Network and Multilayer Network Approaches to Understanding Human Brain Dynamics

Sarah Feldt Muldoon and Danielle S. Bassett*

Network neuroscience provides a systems approach to the study of the brain and enables the examination of interactions measured at different temporal and spatial scales. We review current methods to quantify the structure of brain networks and compare that structure across different clinical cohorts, cognitive states, and subjects. We further introduce the emerging mathematical concept of multilayer networks and describe the advantages of this approach to model changing brain dynamics over time. We conclude by offering several concrete examples of how multilayer network approaches to neuroimaging data provide novel insights into brain structure and evolving function.

Introduction. The human brain is a system of interacting and evolving networks arranged across multiple scales in both space and time (Bassett and Gazzaniga 2011; Feldt, Bonifazi, and Cossart 2011; Bassett and Siebenhühner 2013). At the small scale, individual neurons are linked to one another via synapses, forming a structural web on which information can flow in the

*To contact the authors, please write to: Sarah Muldoon, Department of Mathematics and Computational and Data-Enabled Science and Engineering Program, University at Buffalo, SUNY, Buffalo, NY 14260; e-mail: smuldoon@buffalo.edu. Danielle Bassett, Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104; e-mail: dsb@seas.upenn.edu.

†This work was supported by the John D. and Catherine T. MacArthur Foundation, the Alfred P. Sloan Foundation, the Army Research Laboratory and the Army Research Office through contract numbers W911NF-10-2-0022 and W911NF-14-1-0679, the National Institute of Mental Health (2-R01-DC-009209-11), the National Institute of Child Health and Human Development (1R01HD086888-01), the Office of Naval Research, and the National Science Foundation (BCS-1441502 and BCS-1430087). The content is solely the responsibility of the authors and does not necessarily represent the official views of any of the funding agencies.

Philosophy of Science, 83 (December 2016) pp. 710–720. 0031-8248/2016/8305-0006$10.00
Copyright 2016 by the Philosophy of Science Association. All rights reserved.
form of electrical impulses referred to as action potentials. The activity of individual neurons combines to produce complex oscillatory signals in large-scale brain regions connected by bundles of axons referred to as white matter tracts. Neurophysiological correlates of these oscillations can be recorded using electroencephalography (EEG) or functional magnetic resonance imaging (fMRI), while neurophysiological correlates of white matter can be measured using diffusion imaging. To complement this dynamic information, neurophysiological correlates of white matter tracts can be measured using diffusion tensor or diffusion spectrum imaging (DTI/DSI).

In essence, the brain’s concerted activity and structural linkages form networks (Bassett and Bullmore 2006). Specifically, structural networks can be constructed from anatomical connections that physically link neurons or brain regions, and functional networks can be constructed from functional connections that quantify statistical similarities in activity (Bullmore and Sporns 2009; Bullmore and Bassett 2011). We refer to these two different but equally important types of networks as structural networks and functional networks, respectively. The organizational principles of structural and functional networks are likely deeply intertwined (Sporns 2013), are modulated by cognitive state (Sporns 2014; Medaglia, Lynall, and Bassett 2015), and can be altered in disease and injury (Bassett and Bullmore 2009; Stam 2014).

Understanding the organization of brain networks is arguably necessary to understand cognition. Here, we turn to network theory, an emerging multidisciplinary field that draws on tools from mathematics, physics, and computer science to describe complex interconnection patterns (Newman 2010). Network theory has been successfully applied in a multitude of biological contexts—from anthropology to gene regulation—by mapping systems into graphs. Network neuroscience maps brain systems into graphs in which neurons or brain regions are represented as nodes, and structural or functional interactions are represented as edges. Although this approach can be applied across multiple spatial scales, most recent studies focus on networks constructed at the meso- to macroscale, where noninvasive imaging data exist in humans.

In this article, we describe the basic brain network modeling construct and present several network statistics used to quantify brain network structure (Bassett and Lynall 2013). We summarize fundamental findings that shape our current understanding of brain function, using static snapshots of brain network organization. In light of the vast quantities of dynamic data now available, we further describe the relatively new mathematical formalism of multilayer networks (Kivelä et al. 2014) that enables quantitative characterization and prediction of network evolution over time. These techniques form the backbone of the emerging field of dynamic network neuroscience, which opens up entirely new classes of investigation in human neuroscience.
Building Brain Networks. To model a system as a graph, one must define the elements of the system that will be represented by nodes and the relationships between those elements that will be represented by edges (connections) between nodes. These definitions will be informed by the system being interrogated, the observable data, and the scientific question of interest and will lead to a given set of networks characterized by a specific structure. Thus, different sets of choices can provide complementary information about brain organization and function.

Node Selection. The choice of nodes often depends on the data type. For example, if one has EEG data, network nodes may be surface sensors or hypothesized neural dipoles estimated from source localization techniques. For fMRI data, network nodes may be (i) single voxels or (ii) some aggregate of voxels chosen to represent anatomical (Achard et al. 2006) or functional (Power et al. 2011) areas. In some cases, these atlases can be defined at different spatial scales (Cammoun et al. 2012), where each scale represents a further partitioning of a region into subregions (Zalesky et al. 2010; Bassett, Brown, et al. 2011). Thus, when defining network nodes on the basis of a parcellation, one can also study how the choice of parcellation (Wang et al. 2009; Bassett, Brown, et al. 2011) or parcellation scale (Zalesky et al. 2010; Bassett, Brown, et al. 2011) affects network properties.

Edge Selection. The choice of edge also often depends on the data type. For structural brain networks, one might first perform diffusion tractography to trace white matter tracts through the brain (Sporns 2011). A network edge can then be defined as a weighted connection whose strength is given by the number or average fractional anisotropy of tracts connecting two nodes (Hermundstad et al. 2013). For functional brain networks, one might define edges on the basis of statistical relationships between time series of regional activity, such as interregional correlation, coherence, or phase synchronization that produce undirected networks (A is connected to B with the same strength that B is connected to A). Alternatives include Granger causality that produces directed networks (connection from A to B can have a different strength than the connection from B to A).

Weighted versus Binary Networks. The methods described above produce weighted networks where the value associated with an edge represents the strength of the relationship between nodes. However, concerns regarding measurement noise have traditionally led to the application of thresholding techniques to create binary networks (Achard et al. 2006), where connections between nodes either exist or do not exist. The most common thresholding technique is cumulative thresholding, where only connections whose strength is above a threshold value are retained (Bassett et al. 2012). How-
ever, weak connections have been implicated in network function and dysfunction (Bassett et al. 2012; Santarnecchi et al. 2014), and therefore caution should be exercised when thresholding techniques are employed (Ginestet et al. 2011). Future work in the field should emphasize the importance of analyses and methods that incorporate the strength of connections (Rubinov and Sporns 2011).

Spatial Embedding. Finally, it is important to note that brain networks are an example of a class known as spatially embedded networks (Klimm et al. 2014): nodes have physical locations, and internode distances can be calculated in Euclidean space. In the brain, nodes that are physically close tend to be strongly connected (Kaiser and Hilgetag 2006; Bassett et al. 2010), leading to an efficient info-physical structure (Kaiser and Hilgetag 2006; Bassett et al. 2010; Bullmore and Sporns 2012). It can be useful to take this information into account when designing null models used to compute network statistics (Expert et al. 2011).

Quantifying Brain Networks. A brain network can be characterized using network statistics (Newman 2010). Here we describe some commonly used statistics and the insights that they have offered into the complexity of human brain organization (Rubinov and Sporns 2011).

Network Statistics. We assume that we have an $N$-node graph, $G$, represented by an $N \times N$ adjacency matrix $A$, where the element $A_{ij}$ gives the strength of the connection between node $i$ and node $j$.

Degree. The degree of a node can be computed as

$$k_i = \sum_j A_{ij}. \quad (1)$$

Nodes identified as having an unusually high degree when compared to other nodes in the network are often referred to as “hubs.” The distribution of node degrees can be an important distinguishing feature between different networks: random networks display a Poisson distribution (Newman 2010), whereas brain networks display a heavy-tailed distribution (Achard et al. 2006), indicating the presence of hubs (van den Heuvel and Sporns 2013).

Path Length. To measure the distance between two nodes in topological space, we compute the shortest path length, $d_{ij}$, as the minimal number of edges that must be traversed to walk from node $i$ to node $j$ along edges in the graph. Note that the shortest path length computed in network space is mathematically unrelated to the physical distance between two nodes. We can define the characteristic path length,
as the average shortest path length.

**Clustering Coefficient.** The local clustering coefficient of a node, \( c_i \), can be defined as the fraction of a node’s neighbors that are also connected to one another. For a binary network, we define the global clustering coefficient to be the mean local clustering coefficient values over all \( N \) nodes in the graph:

\[
C = \frac{1}{N} \sum_{i} c_i.
\]  

For weighted networks, one can define related diagnostics that account for the local clustering of edge weights.

**Efficiency.** The global efficiency of a network is defined as

\[
E_{\text{glob}} = \frac{1}{N(N-1)} \sum_{\langle ij \rangle} \frac{1}{d_{ij}}.
\]  

Because this formulation uses information regarding the shortest paths through the network (Latora and Marchiori 2001), efficiency is often interpreted as a measure of how efficiently the network transmits information (although, for caveats on this interpretation in the context of brain networks, see Rubinov and Bassett [2011]).

The local efficiency of a node in a network is given by the average global efficiency of that node’s local subgraph, and therefore the local efficiency of the entire network is given by

\[
E_{\text{loc}} = \frac{1}{N} \sum_{i} E_{\text{glob}}(G_i),
\]

where \( G_i \) is the local subgraph of node \( i \) consisting of the nearest neighbors of node \( i \) and their edges to other nodes in the subgraph.

**Modularity.** An interesting feature of many complex networks is their community structure: the presence of subsets of nodes that are more densely connected among themselves than they are to the rest of the network (Porter, Onnela, and Mucha 2009). A popular method to detect communities in networks is to maximize the following modularity quality function (Newman 2006):

\[
Q = \frac{1}{2m} \sum_{ij} \left( A_{ij} - \frac{k_i k_j}{2m} \right) \delta_{c_i c_j},
\]
where \( m = (1/2m) \sum_{ij} A_{ij} \), \( \gamma \) is a structural resolution parameter (often set to unity), \( \delta_{C_i,C_j} = 1 \) if nodes \( i \) and \( j \) are in the same community, and \( \delta_{C_i,C_j} = 0 \) otherwise. The modularity is computed over the landscape of all possible partitions of the network into communities and is maximal for the partition with the greatest community structure as quantified by the modularity quality index \( Q \). Unfortunately, maximizing \( Q \) is an NP-complete problem, and therefore heuristic algorithms are used such as the locally greedy Louvain algorithm (Blondel et al. 2008).

**Brain Organization.** Network statistics form a toolbox to characterize the organization of human brain networks, offering insight into human cognition (Medaglia et al. 2015). Here, we review three basic features observed consistently in human brain networks: small-world structure, community structure, and rich-club structure.

**Small-World Structure.** The popular Watts-Strogatz model of small-world networks contains high local clustering yet short average path length (Watts and Strogatz 1998) and therefore forms an attractive model for brain organization (Bassett and Bullmore 2006); the locally dense interconnection structure may promote computation, while the long-distance paths may promote the quick dissemination of information (although see Rubinov and Bassett 2011; Bassett and Lynall 2013; Medaglia et al. 2015, for important interpretational caveats). This hypothesized small-world structure is evident in different cohorts, imaging modalities, and spatial resolutions (Bassett and Bullmore 2006; Bullmore and Sporns 2009; Stam 2014).

**Community Structure.** Decades of cognitive neuroscience have demonstrated that different sets of brain regions are important for the performance of different cognitive tasks (Gazzaniga, Ivry, and Mangun 2013). Recent applications of network-based tools to human neuroimaging data have shown that these sets of areas are also strongly connected in brain networks. By applying community detection techniques to neuroimaging data, one can uncover community architecture in structural (Chen et al. 2008; Bassett et al. 2010) and functional networks (Meunier et al. 2009), where communities or modules map onto known cognitive systems. Interestingly, this modular organization tends to be hierarchical—smaller modules exist within larger modules (Meunier et al. 2010; Bassett et al. 2010, 2013; Bassett, Wymbs, et al. 2011)—potentially enabling local computations to be integrated into broader cognitive processes.

**Rich-Club Structure.** Human brain networks also contain hubs that are more likely to connect to each other than they are to regions of low degree, forming what has been called a “rich club” (Colizza et al. 2006). This club
of densely connected hub regions might contribute to the robustness of brain function (van den Heuvel and Sporns 2011) and form a fundamental information-processing feature of neural systems across species (Shih et al. 2015).

**The Future of Network Neuroscience.** Network approaches have fundamentally contributed to our understanding of brain function. These tools are particularly suited to examining single network instances: one subject, one species, one time point, one disease, one task. Yet, they fall short in offering statistically robust methods to translate inferences across networks. In essence, these tools treat each network as an independent sample when in fact there are significant dependencies between networks constructed from different subjects, different species, different time points, different diseases, or different tasks. A principled modeling framework to accurately treat and manipulate these dependencies is critically needed.

**Multilayer Networks.** To address this limitation, we turn to the new mathematical formalism of multilayer networks (Kivelä et al. 2014): representations in which each layer constitutes a different network, for example, constructed from a different subject, patient group, experimental condition, cognitive task, or time point. One can also combine multiple aspects, thus creating a network with more than one multilayer dimension: for example, a multilayer network in which one dimension is subject and a second dimension is time. A node can exist in all layers or in a subset of layers and is linked throughout the layers by an edge representing the node’s identity. Multiple types of edges can also link nodes within and between layers representing different types of relationships between network elements.

The multilayer framework is particularly appealing in the context of understanding statistically significant similarities and differences in network structure between experimental scenarios or across time, a traditionally difficult enterprise (Muldoon, Bridgeford, and Bassett 2016). For example, a perennial question lies in how brain structure constrains brain function (Sporns 2013). To address this question, one might acquire a multimodal data set of structural and functional images and wish to build a multilayer network with two aspects: one layer containing the structural networks from a group of subjects and the other aspect containing the functional networks from a group of subjects. Edges representing nodal identity then link nodes between the two aspects and many layers to ensure that a given region can be identified consistently. A second interesting scenario in the context of disease (Bassett and Bullmore 2009; Stam 2014) lies in building a multilayer network in which one aspect contains brain networks from healthy individuals and a second aspect contains brain networks from a patient cohort. Finally in the context of cognitive neuroscience (Sporns 2014; Medaglia et al. 2015),
one might wish to construct a multilayer network in which each aspect represents a functional brain network constructed during a different task.

The ability to combine different aspects (structure and function, healthy and disease groups, or time windows) within the same overall multilayer network framework allows one to develop network statistics that operate on a single object, using interaspect links to quantify the interactions between layers. This approach directly addresses the inherent dependencies between networks, offering a more principled framework for statistical inference.

Temporal Networks. The first use of multilayer networks in neuroscience addressed the temporal evolution of brain dynamics (Bassett, Wymbs, et al. 2011) in which each layer represented a functional network estimated from a single time window of fMRI data and nodes were linked across sequential layers to indicate identity. This specific instantiation of the framework is referred to as a temporal network (Holme 2012). This framework has proved particularly powerful in characterizing reconfiguration of human brain networks during learning (Bassett, Wymbs, et al. 2011, 2013; Bassett et al. 2015; Mantzaris et al. 2013), where the multilayer formalism enables a statistically robust identification of dynamic changes in community structure that explain individual differences in learning ability (Bassett, Wymbs, et al. 2011, 2013; Bassett et al. 2015). The tools have similarly proved useful in understanding lexical processing (Doron, Bassett, and Gazzaniga 2012), working memory (Braun et al. 2015), and motor behavior (Wymbs et al. 2012), and they complement a growing interest in dynamic functional connectivity in both rest (Hutchison et al. 2013) and task states.

Studies describing, predicting, and manipulating brain dynamics in these temporal networks compose the relatively new field of dynamic network neuroscience (Bassett et al. 2015). By modeling brain networks as evolving, interconnected systems, dynamic network neuroscience seeks to understand the complex nature of neurophysiological phenomena that underlie cognition. However, to achieve this goal, current network statistics must be modified to take into account the multidimensional aspects implicit in multilayer networks, a growing area of innovation (Kivelä et al. 2014). In addition to extending traditional statistics, there is a need to develop new statistics formalized specifically with the features of temporal networks in mind. Examples include network flexibility (Bassett, Wymbs, et al. 2011) and stationarity (Palla et al. 2005), which each quantify temporal changes in community structure. Finally, important challenges remain in defining appropriate statistical approaches to infer nonrandom structure in temporal networks (Bassett, Wymbs, et al. 2011; Bassett, Porter, et al. 2013), as well as tools to use these approaches to discriminate different roles that brain regions may play in network dynamics (Bassett, Wymbs, et al. 2013).
Conclusion. As technology advances, the neuroscience community is becoming increasingly inundated with rich data sets reflecting complex spatio-temporal interactions across multiple scales of the brain. To understand the intricacies of these new data and their utility in uncovering fundamental principles of brain function requires the use of a mathematical framework that incorporates the interconnected nature of the system. Network analysis, and especially dynamic network neuroscience built on multilayer network approaches, has proved to be a powerful tool in describing the complex organization and evolution of the human brain and its relationship to cognition. As the use of dynamic network neuroscience continues to evolve, we will be able to increasingly tease apart the complex relationships between brain structure and function, make predictions about the evolution of disease states, and understand how dynamic interactions between remote brain regions give rise to cognitive states that define human function.

REFERENCES