

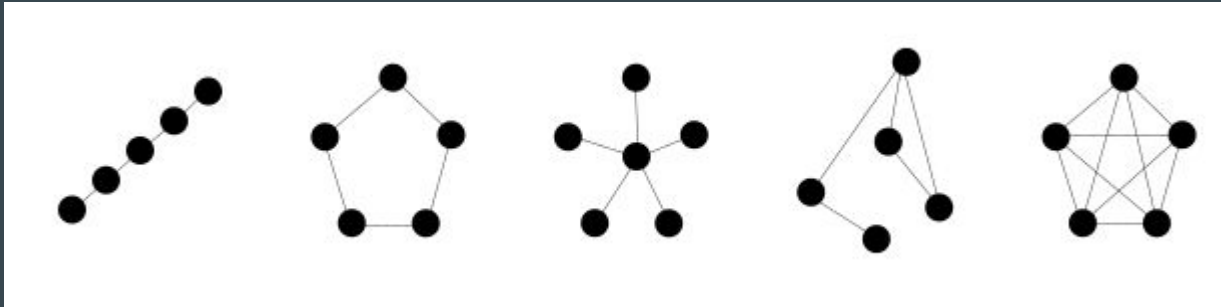
Graphs and Networks

...

Graph analysis

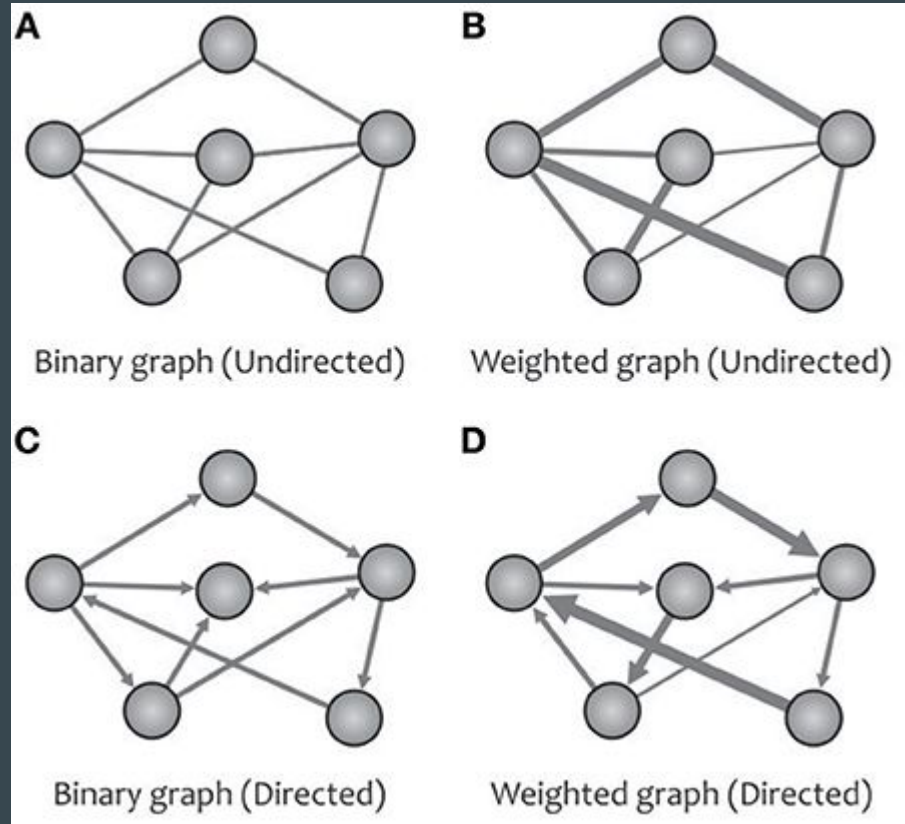
Graphs

Graphs are widely used to model complex systems and analyze data in social networks, biology, computer science. However, characterizing graph complexity is a challenging task as it is a subjective measure that can vary depending on the context and the observer. There is no universal metric for graph complexity, and different measures may emphasize different aspects of the graph's properties.



Graphs

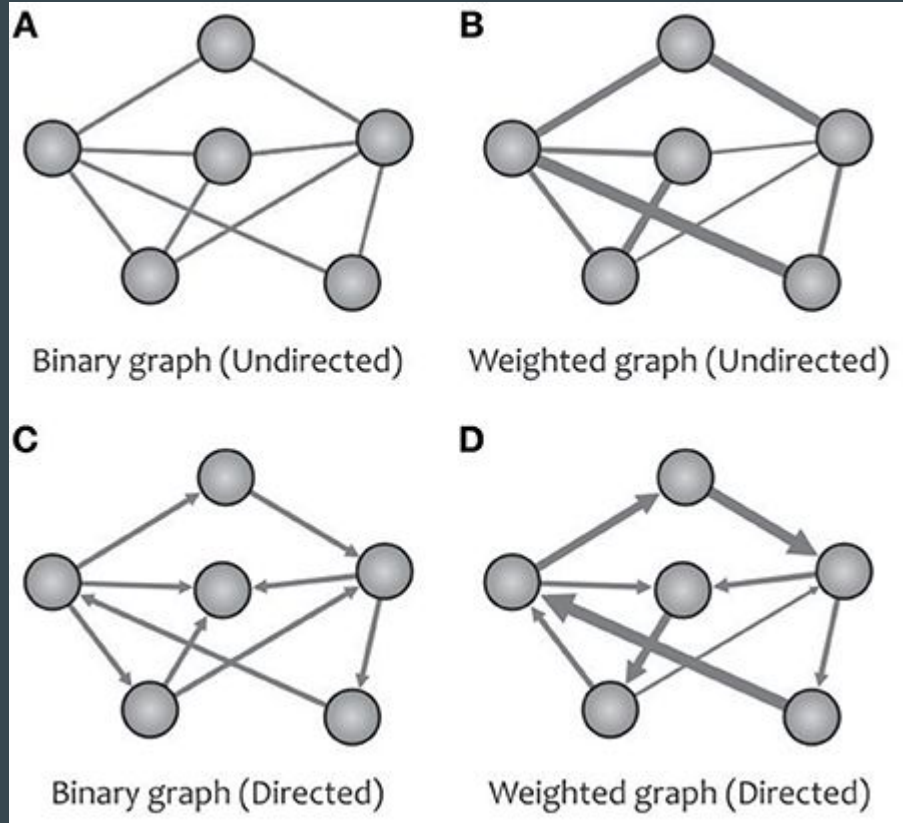
Graphs are typically high-dimensional objects, which are hard to visualize and analyze. In addition, they can exhibit emergent properties that arise from the interactions between their nodes, which can be difficult to predict. Therefore, characterizing a graph requires a better understanding of the underlying structure and properties of the graph, and there is no unique approach to this problem.



Graphs

Graphs can be described by their connectivity matrices.

There are several different formats for connectivity matrices, as the graphs can be undirected, directed or weighted.



Connectivity matrices: adjacency matrix

For undirected graphs, the adjacency matrix is symmetric and it represents the connections between vertices i, j :

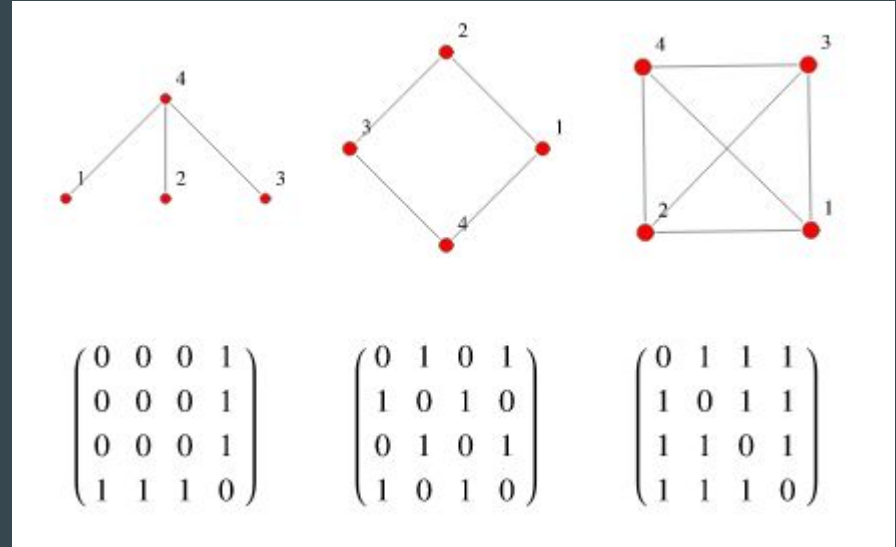
$A(i,j) = 1$, if (i,j) is an edge

0, if (i,j) not an edge

If the graph has weights then the adjacency matrix is respectively:

$A(i,j) = w_{ij}$, if (i,j) is an edge

0, if (i,j) not an edge



Connectivity matrices: adjacency matrix

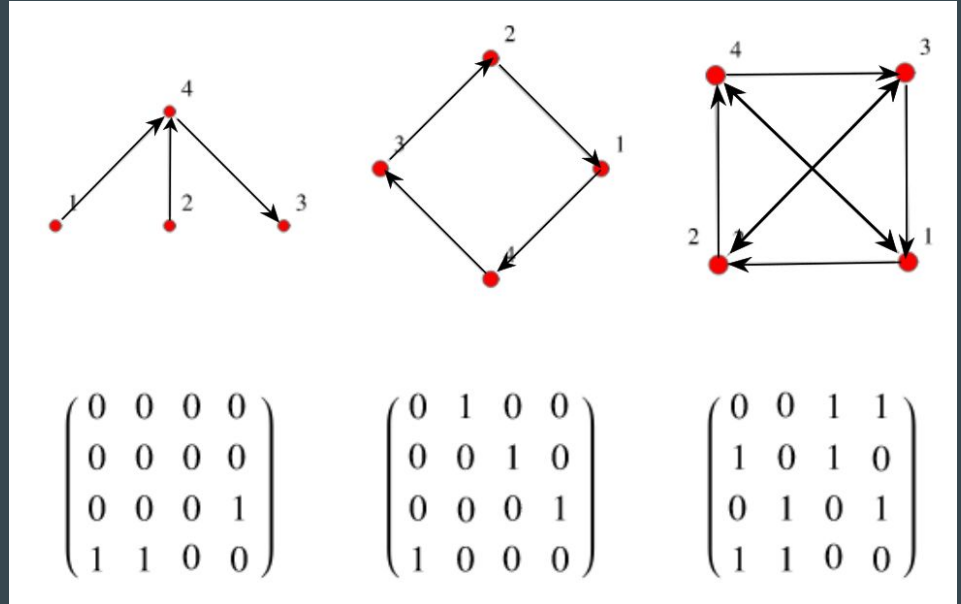
For directed graphs, the adjacency matrix is not necessarily symmetric and it represents the connections between vertices $i \rightarrow j$:

$A(i,j) = 1$, if $(i \rightarrow j)$ is an edge

0, if $(i \rightarrow j)$ not an edge

For example if $(2 \rightarrow 3)$ is an edge but $(3 \rightarrow 2)$ is not then

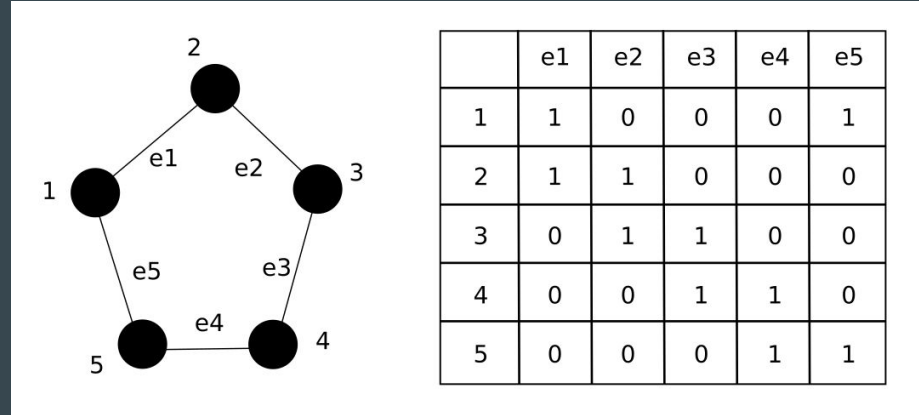
$A(2,3) = 1$ while $A(3,2) = 0$



Connectivity matrices: incidence matrix

For undirected graphs, the incidence matrix is an $m \times n$ matrix that represents the relationship between vertices and edges

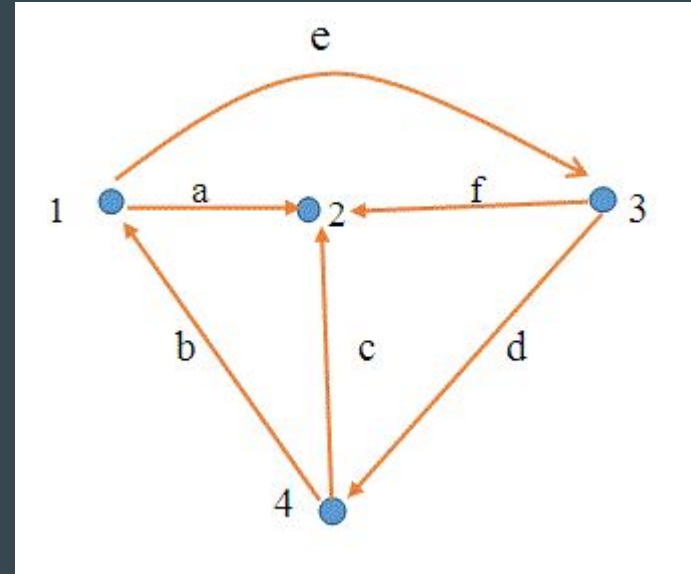
$I(e_i, v_j) = 1$, if e_i is participating in v_j
0, if e_i is not in v_j



Connectivity matrices: incidence matrix

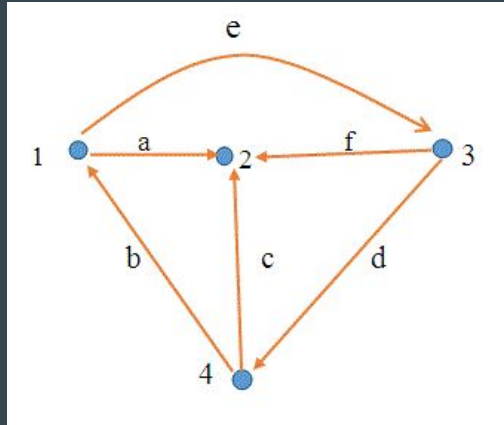
For directed graphs, the incidence matrix is an $m \times n$ matrix that represents the relationship between vertices and edges

$I(e_i, v_j) = 1$, if e_i is the first element in v_j
- 1, if e_i is the second element in v_j
0, if e_i is not in v_j



Connectivity matrices: incidence matrix

For example, for directed graphs the sum of each column is 0 (for undirected it is 2)



branches nodes	a	b	c	d	e	f
1	1	-1	0	0	1	0
2	-1	0	-1	0	0	-1
3	0	0	0	1	-1	1
4	0	1	1	-1	0	0

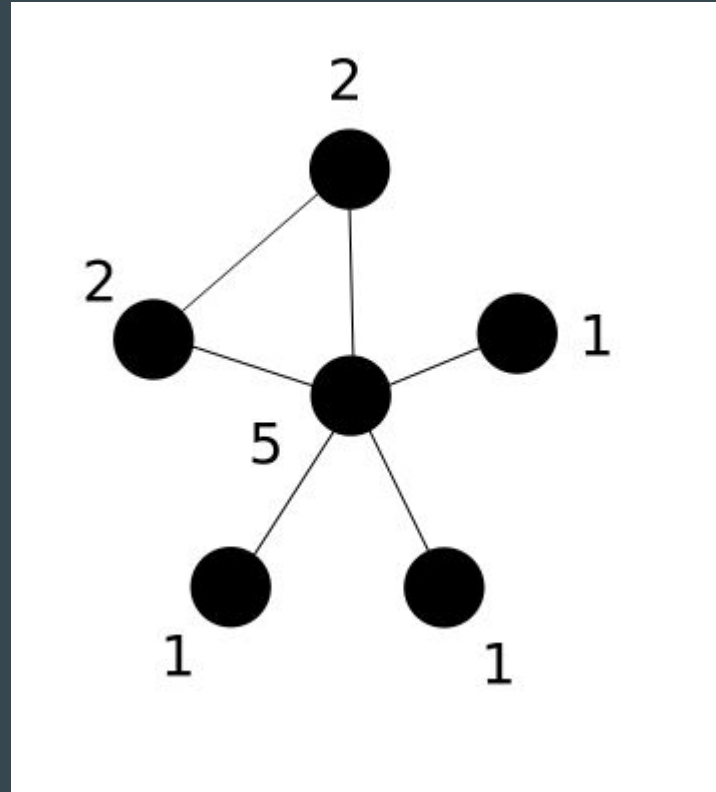
Measurements on graphs : degree of vertices

The degree of a vertex describes the number of edges it participates in.

Given a graph $G(V, E)$, the degree of a vertex v_i is the sum of all its edges:

$$k_i = \sum_{j \in N} a_{ij}.$$

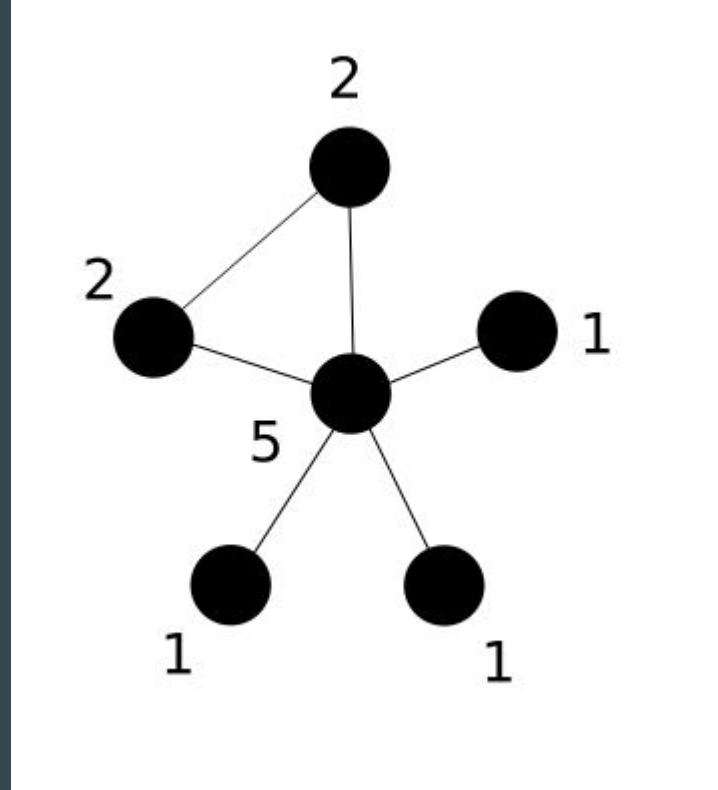
The degree can also be computed efficiently from the incidence matrix. For undirected graphs the degree of a vertex is given by the sum of its rows.



Measurements on graphs : degree of vertices

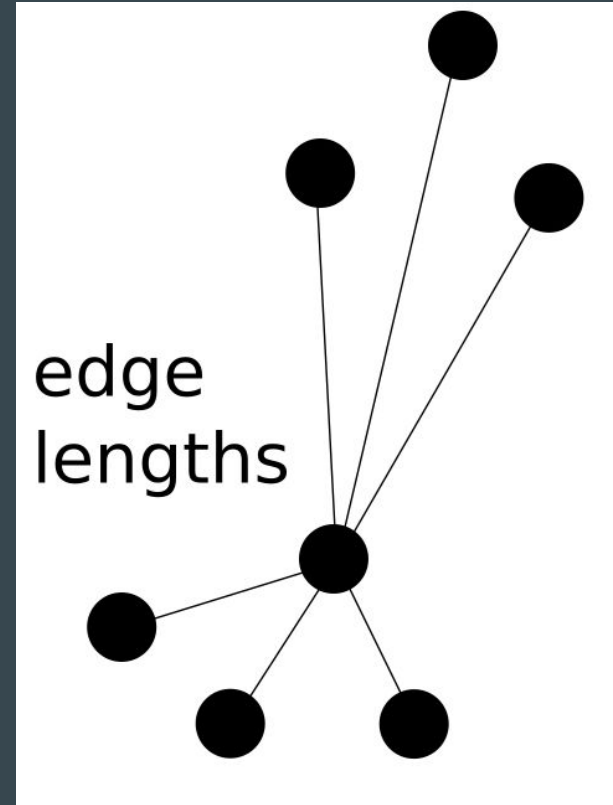
The sum of all degrees in an unweighted graph $G(V, E)$ is twice the number of edges in the graph:

$$\sum_{v \in V} \deg(v) = 2|E|$$



Measurements on graphs : edge length

Edge length is meaningful in weighted graphs and the length usually refers to the weights of the edges. The edge length distribution is a measure of how the lengths are spread out across the range of possible edge lengths and it can provide insights into the structure of a graph, such as the presence of clusters or hubs.



Measurements on graphs : clustering coefficient

The local clustering coefficient of a vertex (node) in a graph quantifies how close its neighbours are to being a clique (complete graph).

For a graph $G=(V, E)$ the neighbourhood N_i for a vertex v_i is defined as its immediately connected neighbours as follows:

$$N_i = \{v_j : e_{ij} \in E \vee e_{ji} \in E\}$$

Measurements on graphs : clustering coefficient

The local clustering coefficient C_i of a vertex v_i is defined as the fraction of the number of connections t_i between the vertices within its neighbourhood N_i divided by the number of all possible connections that could exist between them. For a directed graph, because e_{ij} not equal to e_{ji} for each neighbourhood N_i with k_i neighbors there are $k_i * (k_i - 1)$ possible connections that could exist among them. Thus, the local clustering coefficient for directed graphs is given as

$$C_i = \frac{t_i}{k_i(k_i - 1)}$$

Measurements on graphs : clustering coefficient

For a undirected graph, because e_{ij} is equal to e_{ji} for each neighbourhood N_i with k_i neighbors there are $k_i * (k_i - 1) / 2$ possible connections that could exist among them. Thus, the local clustering coefficient for undirected graphs is given as

$$C_i = \frac{2t_i}{k_i(k_i - 1)}$$

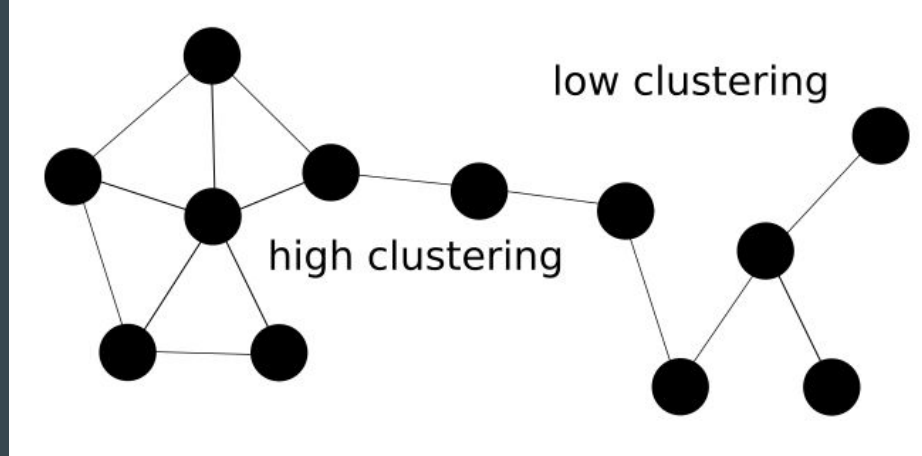
Measurements on graphs : clustering coefficient

For an undirected network, in which N is the set of all vertices in the network, and n is the number of vertices, the clustering coefficient is

$$C = \frac{1}{n} \sum_{i \in N} C_i = \frac{1}{n} \sum_{i \in N} \frac{2t_i}{k_i(k_i - 1)},$$

Measurements on graphs : clustering coefficient

The local clustering coefficient of a vertex (node) in a graph quantifies how close its neighbours are to being a clique (complete graph). High clustering coefficient indicates larger number of connections between neighbors, low indicates small number of connections

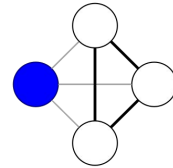


Measurements on graphs : clustering coefficient

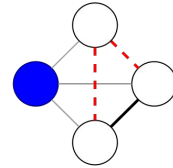
The local clustering coefficient of a vertex (node) in a graph quantifies how close its neighbours are to being a clique (complete graph).

For clustering coefficient 1 the neighborhood is a clique

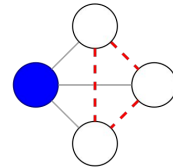
For a clustering coefficient 0 no neighbors are connected



$$c = 1$$



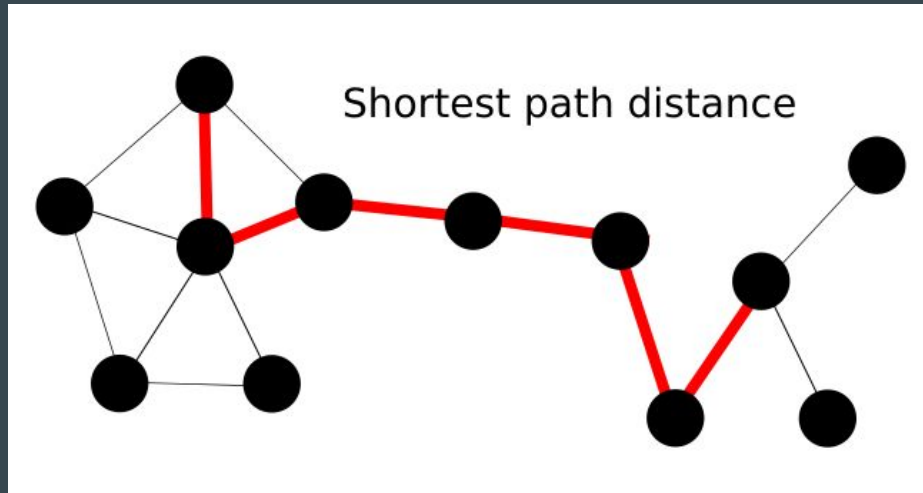
$$c = 1/3$$



$$c = 0$$

Measurements on graphs : shortest path length

In graph theory, the shortest path length is a path between two vertices (or nodes) in a graph such that the sum of the weights of its constituent edges is minimized.



Measurements on graphs : shortest path length

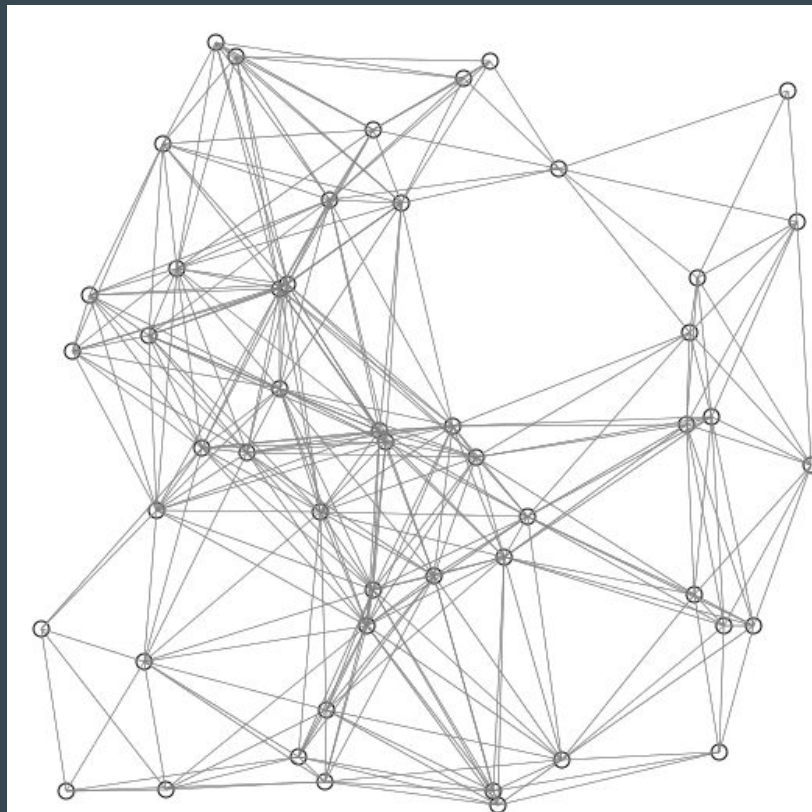
In graph theory, the shortest path length is defined by the path between two vertices in a graph such that the sum of the weights of its constituent edges is minimized.

$$d_{ij} = \sum_{a_{uv} \in g_{i \leftrightarrow j}} a_{uv},$$

Where $g_{i \leftrightarrow j}$ is the shortest path that connects i and j . For a pair of paths that are disconnected d_{ij} is infinite.

Measurements on graphs : shortest path distance

An example of an algorithm to compute the shortest path distances in a graph is shown.



Measurements on graphs : betweenness centrality

Betweenness centrality is a measure of how “central” is a node within a graph, based on the shortest paths. For every pair of vertices in a connected graph, there exists at least one shortest path between the vertices such that either the number of edges that the path passes through (for unweighted graphs) or the sum of the weights of the edges (for weighted graphs) is minimized. The betweenness centrality of a vertex is the number of shortest paths that pass through the vertex.

Measurements on graphs : betweenness centrality

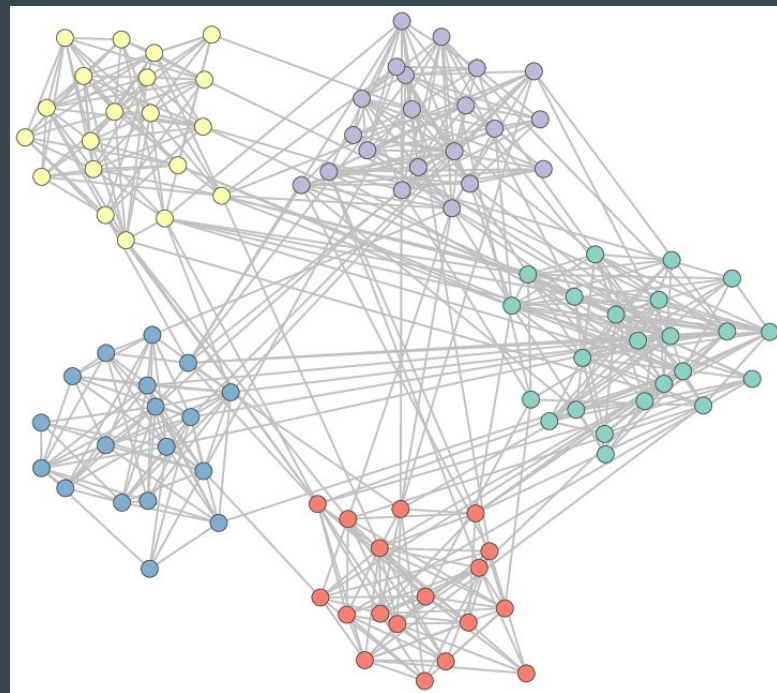
The betweenness centrality of a vertex is the number of shortest paths that pass through the vertex.

$$b_i = \frac{1}{(n-1)(n-2)} \sum_{\substack{h,j \in N \\ h \neq j, h \neq i, j \neq i}} \frac{\rho_{hj}(i)}{\rho_{hj}},$$

Where ρ_{hj} is the number of shortest paths between h and j , and $\rho_{hj}(i)$ is the number of shortest paths between h and j that pass through i .

Measurements on graphs : modularity

Networks can be divided into modules, which represent different functional subnetworks. However, there are many different ways to define modules within a network and there is no optimal way for all types of networks.



Measurements on graphs : modularity

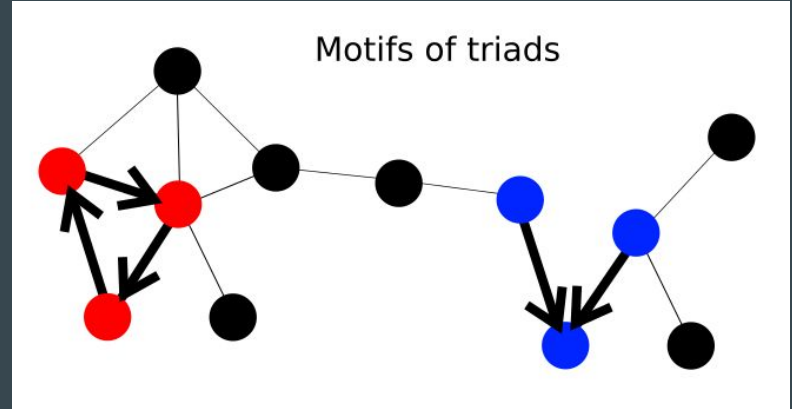
Modularity of the network (Newman, 2004), can be computed as

$$Q = \sum_{u \in M} \left[e_{uu} - \left(\sum_{v \in M} e_{uv} \right)^2 \right],$$

where the network is fully subdivided into a set of nonoverlapping modules M , and e_{uv} is the proportion of edges that connect vertices in module u with vertices in module v .

Measurements on graphs : motifs

Network motifs are recurrent and statistically significant subgraphs or patterns of a larger graph. Each of these sub-graphs, defined by a particular pattern of interactions between vertices, may reflect a framework in which particular functions are achieved efficiently. They have recently gathered much attention as a useful concept to uncover structural design principles of complex networks. Although network motifs may provide a deep insight into the network's functional abilities, their detection is computationally challenging.

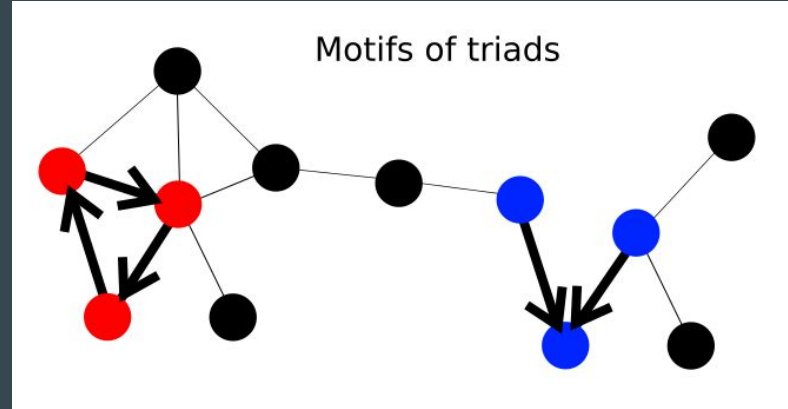


Measurements on graphs : motifs

For example, network triangles around a vertex i are defined as:

$$t_i = \frac{1}{2} \sum_{j,h \in N} a_{ij} a_{ih} a_{jh}.$$

Which is a measurement of segregation in the network.



Measurements on graphs : motifs

J_h is the number of occurrences of a motif h in all subsets of the network (subnetworks). This number can be compared to the probability of motif to appear randomly, and respectively the z-score of a motif can be computed

$$z_h = \frac{J_h - \langle J_{\text{rand},h} \rangle}{\sigma J_{\text{rand},h}}$$

where $\langle J_{\text{rand},h} \rangle$ and $\sigma J_{\text{rand},h}$ are the respective mean and standard deviation for the number of occurrences of h motif in an ensemble of random networks.

Measurements on graphs: small world networks

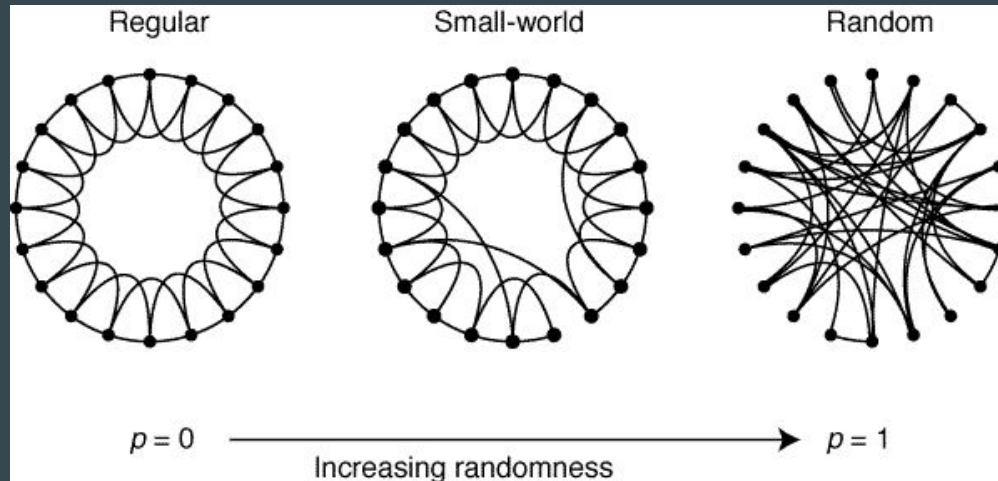
Small-worldness of a network has been quantified by a small-coefficient, σ , calculated by comparing clustering and path length of a given network to an equivalent random network with the same average degree

$$\sigma = \frac{\frac{C}{C_r}}{\frac{L}{L_r}}$$

Effectively this measurement quantifies how “far” from random is a given network. However, it is highly influenced by the size of the network and therefore it is not necessary a reliable metric to characterize a network. See Collective dynamics of ‘small-world’ networks, Watts & Strogatz 1998

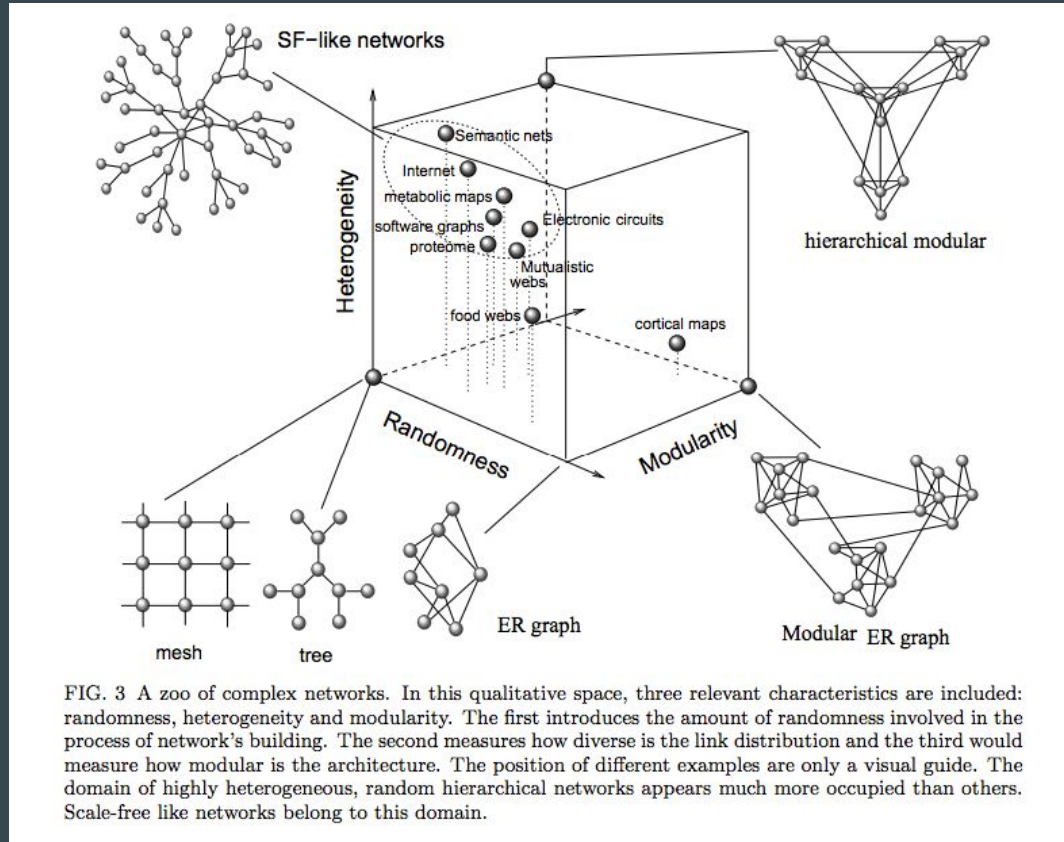
Measurements on graphs: small world networks

Random rewiring: start with a ring of n vertices, each connected to its k nearest neighbours by undirected edges. ($n = 20$, $k = 4$). Choose a vertex and the edge that connects it to its nearest neighbour in a clockwise sense. With probability p , we reconnect this edge to a vertex chosen uniformly at random over the entire ring, considering each vertex in turn until one lap is completed.



Measurements on graphs: small world networks

Various types
of Networks,
Sole and
Valverde
(2004).

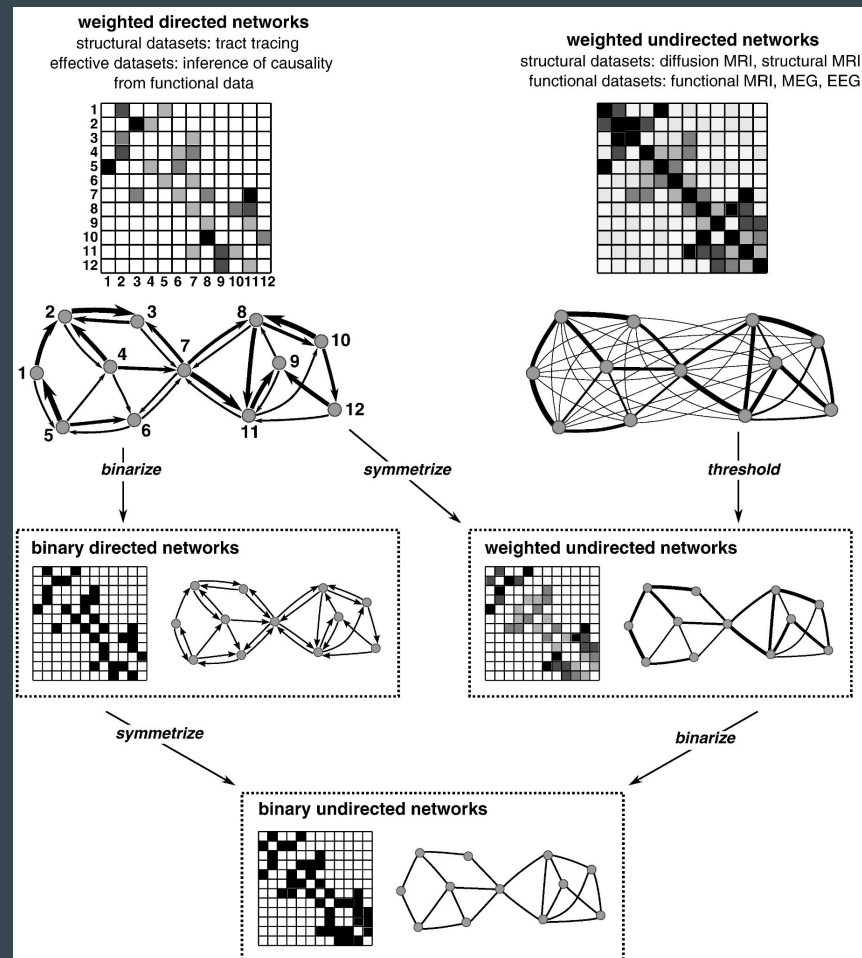


Complex network measures of brain connectivity: Uses and interpretations

Rubinov and Sporns 2010

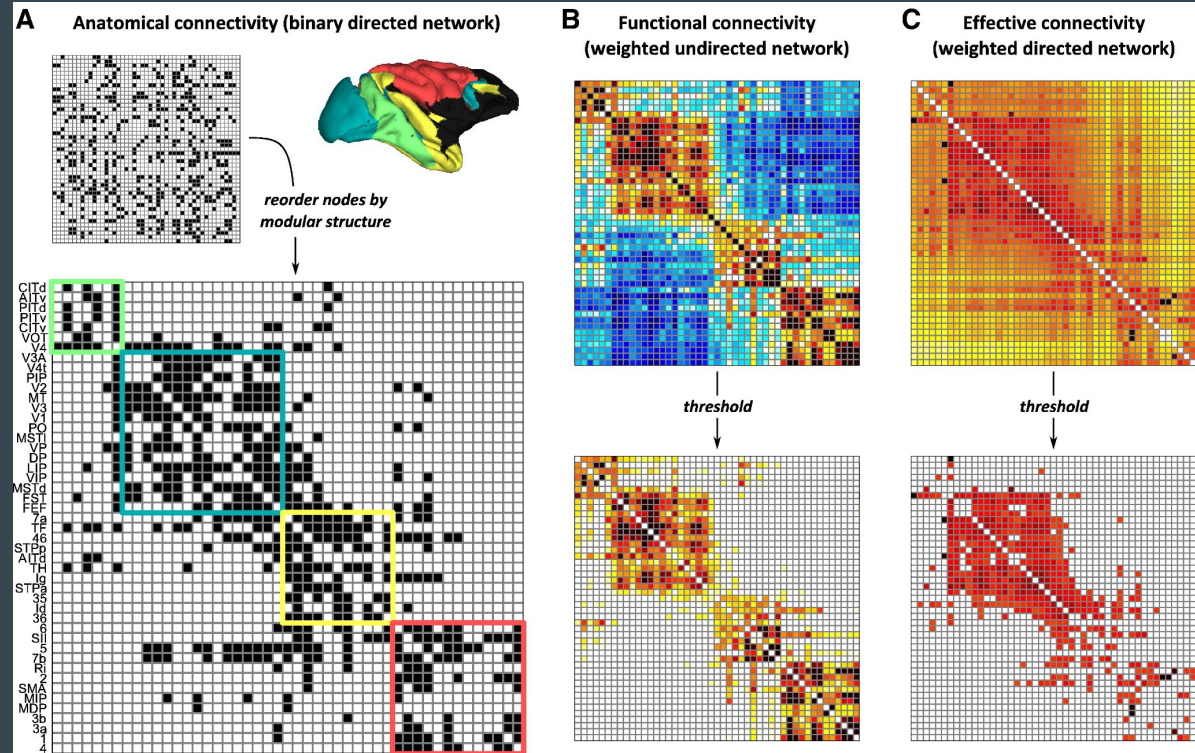
Complex network measures

Construction of brain networks from large scale anatomical and functional connectivity datasets. Structural networks are commonly extracted from histological (tract tracing) or neuroimaging (diffusion MRI) data. Functional networks are commonly extracted from neuroimaging (fMRI) or neurophysiological (EEG, MEG) data. For computational convenience, networks are commonly represented by their connectivity matrices, with rows and columns representing nodes and matrix entries representing links. To simplify analysis, networks are often reduced to a sparse binary undirected form, through thresholding, binarizing, and symmetrizing.

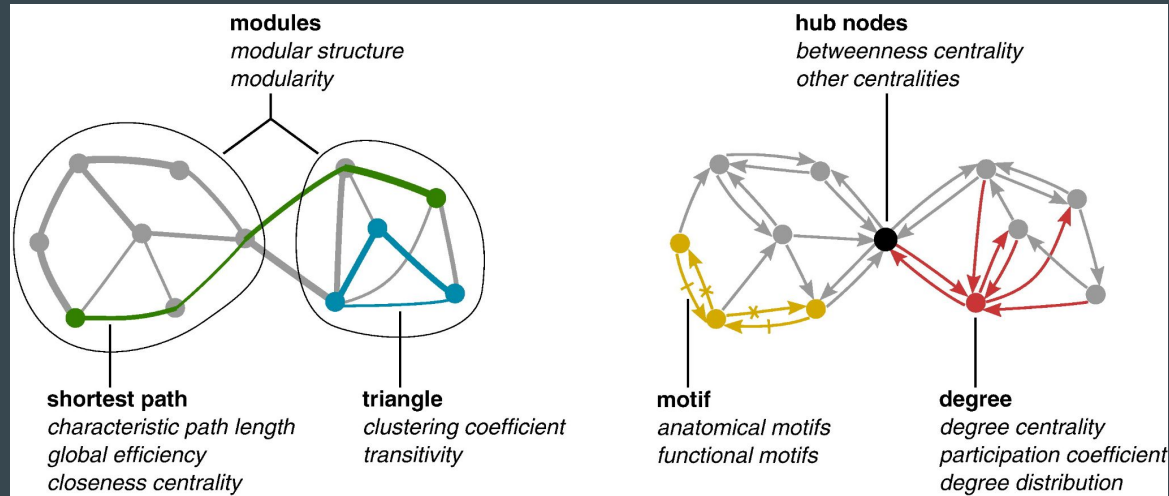


Complex network measures

Anatomical, functional, and effective connectivity networks. Large-scale anatomical connection network of the macaque cortex, including the ventral and dorsal streams of visual cortex, as well as groups of somatosensory and somatomotor regions.

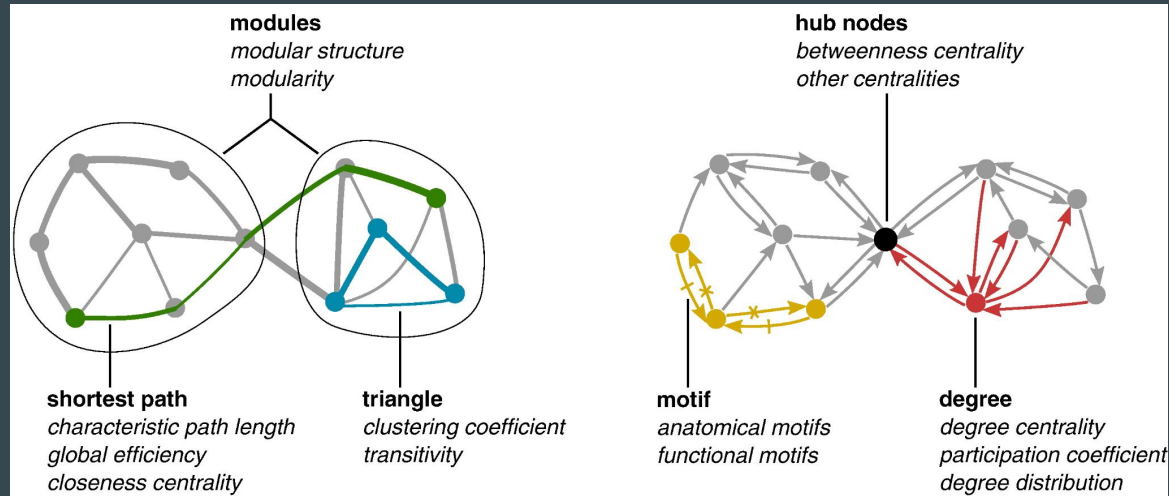


Complex network measures



Key complex network measures. Measures of integration are based on shortest path lengths (green), while measures of segregation are often based on triangle counts (blue) but also include more sophisticated decomposition into modules (ovals). Measures of centrality may be based on node degree (red) or on the length and number of shortest paths between nodes. Hub nodes (black) often lie on a high number of shortest paths and consequently often have high betweenness centrality. Patterns of local connectivity are quantified by network motifs (yellow).

Complex network measures



Key complex network measures. Measures of integration are based on shortest path lengths (green), while measures of segregation are often based on triangle counts (blue) but also include more sophisticated decomposition into modules (ovals). Measures of centrality may be based on node degree (red) or on the length and number of shortest paths between nodes. Hub nodes (black) often lie on a high number of shortest paths and consequently often have high betweenness centrality. Patterns of local connectivity are quantified by network motifs (yellow).

Computational synthesis of cortical dendritic morphologies

Kanari et al. 2022

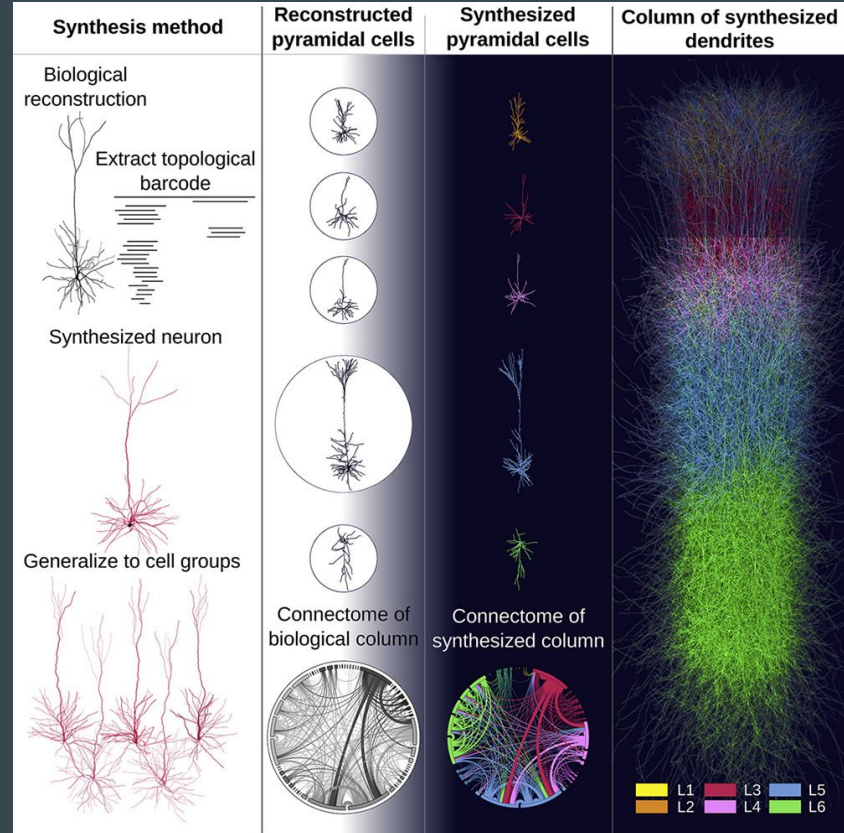
Computational synthesis of cortical dendritic morphologies

Neuronal morphologies provide the foundation for the electrical behavior of neurons, the connectomes they form, and the dynamical properties of the brain. Comprehensive neuron models are essential for defining cell types, discerning their functional roles, and investigating brain-disease-related dendritic alterations. However, a lack of understanding of the principles underlying neuron morphologies has hindered attempts to computationally synthesize morphologies for decades. We introduce a synthesis algorithm based on a topological descriptor of neurons, which enables the rapid digital reconstruction of entire brain regions from few reference cells. This topology-guided synthesis generates dendrites that are statistically similar to biological reconstructions in terms of morpho-electrical and connectivity properties and offers a significant opportunity to investigate the links between neuronal morphology and brain function across different spatiotemporal scales. Synthesized cortical networks based on structurally altered dendrites associated with diverse brain pathologies revealed principles linking branching properties to the structure of large-scale networks.

Computational synthesis of cortical dendritic morphologies

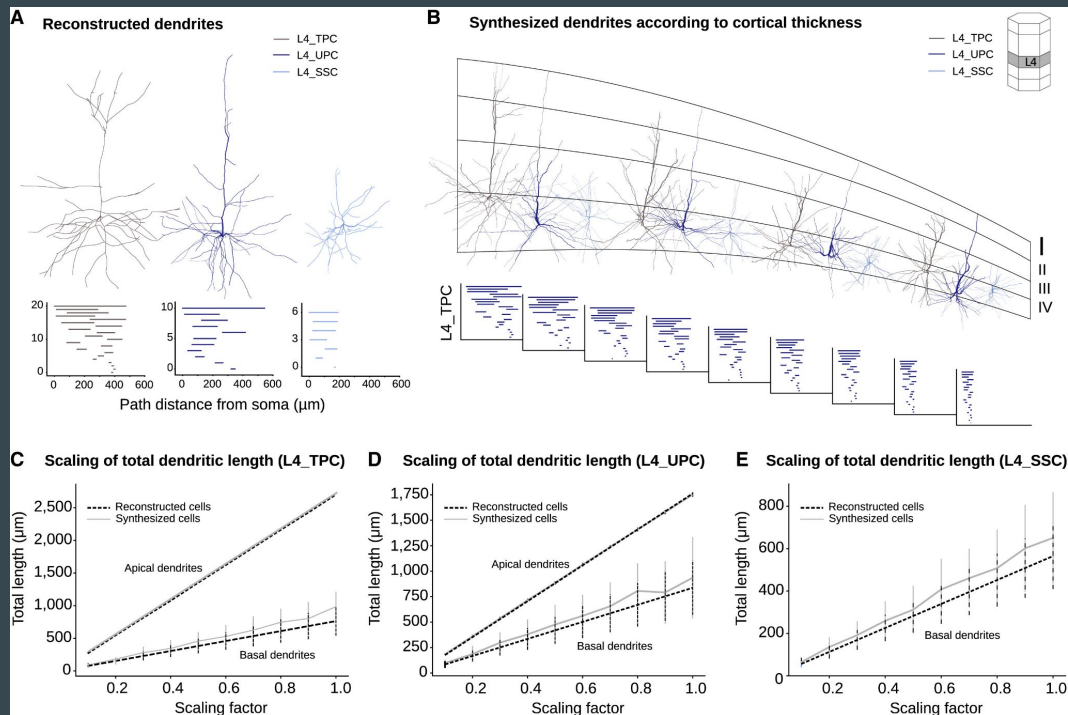
Highlights

- 1) Topological synthesis generates healthy and diseased cortical dendrites
- 2) Synthesized dendrites are indistinguishable from biological reconstructions
- 3) Topological model enables the investigation of the functional roles of cell types
- 4) Links of branching properties and structure of large-scale networks are revealed



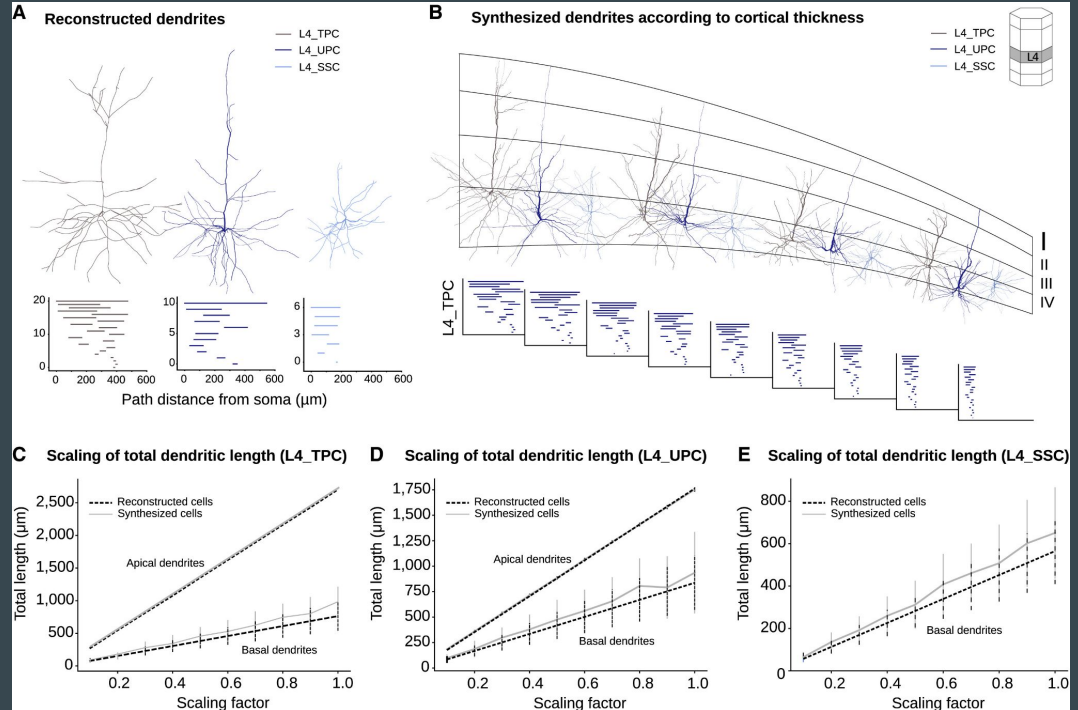
Computational synthesis of cortical dendritic morphologies

Generalization of topological synthesis for varying cortical thickness. (A) Exemplar biological reconstructions of three layer 4 pyramidal cell types: L4_TPC (gray), L4_UPC (deep blue), L4_SSC (light blue), and the corresponding persistence barcodes, used as synthesis input. (B) Scaling of input persistence barcodes and resulting synthesized dendrites ([1.0, 0.8, 0.6, 0.5] of original barcodes). The scaled (from 1.0 to 0.2) barcodes of synthesized L4_TPC apicals presented at the bottom.



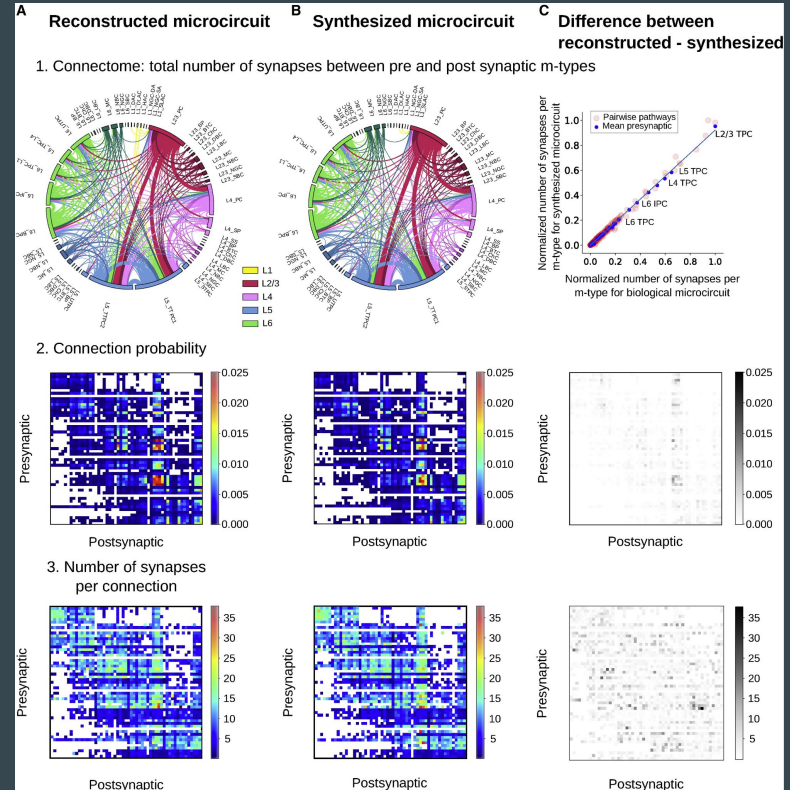
Computational synthesis of cortical dendritic morphologies

(C–E) Total dendritic length of layer 4 cells, as a function of shrinkage factor for basal (bottom) and apical (top) dendrites compared with expected values of scaled biological lengths (black dashed, computed as scaling factor multiplied by total length of reconstructed dendrites) and synthesized (gray continuous) dendrites of L4_TPC (C), L4_UPC (D), and L4_SSC (E). Note that L4_SSC do not have apical dendrites even though they are excitatory cells, therefore only basal dendrite statistics are shown.



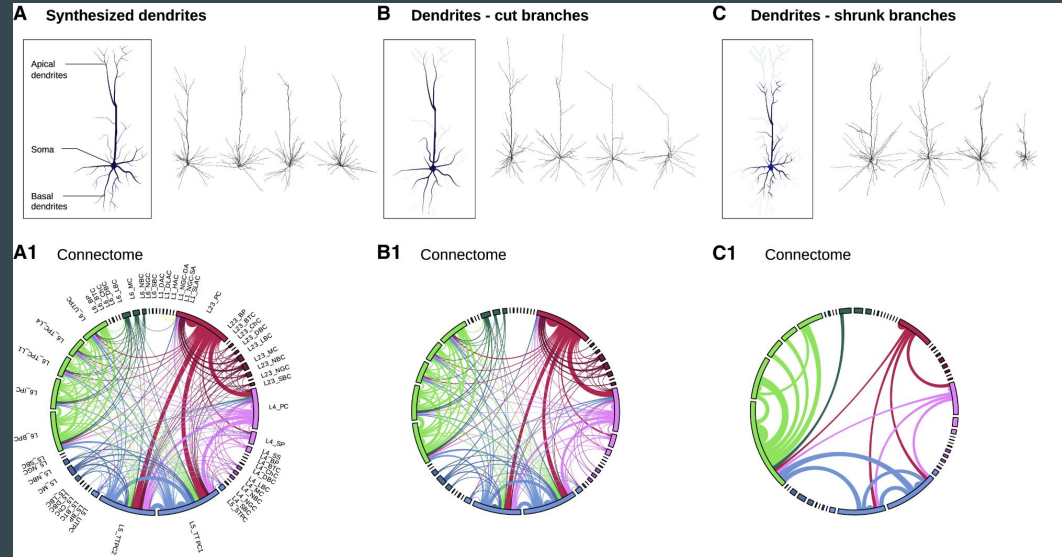
Computational synthesis of cortical dendritic morphologies

Connectivity of synthesized and reconstructed networks. (A) The connectivity properties of a reconstructed microcircuit (Markram et al., 2015). (B) The connectivity properties of a microcircuit of fully synthesized dendrites, and reconstructed axons. (C) Difference between reconstructed and synthesized microcircuits. (1) The connectomes of the microcircuits grouped by m-type. (2) Connection probability. (3) Synapses per connection.



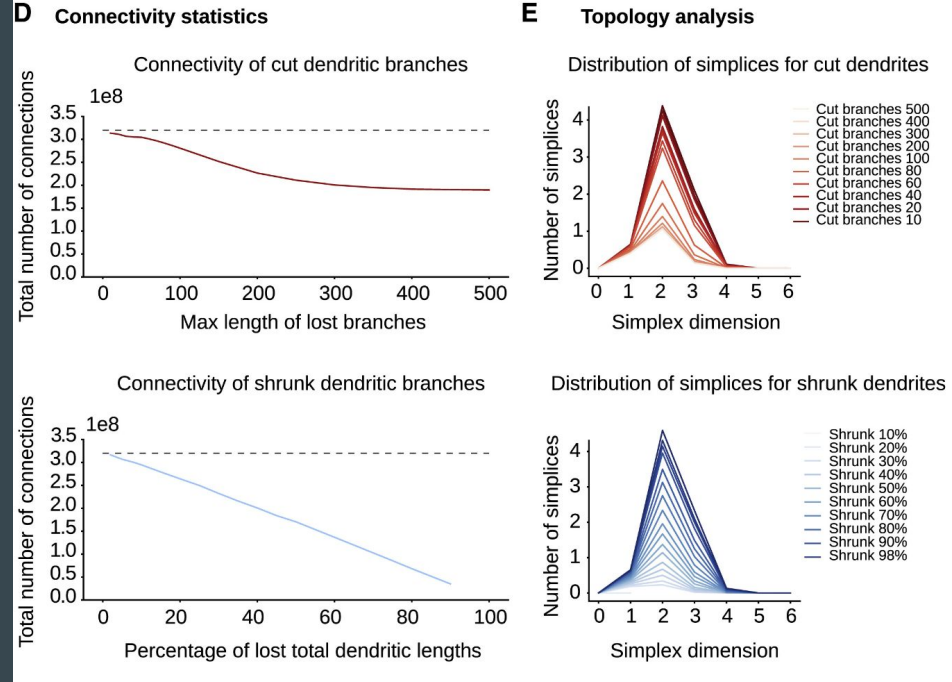
Computational synthesis of cortical dendritic morphologies

Medical applications. (A–C) Connectivity of synthesized networks based on structural alterations of dendritic morphologies. Schematic representation and examples of layer 5 synthesized pyramidal cells (A), in comparison with cut dendritic branches (B) (lengths above 10, 100, 200, and 400 μm), and shrunk dendrites (C) (98%, 90%, 60%, and 30%). Connectome (presented in subpanel 1) of each synthesized microcircuit: (A) synthesized, (B) cut branches of lengths above 400 μm , (C) shrunk dendrites 10%.



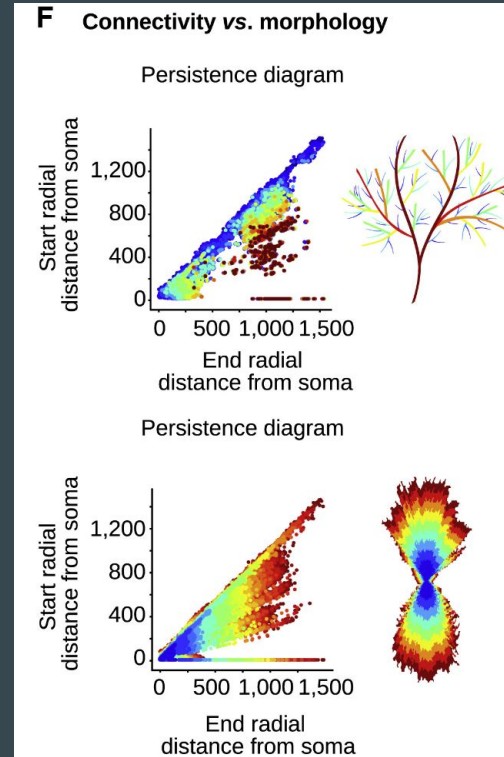
Computational synthesis of cortical dendritic morphologies

(D) Total number of connections for alterations of type B (red) and C (blue) compared with synthesized network A (black). (E) Topological analysis of corresponding networks; distribution of directed simplices for alterations of type B (red, top) and C (blue, bottom).



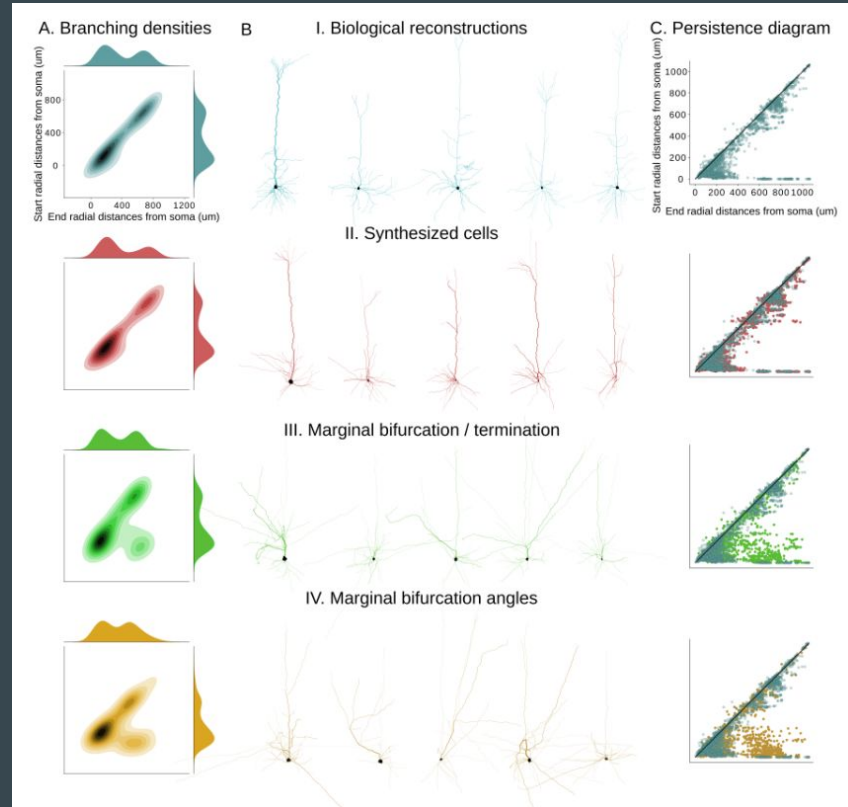
Computational synthesis of cortical dendritic morphologies

(F) Morphological characteristics and connectivity with respect to alterations of type B (top) and C (bottom). The main branches form the majority of connections (top) and larger dendritic extents (bottom) form more connections. Colormap corresponds to normalized number of connections: from maximum number of connections (3.5×10^8 in red) to minimum (10^7 in blue).



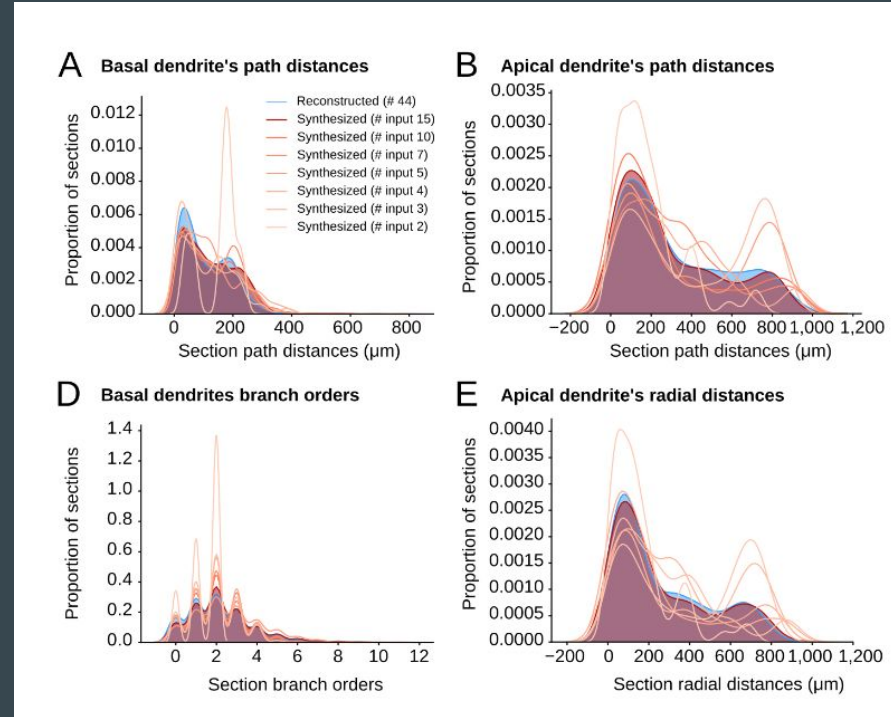
Computational synthesis of cortical dendritic morphologies

Comparison of synthesis methods.
Comparison of synthesized cells for different synthesis methods. A. Density and marginal projections of persistence diagrams for reconstructed cells (I), synthesized cells (II), synthesized without correlation of branching / termination (III), and synthesized without correlation between branching and bifurcation angles (IV). B. Examples for the same data. C. Respective persistence diagrams.



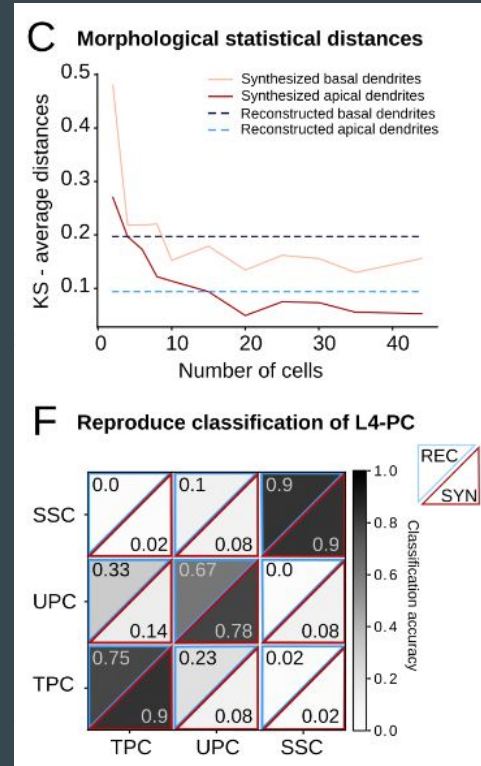
Computational synthesis of cortical dendritic morphologies

Morphological diversity. Comparison of dendrites from 44 reconstructed L4 TPC cells (in blue) to synthesized dendrites (based on subsets of increasing numbers of cells from the original population used as inputs: from 2 to 15, red shades from lighter to darker). Comparison of path distance (A, direct input) and branch order (B, emergent property) for basal dendrites. Comparison of path distances (D, direct input) and radial distance (E, emergent property) for apical dendrites. The original distributions are well approximated by a subset of input cells (15 out of 44).



Computational synthesis of cortical dendritic morphologies

C. Average statistical (Kolmogorov-Smirnov) distance for numerous morphometrics, within reconstructed cells (in blue) and between reconstructed and synthesized cells (in red) as a function of increasing synthesis inputs. F. TMD based classification of three L4 PC types for reconstructed (top left, blue) and synthesized (bottom right, red) cells. Classification accuracy is same or higher for the synthesized population.



Rich cell-type-specific network topology in neocortical microcircuitry

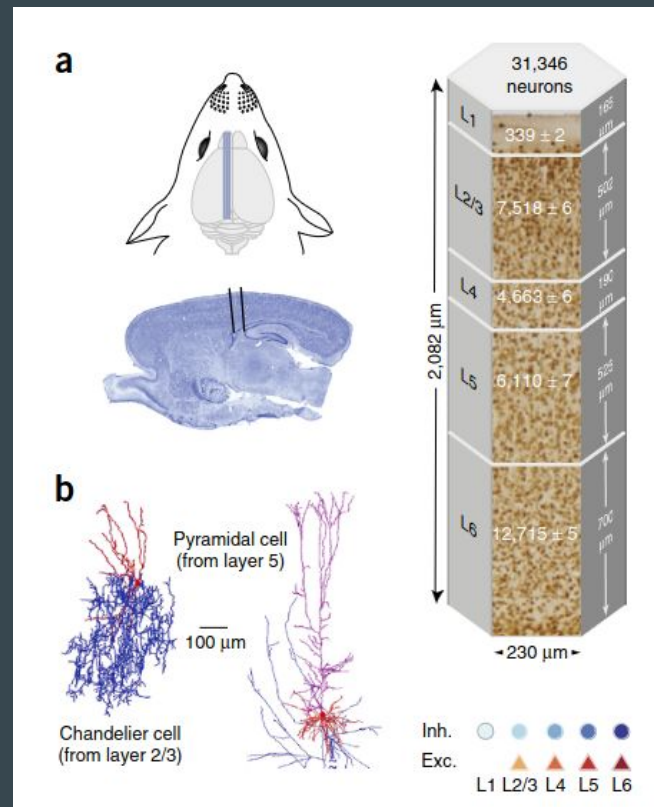
Gal et al. 2017

Rich cell-type-specific network topology in neocortical microcircuitry

Neuronal morphologies provide the foundation for the electrical behavior of neurons, the connectomes they form, and the dynamical properties of the brain. Comprehensive neuron models are essential for defining cell types, discerning their functional roles, and investigating brain-disease-related dendritic alterations. However, a lack of understanding of the principles underlying neuron morphologies has hindered attempts to computationally synthesize morphologies for decades. We introduce a synthesis algorithm based on a topological descriptor of neurons, which enables the rapid digital reconstruction of entire brain regions from few reference cells. This topology-guided synthesis generates dendrites that are statistically similar to biological reconstructions in terms of morpho-electrical and connectivity properties and offers a significant opportunity to investigate the links between neuronal morphology and brain function across different spatiotemporal scales. Synthesized cortical networks based on structurally altered dendrites associated with diverse brain pathologies revealed principles linking branching properties to the structure of large-scale networks.

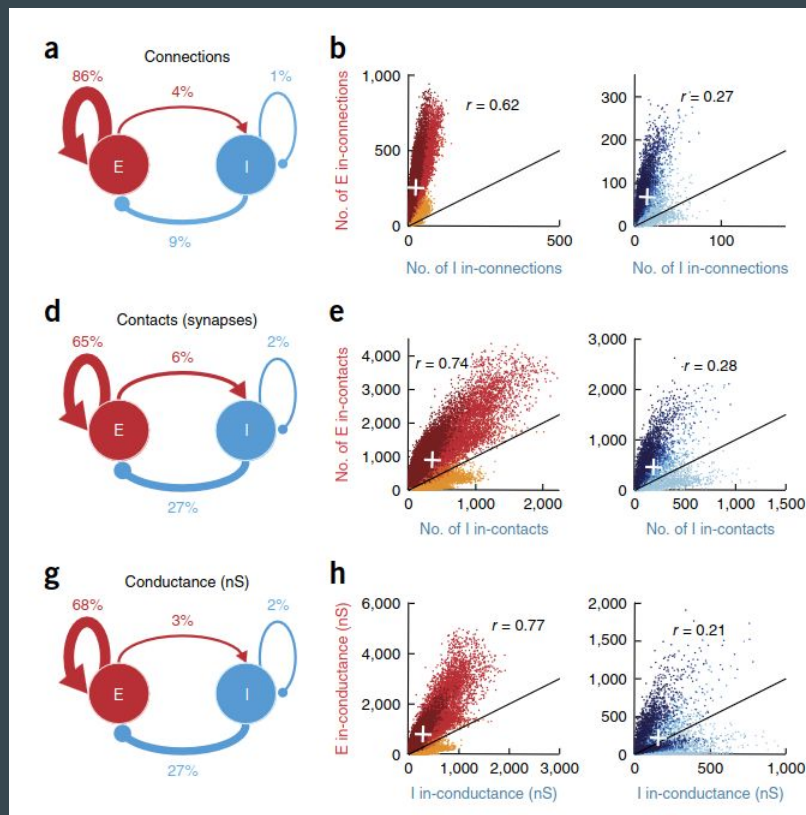
Rich cell-type-specific network topology in neocortical microcircuitry

In silico model of neocortical microcircuitry (NMC) (a) Left: blue stripe in rat cartoon indicates the sagittal plane of the neocortex used to obtain somatosensory slices; black lines in brain image indicate location of the modeled NMC. Right: dimensions and number of cells per layer in the seven instances of the model (mean $\pm$ s.d., N = 7). The total number of neurons in this circuit is depicted at the top. (b) Example of reconstructions of two morphological cell types (layer (L) 2/3 ChC (chandelier) inhibitory interneuron, left, and L5 TTPC2 (thick-tufted pyramidal) neuron, right) out of the 55 neuron types used in this study.



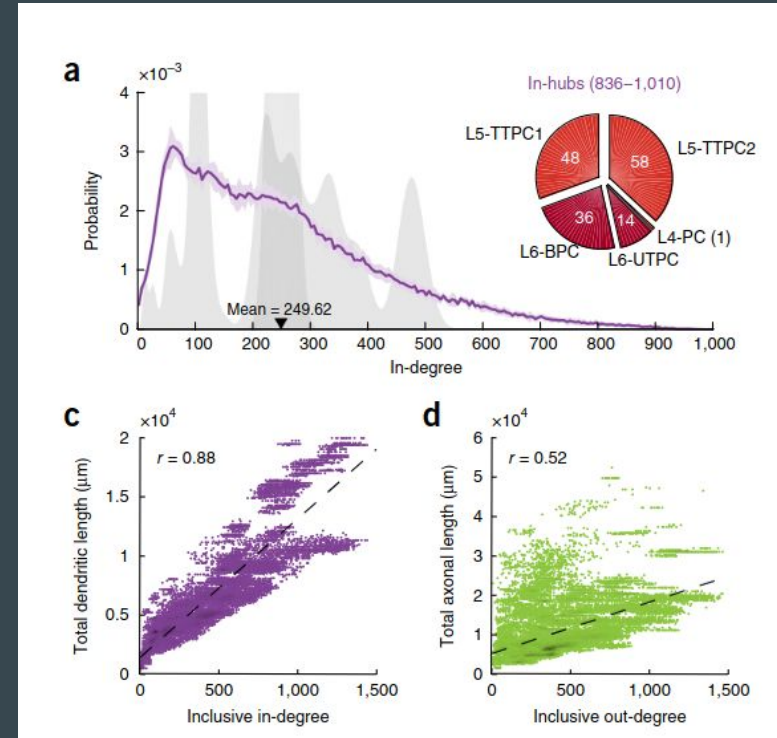
Rich cell-type-specific network topology in neocortical microcircuitry

Cellular-level E/I balance is emphasized when considering synaptic conductance. (a) The percentage of E (arrow head) and I (circular head) pathways in terms of the total number of connections ($7,824,436 \pm 104,092$). Line thickness illustrates the percentage of connections in the corresponding pathway. (b) Incoming E/I connections for excitatory (left, $n_E = 26,567$) and inhibitory (right, $n_I = 4,779$) cells. (d–i) Total number of synaptic contacts (synapses: $36,471,080 \pm 554,503$) and (g–h) total peak synaptic conductance (30.74 ± 0.47 mS). Note the excess excitation vs. inhibition in all three cases.



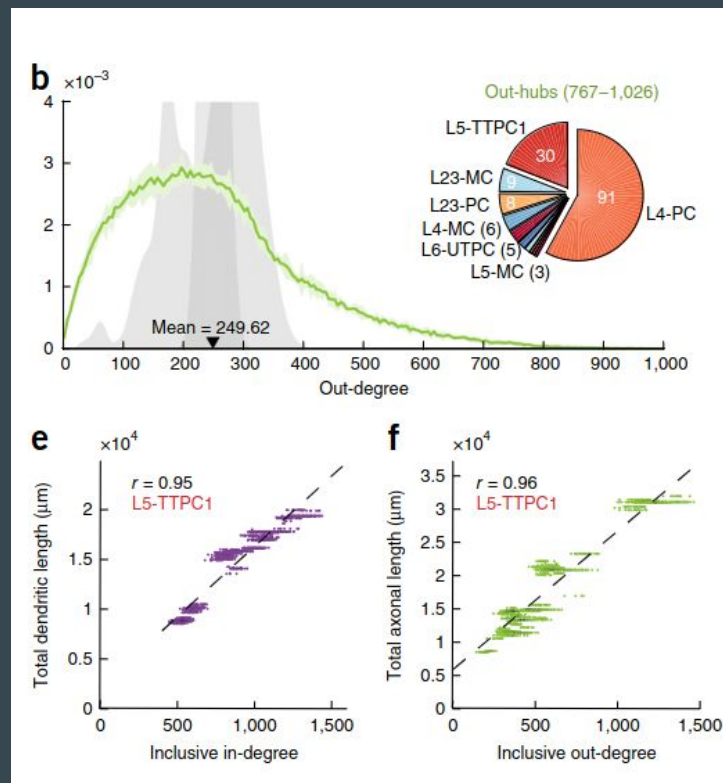
Rich cell-type-specific network topology in neocortical microcircuitry

In silico model of neocortical microcircuitry (NMC)
(a) Left: blue stripe in rat cartoon indicates the sagittal plane. In-hub and out-hub neurons belonging to a small subset of cell types and forming a rich club. (a) Long-tail distributions of in-degrees (number of presynaptic cells per neuron) in the seven NMC instances. The top 157 (0.5%) in-degrees (in-hubs) arise primarily from only four cell types residing in deep layers (pie chart). (c) Within the central NMC, a neuron's inclusive in-degree (when taking the extrinsic connections from surrounding NMCs into account) is correlated with its total dendritic length. (d) A neuron's inclusive out-degree is less correlated with its total axonal length. P values of Pearson correlations coefficients, r , are both < 0.001 ($N = 31,346$ neurons).

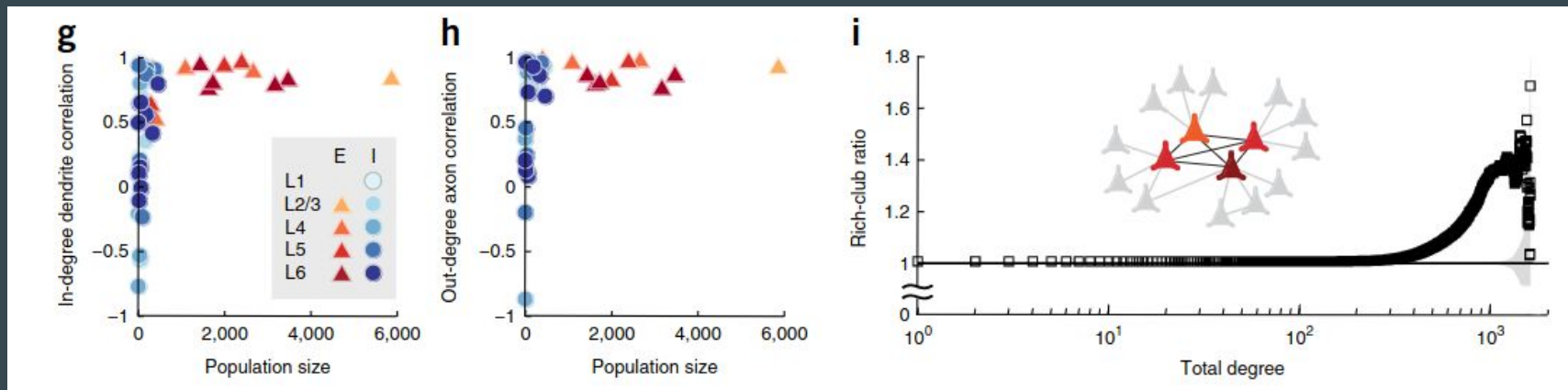


Rich cell-type-specific network topology in neocortical microcircuitry

In silico model of neocortical microcircuitry (NMC) (b) Long-tail distribution of out-degrees (number of postsynaptic cells per neuron); the top 157 out-degrees (out-hubs) arise from multiple cell types, the majority of which are pyramidal and Martinotti cells from intermediate layers (pie chart). (e,f) As in c and d but for thick-tufted pyramidal neurons from layer 5 (L5-TTPC1) only (N = 2,403).



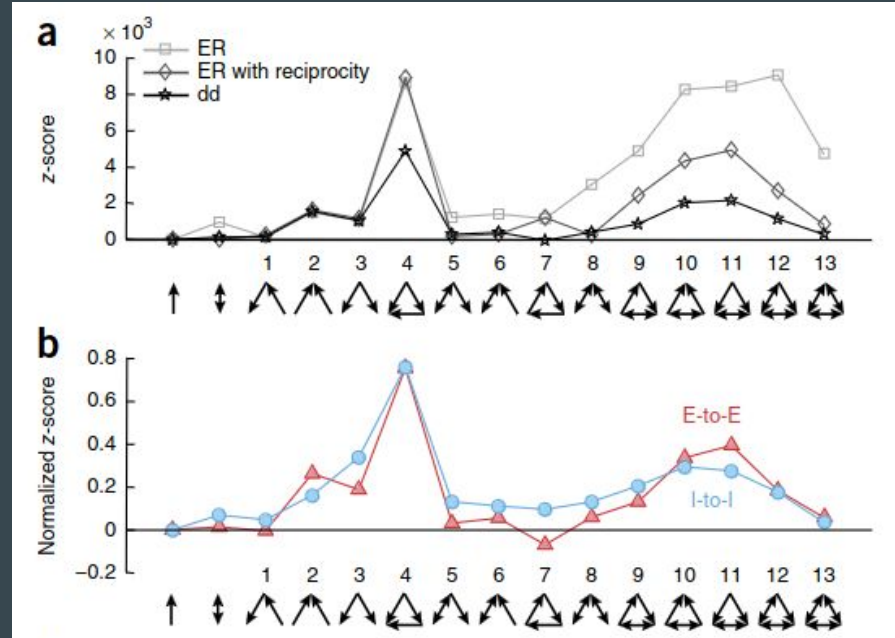
Rich cell-type-specific network topology in neocortical microcircuitry



(g) Correlations between inclusive in-degrees and total dendritic lengths for all 55 cell types. (h) Correlations between inclusive out-degrees and total axonal lengths for all 55 cell types. (i) Ratios of the number of connections among NMC neurons whose total degree (in-degree + out-degree) $> d$ to the number of connections expected from random networks with matching degree sequence. This ratio is > 1 for high-degree neurons ($P < 0.001$, Monte Carlo), reflecting the presence of a rich club of the neocortex used to obtain somatosensory slices.

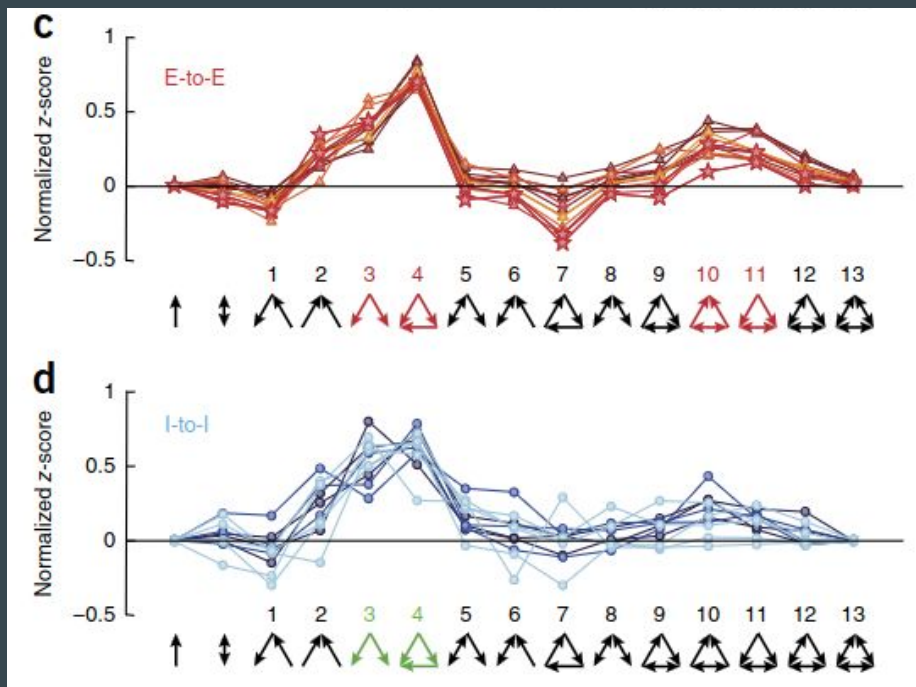
Rich cell-type-specific network topology in neocortical microcircuitry

Local wiring-specificities within the NMC show overrepresented three-neuron network motifs. (a) Significance level (z-score) for all 13 triads for all cells in the NMC with respect to three types of random networks (Monte Carlo with $N = 100$). ER random networks with matching numbers of overall connections, light gray with squares; ER with additional matching of reciprocal connections, dark gray with diamonds; dd-matched random networks, black with stars. (b) Normalized z-scores for all 13 triads in both excitatory (black) and inhibitory (gray) subnetworks with respect to the dd-matched random networks (Monte Carlo with $N = 100$).



Rich cell-type-specific network topology in neocortical microcircuitry

(c,d) As in b but for cell-type-specific subnetworks (c) of all 13 excitatory neurons (each layer is depicted by a different color) and for (d) the 11 largest inhibitory cell type populations (total number of cells > 150 cells per type). Color codes correspond to specific layers, as in Figure 4. Red motifs in c were found to be overrepresented in experiments performed on L5–L5 thick-tufted pyramidal cells by Perin et al.16, and green motifs in d were found in cerebellar inhibitory cells by Rieubland et al.31. Both of these motifs were found in this study to be common to all excitatory and inhibitory cell types. Monte Carlo with $N = 1,000$ was used in c and d.



Questions?