nodes with lower degrees than the cut-off value k are removed. Since this coefficient is a function of club size k, it is hard to visualize its evolution over time for all possible values of k. Hence, the values of this coefficient were only plotted for the final state of the networks. We use the `brainGraph` package (citation) to calculate rich club coefficient.

EQ 5 {Citation}

The absolute value of the rich club coefficient is hard to interpret and is not comparable among networks of different sizes, densities, and degree distributions. Hence, this coefficient is often normalized by the average rich club coefficient of random networks of the same size with similar degree sequence. We simulate 200 such networks for each model and, for each club size k, normalized the values of the rich club coefficients of the model by the average among the random networks. Moreover, for each k, we performed 1-sample t-test to determine if the non-normalized RC is significantly higher than its average among the randomly generated networks.

For a certain k, the normalized rich club coefficient is larger than one indicates that nodes with degree k tend to connect to the "rich" nodes (i.e., those with degrees equal to or higher than k), thus forming "rich clubs." Conversely, implies that the nodes with degree k have a tendency to connect to nodes with lower degrees. Finally, suggests that nodes with degree k show no preference to connect to nodes with lower or higher degrees.

### Investigating resemblance between models

To studying the effect of heterogeneous parameterization of the models on their structures, we need to compare families with each other. In line with Mheich et al. (2018), we assume that the structural information embedded in networks can be summarized by the distributions of local network measures. Comparison of networks is thus reduced to comparing these distributions. To obtain measures of distributional distances, we use NetSimile (Mheich et al., 2018) and HHG (Heller et al., 2013) methods, detailed in the Supplementary Materials. NetSimile suggests the degree of dissimilarity between the distributions attributed to the networks being compared. HHG, on the other hand, provides p-values for a test of independence among the distributions; lower p-value derived from HHG (e.g., below the threshold of $alpha = 0.05$) provides evidence in favor of distributional dependence. Thus, loosely speaking, HHG can be regarded as an indicator for dissimilarity; higher values of this measures entail smaller "resemblance" (or dependence) between the networks. Yet, HHG must be interpreted with caution and is hardly meaningful unless put in parallel with another dissimilarity measure such as NetSimile.

We first make pairwise comparisons among the 1225 unique pairs of model instantiations at their final state after 1 million rewiring attempts. Having quantitative measures for dissimilarities among the networks, we quantifiy the within-family resemblances and between-family contrasts among the models. Finally, we define a measure for family distinction and compare how families vary with respect to this measure.

#### Family resemblances and differentiations

The outcomes of pairwise comparison of the networks using NetSimile and HHG were stored in four 50X50 matrices of form for networks N (either M or F, respectively for anatomical and functional connectivities) based on method m (either NetSimile or HHG). To ease the visual comparison of these measures, the matrices of NetSimile distances, i.e., and , were normalized by the highest value in each matrix so their values range from zero to one. Then, within- and between-family contrast aggregate scores were calculated by averaging the elements of dissimilarity matrices that belong to the families being compared as shown in Equation 6.

EQ 6

Finally, a differentiation score was calculated for each family to quantify the degree to which models belonging to family $f\_i$ resemble each other and, at the same time, diverge from the members of other families, as shown in Equation 7.

EQ 7

In this equation, the numerator is the within-family resemblance of networks belonging to family $f\_i$. The denominator is the mean of between-family resemblance of $f\_i$ to other families. This score will be larger than one if family $f\_i$ differentiates from other families.

# Results

## Network structures

Using the `seriation` and `igraph` packages, we plot the raw (unserialized) and ordered (serialized) adjacency matrices and the graph diagrams of the anatomical connectivities at their final state. In the plots, the minority and majority nodes are colored sky blue and pink, respectively. In both matrix visualizations and graph diagrams, the within-minority and within-majority edges are colored blue and red, respectively. The inter-partition edges, connecting nodes of minority subset to nodes of majority, are colored green. Although there are variations among models, either within- or between-families, in all models (except for the OT2, OT3, UC1, and UC3; see below), several densely coupled sets of nodes, i.e., modules, emerged. These modules can be identified as squares in the serialized adjacency matrices. All plots are available online on Open Science Framework (<https://osf.io/625d8/>). Figure 1 shows two representative networks per family. As evident in the plots, the networks manifest a range of different structures. Yet, similarities can be observed among models belonging to the same families.

The baseline models (BL2 and BL8, Fig 1) typically include three densely coupled modules, and a few larger, sparser sets of nodes. The modules are not isolated from the rest of the network, as inter-modular edges keep them connected to other nodes The unserialized adjacency matrices show that the density of edges are quite uniform over subsets of nodes.

In the OC family, wherein the minority nodes have relatively higher coupling strengths, the edge density is higher in the minority subgraph. Moreover, the OC networks have more distinct modules than the BL models. OC2 (Fig 1), for instance, only has two modules, both of which are highly connected. Higher edge density in the minority subset and highly distinct modules are also the case for UT family, which has lower turbulence parameter in the minority nodes (UT5 and UT 8, Fig 1). The similarity between OC and UT models was expected, as the nodes with lower turbulences and higher coupling strengths can synchronize more easily. However, it is interesting to note that the effect of better synchrony is not limited to the minority nodes; highly connected modules also emerge among the majority nodes.

In the OT family, the edge density is lower in the minority partition. Moreover, the higher level of turbulence parameter for minority nodes resulted in highly connected modules among the majority nodes (see OT4 in Fig 1). In the UC family, the edge density of the minority (which had lower coupling strengths) is lower than that of the majority. Moreover, lower coupling strength of the minority prevented minority nodes to form modules, and they were absorbed into modules formed mainly by the majority nodes. See UC5 and UC7 in Fig 1.

## Network statistics

The evolution of clustering coefficient, modularity, edge density, small-world index, assortativity, and average path length were plotted for anatomical networks of all models and their sub-networks (viz., within-minority, within-majority, and interpartition). Figure 2 shows these plots for the minority, majority, and whole networks, stacked per family. The individual plots (including those of the interpartition sub-network) are available on the OSF repository. Furthermore, the normalized rich club coefficient of the final state of anatomical networks are plotted in Fig 3.

Let us first consider the evolution of network statistics for the whole networks. As evident in the plots, modularity, clustering coefficients, small-worldness, and degree assortativity of all models increase noticeably after 60 thousand rewiring attempts (roughly by a factor of 4.5, 3.7, 3.9, and 3, respectively). Meanwhile, the average path length increases no more than 15%. The network statistics fluctuate afterward, and the fluctuation is more pronounced for degree assortativity. The majority subnetworks show similar trends. In the minority subnetworks—although with higher fluctuations and instability in OC, UT, and UC—we observe an increase in these statistics after 60 thousand rewiring attempts. The network statistics of minority subnetworks of OT and BL families resemble those of the whole networks in the same families. The edge density in the minority subnetworks, although unstable, are often bigger than one. This means that the minority nodes tend to connect more strongly to each other compared to other nodes.

Figure 3 shows the normalized rich club coefficient of the anatomical networks, grouped by families, as a function of club size k. above one (dashed line) suggest existence of rich clubs. The values significantly larger than one (p < 0.05, based on 1-sample t-test) are marked by solid circles. Despite differences among families, they all show significant rich club structures for larger club sizes. More specifically, OC has clubs of sizes 80<k<145, most of them significant, with consistent values among members. Almost half of UT models form significantly rich clubs with 75<k<150 yet others fail to. Other families, i.e., OT, BL, and UC, form clubs with of sizes ranging from 70 to roughly 135, with higher (but less consistent) values compared to OC and UT. Fewer of are significant in OT, BL, and UC compared to OC.

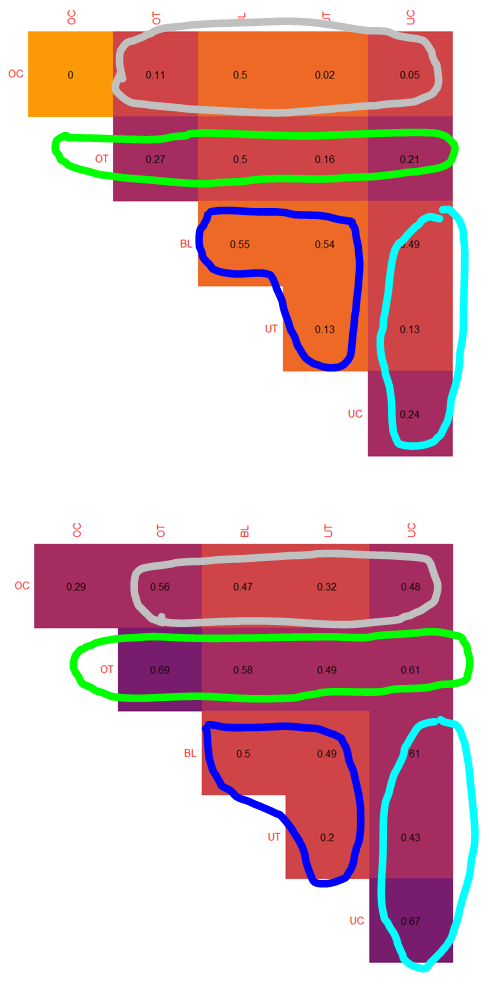
## Odd cases

As evident in the evolution plots, four models (viz., OT2, OT3, UC1, and UC3) stop evolving before 10 thousand rewiring attempts. Visual inspection of their anatomical connectivity reveals that, at some point in their evolution, one node reaches maximum degree and is connected to all other nodes (Figure XX). Consequently, in the next rewiring step, an element of the coupled logistic map reaches singularity because of a zero by zero division which is mathematically undefined and represented as `NaN`, i.e., "not a number," in R language. Since the node activations update through matrix algebra, the `NaN` spreads to other elements of activation vector $X\_t+1$, causing a halt in activities.

## Family comparisons

To more closely observe how NetSimile and HHG similarity measures diverge, they were composed in new matrices wherein the lower triangle belongs to and the upper triangle belongs to , i.e., . The matrices of are plotted in Figure 4 (left) as heat maps using the `ComplexHeatmap` R package (Gu et al., 2016). In this figure, lower dissimilarity (hence higher similarity) measures are coded by brighter colors.

The matrices of are plotted as upper triangular matrices in the Figure 4 (right). The cell colors, coded similarly to the heat maps, denote average contrast measures derived from NetSimile algorithm while average HHG p-values (i.e.,) are indicated in each cell. As can be seen in this figure, the within- and between-family NetSimile contrasts of both anatomical and functional networks are in agreement. (Better phrasing?) More specifically, CANNOT PUT IN WORDS, LOOK AT THIS PLOT.



The differentiation scores for both anatomical and functional connectivities of families are plotted in Figure 5. Differentiation values above one (dashed line) imply that the within-family resemblance of network structures of family $f\_i$ is higher than the average resemblance of its members to the members of other families. We observe elevated differentiations in both anatomical and functional networks of OC, BL, and UT. This measure is remarkably higher for the anatomical network of OC family.

Finally, in order to have both family resemblance and family differentiation in a single frame, we summarized their values in the graphs shown in Figure 6. In these graphs, individual nodes represent families of models. Edge color and size code between-family contrast and node color capture within-family contrast. The size of each node is proportional to the value of the differentiation score of its corresponding family. The families with are marked with asterisks. It can be noticed that

# Discussion

TO BE WRITTEN

Start with a brief summary of the main results

-BL networks evolve in accordance with the literature to SWN with modular and rich club structure?

- The other families do so too(?) to what extent?

-Evolution of these properties is robust against changes to the turbulence and connectivity parameters, so these parameters can be used to implement cognitive functions in evolving networks

-The families differentiate, both in structure and function from the BL models (to different extents; which ones more?

-Differentiation can be used to identify functional components in these networks. This warrants the use of structural and functional connectivity measures in neuro-imaging.

# References

Avena-Koenigsberger, A., Misic, B., & Sporns, O. (2018). Communication dynamics in complex brain networks. *Nature Reviews Neuroscience,* 19, 17.

Butz, M., Wörgötter, F., & van Ooyen A. (2009). Activity-dependent structural plasticity *Brain Research Review,* 60, 287-305.

Gong, P. & van Leeuwen, C. (2003). Emergence of scale-free network with chaotic units. *Physica A, Statistical mechanics and its applications*, *321*, 679–688.

Gong, P. & van Leeuwen, C. (2004). Evolution to a small-world network with chaotic units *Europhysics Letters,* 6, 328–333.

Hellrigel, S., Jarman, N., & van Leeuwen (2019). Adaptive rewiring of weighted networks. *Cognitive Systems Research, 55*, 205-218.

Jarman, N., Steur, E., Trengove, C., Tyukin, I., & van Leeuwen, C. (2017). Self-organisation of small-world networks by adaptive rewiring in response to graph diffusion. *Scientific Reports*, 7, 13518.

Jarman, N., Trengove, C., Steur, E., Tyukin, I., & van Leeuwen, C. (2014). Spatially constrained adaptive rewiring in cortical networks creates spatially modular small world architectures. *Cognitive Neurodynamics, 8, 479-497.*

Papadopoulos, L., Kim, J.Z., Kurths, J., & Bassett, D.S. (2017). Development of structural correlations and synchronization from adaptive rewiring in networks of kuramoto oscillators, *Chaos,* 27, 073115.

van Leeuwen, C. & Raffone A. (2001). Coupled nonlinear maps as models of perceptual pattern and memory trace dynamics. *Cognitive Processing*, *2*, 67–116.

Rubinov, M., Sporns, O., van Leeuwen, C. & Breakspear, M. (2009).Symbiotic relationship between brain structure and dynamics. *BMC Neuroscience, 10*, 55.

van Leeuwen, C., Steyvers, M., & Nooter, M. (1997). Stability and intermittency in large-scale coupled oscillator models for perceptual segmentation. *Journal of Mathematical Psychology, 41*, 319–344.

# Figures

Figure 1. Network structures of representative models.

Figure 2. Evolution of network statistics

Figure 3. Rich club coefficients

Figure 4. Heat maps of pairwise similarities of networks and family comparisons.

*Figure 5*. Differentiation score of families for the anatomical and functional networks.

*Figure 6*. Graph representation of family resemblance and differentiation.

# Supplementary Materials

## Network Comparison

Dozens of methods have been proposed to be used in network comparison (for a review, see citation needed). The well-known methods are adapted to cases where the nodes are labeled (i.e., distinguishable) a priori. These methods are not always suitable for non-labeled networks (wherein the nodes are not assigned to any labels, which is the case for our models), as they require a primary step of inferring (or estimating) node labels or matching the nodes for the networks being compared (also known as node correspondence problem), which is computationally very costly (citation needed).

### Networks as distributions

As we have seen, there are numerous network statistics that can describe network structures either locally or globally. The local measures are suitable for node-wise (or clique-wise) comparisons, while the global measures are aggregates of some local properties that provide "summary statistics" for the structure. The local measures hardly lead to holistic description of networks as the nodes are usually described in isolation from other nodes. On the other hand, valuable structural information is sacrificed along aggregation taking place in deriving global structural measures. Therefore, neither local nor global measures are optimal for comparison of networks. A solution to this issue would be taking an intermediary (?) approach by putting the distribution of various local measures under the spotlight. This way, we get to keep all[[3]](#footnote-4) the information we can harvest from structure. Then, we can decide how to use this information in network comparison.

Berlingerio et al. (2012) suggest characterizing each node i of the network with a seven-dimensional feature vector consisting of the following local measures that capture characteristics of the node and its surrounding subset of the network: , degree (i.e., number of neighbors); , local clustering coefficient (i.e., the number of triangles connected to node i over the number of connected triples centered on node i); , average degree of node’s neighbors; , average clustering coefficient of node’s neighbors; , the number of edges in the egonet[[4]](#footnote-6) of node i; , the number of outgoing edges from the egonet of node i[[5]](#footnote-7); and , the number of neighbors of egonet of node i. Although one can add more local features to this vector, Berlingerio and colleagues have shown that these features suffice for decent comparison of networks in their algorithm, i,e, NetSimile. Having the feature vectors of all nodes, we reach a 7-variate distribution which can be used to compare networks.

### The NetSimile method.

Given the distributions of local features, one can compare the distributions by means of comparing their summary statistics. In NetSimile, the feature distribution (which is a nodesXfeatures matrix) is aggregated into a 35-dimensional signature vector consisting of five summary statistics for each feature: median, mean, standard deviation, skewness, and kurtosis. The comparison of networks is thus reduced to calculating distances (or similarities) of the signature vectors. NetSimile is superior to other methods of inferring network similarity as its computational complexity grows linearly with the size of the networks, and more importantly, it allows comparison of networks of different sizes. One now must define an appropriate method of calculating the distances.

As we know (citation needed?), the ranks of summary statistics characterize the overall shape of distributions thus is a highly discriminative metric in their comparison. Then, the values of these summary statistics provide information about the similarities among the distributions on top of their overall shapes. Hence, the signature vectors are akin to ranked lists. It has been shown that the Canberra distance, defined in Equation S1, is an appropriate measure of dissimilarity for ranked lists (Jurman, Riccadonna, Visintainer, & Furlanello, 2009) as it is sensitive to small distances from zero and normalizes the pairwise distances of features by their absolute values. Moreover, Berlingerio and colleagues (2012) report high discriminative power of Canberra distance in comparison of signature vectors, a good property of a dissimilarity for the task at hand.

EQ S1

In this study, we used this dissimilarity metric in the pairwise comparison of the signature vectors derived from NetSimile algorithm. However, NetSimile does not allow hypothesis testing to infer significance levels for the distances. Berlingerio et al. (2012) suggest hypothesis testing for independence of the distributions by pairwise comparison of the univariate distributions of the features and aggregating their p-values through averaging or choosing the maximum values. They report that neither Mann-Whitney nor Kolmogorov-Smirnov tests—which are nonparametric tests without any assumption for the distributions being compared—yield amply meaningful discrimination among the networks being compared. Their approach of hypothesis testing ignores the multivariate dependencies of the features. Hence, we use another method to test independence of distributions that is discussed below.

### Hypothesis testing for similarities of network distributions.

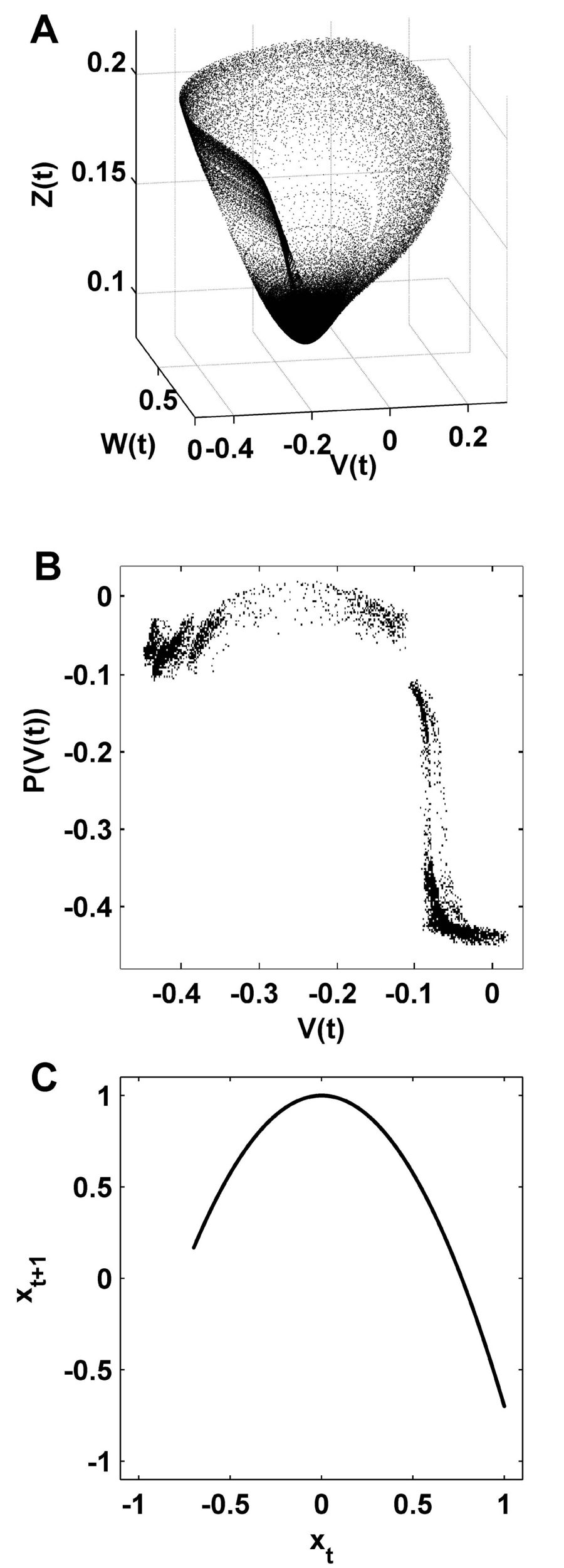
The significance tests used by Berlingerio and colleagues (2012) posits multivariate independence among features and lacks what they call "discrimination power". To tackle this issue, one needs use multivariate dependence tests. Since parametric dependence tests rely on assumption for the distributions being compared, we used HHG nonparametric permutation test of multivariate dependence (Heller, Heller, & Gorfine, 2013) implemented in `HHG` R package (Brill, Heller, & Heller, 2018). HHG is a consistent omnibus test for the null hypothesis that the joint distribution of two multivariate random variables X and Y is equal to the multiplication of the marginal distributions of those variables. Equation S2 shows the null and alternative hypotheses:

EQ S2

HHG has a reasonable computational complexity and uses norm distance matrices of the samples taken from X and Y separately. The technical details of this method are beyond the scope of this paper. In short, HHG iteratively forms hyper spheres in the joint space of and based on the implications of the null hypothesis, quantifies evidence against by likelihood ratio or Pearson’s Chi-square tests statistics over the contingency tables. From these tests, one can drive permutation p-values that can be interpreted as evidence against null hypothesis of the independence of the distributions. Hence, the lower the p-value, the more evidence favoring the dependence of the distributions being compared. Loosely speaking, one can treat the p-values derived from HHG methods as a form of distributional dissimilarity; the lower the value, the more "similar" the distributions are to each other. This interpretation (is it an interpretation, or some other term?) is rather unorthodox (?) in hypothesis testing, yet it allows us to compare non-significant p-values as relative measures of resemblance.

The `hhg.test()` function in `HHG` package runs the test for a number of permutations on distance matrices of the samples in X and Y and outputs four different permutation p-values based on sums or maximum values of likelihood ratio or Chi-square test scores of al 2X2 contingency tables (elaborate? Clarify?). In this study, we let HHG run for 2000 permutations for each pairwise comparison and extracted permutation p-value for the maximum of likelihood ratio score statistics as it yielded higher discriminative power compared to other test statistics.

## Supplementary figures



**Figure S1. From Rubinov et al., 2009. Dimension reduction of nonlinear neuronal dynamics**. (A). Phase space attractor of a three-dimensional neural mass flow. This attractor is an illustration of the dynamics generated by the flow of a neural mass model (see Breakspear et al. [[33](https://bmcneurosci.biomedcentral.com/articles/10.1186/1471-2202-10-55#ref-CR33)]). The dynamical variables represent the mean membrane potential of pyramidal (*V*) and inhibitory (*Z*) neurons, and the average number of open potassium ion channels (*W*). (B). Poincaré first return map from the same attractor [[33](https://bmcneurosci.biomedcentral.com/articles/10.1186/1471-2202-10-55#ref-CR33)]; this map captures key features of the neural mass flow, by following each trajectory from one intersection (*V*) of the attractor to the next (*P*(*V*)). (C). The quadratic logistic map. This map has the same unimodal topology as the neural mass Poincaré return map. While the logistic map lacks the "thickness" of the neural mass map, it is several orders of magnitude faster to compute, hence allowing the detailed quantitative analysis in the present paper.

Figure S2. Feigenbaum diagram of the values of the 200 draws of logistic maps (after a burn-in period of 4000 iterations)

1. Also known as elementwise multiplication of matrices where the corresponding elements of matrices are multiplied. Hadamard division is defined similarly. [↑](#footnote-ref-2)
2. While programming the analyses, a matrix of distances was calculated to increase the versatility of the code for other rewiring algorithms. [↑](#footnote-ref-3)
3. Note that, since the nodes are not labeled and are indistinguishable, the order of node-specific values for local measures hold no information. [↑](#footnote-ref-4)
4. Egonet of node i, referred to as ego(i), is the subset of the network including node i, its first-order neighbors (N(i)), and the edges among N(i). [↑](#footnote-ref-6)
5. Since our networks are undirected, this value would be equal to . [↑](#footnote-ref-7)