Network Models in Neuroscience

Danielle S. Bassett, Perry Zurn, and Joshua I. Gold

Abstract

From interacting cellular components to networks of neurons and neural systems, interconnected units comprise a fundamental organizing principle of the nervous system. Understanding how their patterns of connections and interactions give rise to the many functions of the nervous system is a primary goal of neuroscience. Recently, this pursuit has begun to benefit from the development of new mathematical tools that can relate a system's architecture to its dynamics and function. These tools, stemming from the broader field of network science, have been used with increasing success to build models of neural systems across spatial scales and species. This chapter discusses the nature of network models in neuroscience. It begins with a review of model theory from a philosophical perspective to inform our view of networks as models of complex systems, in general, and of the brain, in particular. It summarizes the types of models that are frequently studied in network neuroscience along three primary dimensions: from data representations to first-principles theory, from biophysical realism to functional phenomenology, and from elementary descriptions to coarse-grained approximations. Ways to validate these models are then considered, with a focus on approaches that perturb a system to probe its function. In closing, a description is provided of important frontiers in the construction of network models and their relevance for understanding increasingly complex functions of neural systems.

Introduction

The brain is composed of intricate networks that operate at many different levels of organization. At small spatial scales, gene regulatory networks direct neuronal cell fate, and both chemical and electrical synapses define the accessible routes of information transmission between neurons (Francis et al. 2003). At intermediate spatial scales, laminar architecture in cortex is accompanied by stereotyped interlaminar connectivity thought to support ensemble dynamics and resultant computations (Sherman et al. 2016). At even larger spatial scales, the anatomical locations of inter-areal projections display a

precise spatial arrangement associated with a diverse repertoire of functional processes (Betzel and Bassett 2018; Betzel et al. 2018).

Although networks are fundamental to brain structure, the complexity of these networks poses challenges to understanding their function. Unlike a sphere, which we can quickly guess to have the capacity to be rolled or thrown, a network—with its tangle of wires—defies any simple conspectus. Thus, even though more than a century has passed since Camillo Golgi, Santiago Ramón y Cajal, and other neuroanatomists introduced to the world the intricate beauty of the networks of neurons that comprise our nervous system, our understanding of how those networks give rise to perception, learning, memory, cognition, action, and other aspects of brain function remains incomplete.

To understand relationships between the brain's networked architecture and its many functions, one fruitful set of approaches stems from the emerging field of network science. This discipline addresses the study of systems whose structure, function, or dynamics depend upon the pattern of interconnections between units (Albert and Barabasi 2002). Network science is inherently interdisciplinary, drawing on and integrating among recent advances in mathematics, physics, computer science, and engineering (Newman 2010, 2011). Although early work in the field was largely devoted to the study of social systems, efforts over the last decade have focused increasingly on the study of neural systems across spatial scales, temporal scales, and species (Bullmore and Sporns 2009; Fornito et al. 2016; van den Heuvel et al. 2016). These newer efforts, collectively referred to as *network neuroscience*, model neural systems as networks to distill the dependence of brain function and dysfunction on interconnection architecture (Bassett and Sporns 2017).

Here, we review recent work in network neuroscience that has been applied to our collective quest to understand the brain, and emphasize the diversity of approaches that now fall under this general framework (Bassett et al. 2018). Because network neuroscience is fundamentally a modeling endeavor, we begin with a broad philosophical perspective on model theory. We then consider how networks are models, before turning to a discussion of the types of network models that are commonly used in neuroscience. Motivated by approaches to validate network models via prediction, we discuss the importance of perturbation-based techniques for understanding network function. We close by outlining important directions for future work to build, use, and validate network models in neuroscience.

Model Theory: A Philosophical Perspective

The term "model" can evoke quite different pictures in the mind's eye:

- a small, inexact replica of a 1910 Schacht Roadster
- a miniature cityscape coarsely true to the form of Cambridge, England

- Rodin's "The Thinker"
- a contemporary reworking of a piece from Greek mythology (Eugenides 2002)
- a person one wishes to emulate or an ideal one hopes to become
- IBM's Watson
- a cerebral organoid—miniature organ *in vitro* (Huang et al. 2017)—or a human blinking eye-on-a-chip (Chan et al. 2015)
- a set of interdependent partial differential equations producing dynamics reminiscent of a real-world system

In mentally traversing these diverse examples, one immediately realizes that the space of model types is exceedingly large, and one wonders whether it is even possible to define what a model actually is, or what it is not.

The question of how to define the term "model" is the focus of a branch of philosophy known as model theory, which aims to identify the essential elements that make models what they are and to disambiguate the characteristics that distinguish different types of models from one another (Gelfert 2016). At its most basic level, a model is a representation of one or more aspects of the world. It aims to increase understanding of what something is by measuring and imaging what something does. As such, models inherit a basic philosophical conundrum (Papineau 1987): What precisely is their relationship to the target systems they model, and from whence do they derive their truth value? Must they simply evidence functional coherence and have pragmatic purchase, or must they also meaningfully correspond to what they represent? If so, how is that meaningfulness determined?

In science, at least four types of models have been recognized, each of which provide different answers to these questions: scalar, idealized, analogical, and phenomenological models (Frigg and Hartmann 2012; Hartmann and Frigg 2012). Scalar models, much like the Roadster replica, either magnify or reduce their target systems. Idealized models abstract and isolate a limited set of features from their target systems. Analogical models highlight relevant similarities between two target systems, whether those similarities are shared properties or comparable structures. Finally, phenomenological models represent only the observable elements of their target systems, without postulating any theoretical explanation as to why those elements are what they are.

Recent work in model theory explores the intriguing possibility that all of these forms of scientific models are heuristic devices not unlike literary fictions (Suarez 2009; Frigg and Hunter 2010; Toon 2012). The models-as-fictions theory reconceptualizes models as fictional entities that aim to narrativize certain features of a target system (Barbrousse and Ludwig 2009; Godfrey-Smith 2009; Frigg 2010a, b, c; Garcia-Carpintero 2010; Toon 2012; Frigg and Nguyen 2016, 2017). Accordingly, models—much like fictions—may imaginatively isolate and abstract or distort and exaggerate certain features of the world in such a way as to facilitate epistemic access (Elgin 2010). They may

creatively instantiate either analogical or phenomenological substructures of their target systems in order to crystallize insight. Scientific models are therefore subject to evaluation at the level of both artistry (clarity, elegance, originality) and function. Moreover, just as the literary tradition provides new fiction with meaningful constraints in advance, the scientific community provides the parameters within which new models are developed and applied. Scientific models are, in this sense, accountable to the scientific communities that use them in the exploration of target systems that are already of particular value and interest (Almeder 2007). Finally, different sorts of models can be used together to build a multiscalar narrative architecture, modeling complementary features of a target system beside one another.

Whether used in isolation or in conjunction, scientific models illuminate the overarching structure of a target system precisely through the practice and provocation of creative imagination.

Networks as Models

The incipient challenge in modeling biological systems is to identify the most meaningful characteristics of the system that are distilled into a sensible representation (Bellomo et al. 2015). That is, biological models are inherently idealized models of complex systems, and their construction requires identifying first the form and degree of abstraction to use. This process requires a set of value judgments (Which characteristics are most meaningful?) and a commitment to epistemic cleanliness (What details of biology can we defensibly ignore?). These principles of valuation and purposeful ignorance are manifest even when exercising simple visual depictions, which arguably comprise the most impoverished of modeling approaches (Tufte 2001). One similarly faces choices of what to depict and what not to depict when building any simple mathematical representation of the system. For example, when building a differential equation to represent a system, one must choose which processes to encapsulate or not in a variable.

The fundamental assumption of network neuroscience is that idealized models of the brain should be constructed using analogical principles that focus on the networked architecture of the nervous system. As Cajal saw under his microscope, the nervous system is composed of individual neurons that are interconnected in complex ways. Accordingly, the earliest network models were idealized versions of this network structure, with nodes representing neurons and edges representing the connections between them. More recently, network models have been developed within and across multiple spatial and temporal scales, at the level of interconnected neurons as well as involving networks of subcellular components, multicellular systems, or both. As detailed below, these models can also be phenomenological, based on measured elements of the nervous system, or more theory driven. Despite

this diversity, these models retain key features that can be understood in terms of their basic idealized and analogical structure: an architecture built using interconnected units.

Such an architecture is typically encoded in a graph: an object composed of nodes, representing units of the system, and edges, representing interactions or links between those units (Bollobás 1979, 1985). Studies of graphs can be neatly separated into two categories: those who consider artificial graphs with arbitrary wiring principles (Harary 1969) and those who consider them to reflect the architecture of a real system (Cohen and Havlin 2010). In both cases, one seeks to describe the mathematical properties of the graph with the goal of understanding the function of the system. The patterns of which units can and cannot (or do and do not) interact with one another can allow one to deduce where information might be relatively more densely or relatively more sparsely located, where vulnerability might exist to injury or perturbation, and where circumscribed instances of collective dynamics might emerge (Albert et al. 2000; Cisneros et al. 2002; Gomez-Gardenes et al. 2007; Simonsen et al. 2008).

In simple graphs, all units are represented by identical nodes, and all edges are represented as either existing or not existing (Figure 7.1). These representations can be encoded using a binary weighting scheme. Furthermore, interactions are assumed to be bidirectional: if an edge exists between node *i* and node *i*, then an edge also exists between node *i* and node *i*. The very first formal network models of neural systems employed such binary, undirected graphs (Felleman and Van Essen 1991; Young et al. 1994; Scannell et al. 1995; Sporns et al. 2005; Achard et al. 2006; Kaiser and Hilgetag 2006). Nonetheless, it is relatively straightforward to adapt this encoding to a continuous weighting scheme, as well as to specify distinct weights for the edge from node *i* to node *j*, and for the edge from node *j* to node *i*. With the continued refinement of empirical measurement techniques, the inclusion of edge weights in network models has become increasingly prevalent, providing richer insights into system function and dynamics (Rubinov and Sporns 2011; Markov et al. 2013; Oh et al. 2014; Bassett and Bullmore 2017; Betzel and Bassett 2018).

Network models are simple constructs. They can be used effectively to study social, biological, technological, and physical systems (Newman 2010). Yet this flexibility is a marked reminder that the intuitions one gains from a network model depend strongly on what the nodes and edges are chosen to represent. A structural motif in a network of humans interlinked by friend-ships can mean something quite different than the same structural motif in a network of neurons interlinked by synapses. Thus, in any endeavor that translates a complex system into a network model, it is critical to specify exactly what the nodes and edges (or more complicated model components) represent, and to ensure that interpretations are drawn in accordance with those choices (Butts 2009).



Figure 7.1 Schematic of network models. Upper left: The simplest network model for neural systems is one that represents the pattern of connections (edges) between neural units (nodes). More sophisticated network models can be constructed by adding edge weights and node values, or explicit functional forms for their dynamics. Multilayer networks can be used to represent a set of interconnected networks; dynamic networks can be used to understand the reconfiguration of network systems over time. Bottom (left to right): Common measures of interest include degree (the number of edges emanating from a node), clustering (related to the prevalence of triangles), hubness (related to a node's influence), cavities (the absence of edges), communities (local groups of densely interconnected nodes), paths (which determine the potential for information transmission), and core-periphery structure, which facilitates local integration of information gathered from or sent to more sparsely connected areas.

Types of Network Models in Neuroscience

As a field, network neuroscience aims to build, exercise, and validate network models of neural systems with the explicit goal of better understanding brain structure and function, as well as cognition, behavior, and disease (Sporns 2014; Stam 2014; Medaglia et al. 2015; Fornito et al. 2017; Braun et al. 2018). The types of network models that are built share a similar analogical basis that emphasizes the importance of network-based architectures across spatial and temporal scales. These models, however, differ from one another in many important ways, which directly impact the sorts of inferences that can be justifiably drawn from them. Here we briefly describe recent efforts to systematize the study of network models in neuroscience by organizing these similarities and differences according to three dimensions (see Figure 7.2) that reflect the model categories described above (Bassett et al. 2018):



Figure 7.2 Efforts to understand mechanisms of brain structure, function, development, and evolution in network neuroscience can be organized along three key dimensions of model types. The first dimension (a) extends from elementary descriptions to coarse-grained approximations. The second (b) extends from biophysical realism to functional phenomenology. The third (c) extends from data representation to first-principles theory.

- their phenomenological basis, ranging from representations of measured phenomena to first-principles theory;
- their target of idealization, from biophysical to functional features; and
- their scalar focus, ranging from elementary descriptions to coarsegrained approximations.

The dimension from data representation to first-principles theory is arguably the most fundamental to network modeling efforts in neuroscience (Abbott 2008). Modeling efforts of the former type begin with empirically acquired data. They then seek to build a representation of those data by stipulating which part of the data to represent as a network node, and which part of the data to represent as a network node, and which part of the data to represent as a network node, and which part of the data to represent as a network node, and which part of the data to represent as a network edge. Intuitively, the data representation provides an abstract, nonvisual depiction or description of the system (for examples, see Young et al. 1994; Scannell et al. 1995; Watts and Strogatz 1998; Hilgetag et al. 2000; Stam 2004; Sporns et al. 2005; Achard et al. 2006; De Vico Fallani et al. 2006; Kaiser and Hilgetag 2006; Micheloyannis et al. 2006; Bettencourt

et al. 2007). In contrast, to make a prediction about system behavior either now or in the future, one must turn to models that instantiate first-principles theories. These models combine a network with a mathematical expression specifying the dynamics of network nodes, network edges, or collections of nodes and/or edges (see, e.g., Ritter et al. 2013; Roy et al. 2014; Gu et al. 2015; Falcon et al. 2016; Bezgin et al. 2017; Breakspear 2017; Melozzi et al. 2017; Yan et al. 2017; Kim et al. 2018). Data-driven network models enjoy the benefits of biological realism, whereas theory-based models have the capacity to make predictions and unearth function.

The dimension from biophysically to functionally defined features differentiates models that are physical in nature from those that are statistical in nature. Network models with biophysical realism are composed of nodes that represent physical units, including neurons, cortical columns, or Brodmann areas; and of edges that represent physical links, including synapses, projections, or white-matter tracts (see, e.g., Sporns et al. 2005; Kaiser et al. 2009; Bassett et al. 2010; Varshney et al. 2011; Nicosia et al. 2013; Oh et al. 2014). Network models addressing functional phenomenology are comprised of nodes and edges that are not necessarily physically instantiated but may instead be defined as statistical abstractions (Achard et al. 2006; Bettencourt et al. 2007; Chu et al. 2012; Burns et al. 2014; van Diessen et al. 2015; Khambhati et al. 2016). Common examples of such abstract edges are those that offer estimates of effective connectivity or functional connectivity, the latter of which are also referred to as noise correlations (Brody 1999a, b; Friston 2011a). It is often important to distinguish between these two types of models because they have distinct utility in assessing a network's physical constitution versus inferring its functional capacities.

The dimension from elementary descriptions to coarse-grained approximations is critical to support a multiscale understanding of brain structure, function, and dynamics. In general, network models can encode the organization of interconnections among cells, ensembles, cortical columns or subcortical nuclei, and large-scale brain areas. As evidenced by the diversity of scales represented in current empirical and theoretical investigations, no single level of description can provide a complete explanation for cognitive function and behavior. In many cases, however, it is worthwhile or at least practical to consider a single scale for a given study, and then to use insights gained at that scale to inform larger theories of multiscale function. The challenge in developing an appropriate network model at a particular scale is to ensure that the network nodes represent well-defined, discrete, non-overlapping units, and that network edges represent organic, irreducible relations (Butts 2009). Whereas models at the final spatial scale consider elementary building blocks (Brody 1999b; Sautois et al. 2007; Tang et al. 2008; Feldt et al. 2011; Teller et al. 2014; Kim and Lim 2015; Kaiser 2017; Mahadevan et al. 2017; Betzel et al. 2018), models at the coarse spatial scale consider emergent functions (Breakspear 2017).



Figure 7.3 Density of study in the three-dimensional space of network models. Density plot showing varying levels of study across the three-dimensional space of model types in network neuroscience. The center of the three-dimensional space (aubergine) is becoming increasingly accessible due to empirical and computational advances over the past few decades. The least well-studied space (periwinkle) represents first-principles theories of functional phenomenology at the elementary level of description. Clear volumes indicate spaces that are most commonly studied.

Together, these complementary dimensions define a three-dimensional space of network models that can be used to enhance our understanding of brain structure and function (Figure 7.3). Notably, prior modeling efforts have not been pursued with equal vigor in all volumes of this space, partly reflecting historical factors and the changing state of neurotechnologies. Early work focused on large-scale network models of anatomy, which represent coarse-grained data representations with biophysical realism (Bassett and Bullmore 2017; Liao et al. 2017). Less well-studied are first-principles theories of functional phenomenology, particularly among elementary units. With a marked increase in the pace of data acquisition and the capacity for data analysis, we anticipate increasing success developing network models at the center of the volume. In other words, we anticipate a wider array of studies that inform first-principles theories with data, complement physical models with statistical inferences on informational capacity, and build explicit multiscale accounts of network function. These multifaceted network models will enhance our ability to explain different parts, processes, or principles of the nervous system.

Modeling Perturbations to Networks

When building network models, we are often concerned with demonstrating their validity (Bassett et al. 2018). In prior work, we followed canonical principles for validating animal models of disease to suggest that network models can display three distinct types of validity (McKinney and Bunney 1969; Willner 1984; Shmueli 2010; Belzung and Lemoine 2011): descriptive, explanatory, and predictive. Intuitively, descriptive validity requires that the model resembles the system under study, a concept akin to face validity in animal models (McKinney and Bunney 1969; Willner 1984; Shmueli 2010; Belzung and Lemoine 2011; Willner 2017). For instance, a model with descriptive validity might accurately reflect the specific pattern of nodes and edges observed in anatomical or functional data. By contrast, explanatory validity requires that the model can be used to define statistical tests, for example by assessing causal relations based on the network's architecture. Finally, predictive validity is attained when there is a correlation between a network model's response to perturbation and an organism's response to that same perturbation (Belzung and Lemoine 2011). Such perturbative studies can be operationalized using stimulation, lesion, ablation, or drugs.

Predictive validity is often the final goal of any scientific domain of inquiry (Shneider 2009). Because predictive validity depends on understanding its response to a perturbation, it is of interest to consider the different ways in which a network model can be perturbed. Recent work in the physics and engineering communities has begun to focus on the means by which the architecture of a network determines how perturbations affect its function. A simple way in which to parse these studies is to consider separately perturbations applied: (a) to a single node or to a single edge ("point perturbations"), (b) to a set of nodes or to a set of edges, and (c) across a fixed area or volume of the network's topology. In the context of network neuroscience, these different types of perturbations may be accessible to distinct empirical techniques and collectively could be used to better understand both endogenous and exogenous control, thereby informing clinical intervention (Tang and Bassett 2018).

From a modeling perspective, point perturbations are perhaps the simplest type to study. Initial modeling approaches focused on point perturbations in the form of node or edge removal. Referred to as virtual lesioning at times, this approach was developed to quantify the robustness of a network by estimating the difference between the value of a graph statistic estimated before node or edge removal, and the value of that same graph statistic estimated after node or edge removal (see figure 1b in Dong et al. 2013). When nodes are removed at random, the approach is referred to as a random attack. When nodes are removed based on their topological role in the network, estimated using the values of various graph statistics such as degree and betweenness centrality, the approach is referred to as a targeted attack (Achard et al. 2006). These approaches have recently been used to better understand the impact of regional

dysfunction in schizophrenia and stroke (Alstott et al. 2009; Lynall et al. 2010; Lo et al. 2015).

Other approaches address how a perturbation of the activity of a node or edge can change the activity of other parts of the network. This approach is central to network control theory and built on the foundations of linear systems theory (Kailath 1980; Motter 2015). Here one considers the pattern of interconnections between units as well as a model of dynamics that specifies how the activity at one node can travel along edges to other nodes in the graph (Tang and Bassett 2018). By modeling both the connectivity and the dynamics, one can identify "driver" nodes with time-dependent control that can guide the system's activity (Liu et al. 2011). A recent application of these techniques to the connectome of Caenorhabditis elegans demonstrated that a particular network model had striking validity in predicting the effects of single-cell ablasions on the organism (Yan et al. 2017). The approach can also be extended beyond the identification of drivers controlling all dynamics to the identification of drivers controlling specific dynamics (Pasqualetti et al. 2014). The specificity of this extension allows for the study of unique control strategies within neural systems. A recent application of this technique provided an explanation for the anatomical location of areas of the brain involved in executive function, as those most capable of enacting modal controllability (Gu et al. 2015).

Despite their analytical tractability, point perturbations can be the most difficult to enact and interpret in the context of real neural systems. On a conceptual level, a single, functional node or edge used in a model might not have an obvious, well-defined anatomical substrate in the brain to target. On a practical level, even given a well-defined target, it may not be possible to perturb cleanly just that target, given the lack of complete specificity associated with current microstimulation, optogenetic, and pharmacological methods. Nonetheless, point perturbations represent a useful starting point in considering the validation of network models.

Moving beyond point perturbations, it is also of interest to consider perturbation to multiple points in the network, or to entire areas or volumes of neural systems. Intuitively, multipoint control is a natural reflection of circuit activity, where several areas may be activated simultaneously to orchestrate a change in communication or dynamics (Palmigiano et al. 2017). Multipoint control could also be fruitfully applied to the development of stimulation therapies to quiet seizure dynamics using implantable devices (Ehrens et al. 2015; Taylor et al. 2015; De Ridder et al. 2017; Jobst et al. 2017). Fortunately, the general network control framework is readily extended to account for the activity of multiple control points simultaneously, and can be used to model the propagation of stimulation directly along white matter tracts to predict distant effects on regional activity (Muldoon et al. 2016; Stiso et al. 2018). Extending these tools to affect control over continuous areas or volumes of a network is more difficult and remains an important area for future work. Progress in this area

is critical for extending network models to account for other chemical mechanisms of transcellular communication and the effects of glia, neuromodulatory systems, and other mechanisms on brain function and behavior (Borroto-Escuela et al. 2015; Safaai et al. 2015; Bruinsma et al. 2018; Savtchouk and Volterra 2018).

Modeling Network Growth and Evolution

The study of network perturbations, while useful for understanding endogenous and exogenous mechanisms of control, is also pertinent to an understanding of how neural systems came to be, how they develop, and how they age. A change in gene expression can alter the natural progression of cell fate from pluripotent stem cell through neuroprogenitor cell and eventually neuron (Mahadevan et al. 2017).

A fluctuation in chemical gradients can comprise a perturbation that alters the course of neuronal migration and, by extension, the location and density of synapses (Wrobel and Sundararaghavan 2014). In fully developed adult neurons, Hebb's rule essentially postulates that perturbations to neuronal firing can alter cellular-level network architecture (Bi and Poo 2001). Even at the large scale in humans, long-term training can induce changes in white matter architecture evident in noninvasive neuroimaging (Scholz et al. 2009). Understanding how perturbations of the organism or part of the organism over both short and long timescales affect network growth and evolution is an important open area of research.

Some progress has been made over the last few years in constructing socalled generative network models. Such models stipulate a wiring rule in the hope of producing a network architecture that displays topological or functional properties that are similar to those observed in the networks representing real systems (Betzel and Bassett 2017). The basic idea is that a wiring rule which produces a network topology similar to that observed in the real system is a candidate mechanism for network generation. The inference is made stronger if the wiring rule also displays characteristics thought to be consistent with biology, such as parsimony and efficiency. A common way of testing the pragmatic utility of the generative network model is to determine if it can be used to make out-of-sample predictions about held-out network data.

Generative network models tend to be built in one of three types: single-shot models, growth models, or developmental models (Figure 7.4). A single-shot model specifies a form for the connection probabilities, from which all edges and their weights are then drawn (Vertes et al. 2012, 2014; Beul et al. 2015, 2017; Betzel et al. 2016; Hilgetag et al. 2016). A growth model specifies a time-dependent wiring rule that indicates how nodes and possibly even edges are added over time (Klimm et al. 2014). Developmental models extend the biological realism of the effort even farther by specifying



Figure 7.4 Distinct classes of generative network models. Generative network models exist in three main classes that differ in the timescales over which they operate. Single-shot models specify a functional form for the probability with which any two nodes are linked with one another. Growth models specify rules by which nodes and/or edges are added to the network over time, which is commonly discretized in arbitrary units. Developmental models specify wiring rules with fixed timescales in actual units of seconds, minutes, days, months, years, etc., in an effort to match the true growth mechanisms of an organism better. From Betzel and Bassett (2017).

wiring rules in which the timescales of the model match the timescales of development in the organism under study (Nicosia et al. 2013). Together, these three types of generative network models vary in the timescale over which they operate and in their neurobiological plausibility.

Recently, generative network models have been developed and applied to explain neurophysiological and neuroanatomical data across both elementary descriptions and coarse-grained approximations (Beul et al. 2015, 2017). For example, a particularly striking single-shot model of the neural connectome of the nematode C. elegans demonstrated that a wiring rule based on the random outgrowth of axons, in combination with a competition for available space at the target neuron, was able to recapitulate the empirical network's edge length distribution (Kaiser et al. 2009). A developmental model of the same organism combined information regarding the pattern of interconnectivity between neurons with information regarding the birth times of neurons and their spatial locations (Nicosia et al. 2013). This study provided compelling evidence for a trade-off between the network's topology and cost that appears to be differentially negotiated over different developmental time periods. With a few exceptions (Vertes et al. 2012), most generative network models have focused on data representations more so than first-principles theories, and biophysical realism more so than functional phenomenology. Expanding efforts to fill the

full space of model types will be an important area for future work in generative modeling.

Future Directions

In considering the future utility of network models for advancing our understanding of neural systems, it is worth pointing out that the models used to date are relatively simple from a mathematical perspective. It remains an open question whether more complex network models might prove useful or merely obfuscate inference. To address this question, one could rationally assess whether some aspect of a known neurophysiological process remains unaccounted for by existing models. For example, to increase descriptive validity, one might wish to build an annotated network, where nodes can be assigned values or properties, reflecting for example cerebral glucose metabolism estimates from fluorodeoxyglucose (FDG)-positron emission tomography (PET), blood oxygen level dependent (BOLD) contrast imaging, magnetoencephalographic (MEG) or electroencephalographic (EEG) power, gray matter volume or cortical thickness, or cytoarchitectonic properties (Murphy et al. 2016; Newman and Clauset 2016). Furthermore, to increase explanatory validity, one might wish to build multilayer networks where the nodes and edges in each layer are obtained from different types of measurements (Kivel et al. 2014; Tewarie et al. 2014; Yu et al. 2016), and where the architecture of the network is allowed to vary over time in concert with system function (Holme and Saramaki 2012; Kopell et al. 2014; Breakspear 2017; Khambhati et al. 2017; Sizemore and Bassett 2018). Such richer models could allow one to test how network dynamics at one time point or in one modality might cause a change in network dynamics at another time point or in another modality.

A second way in which to address the question of whether more complex network models might prove useful is to consider whether the testing of a particular hypothesis requires a novel network model. For instance, recent efforts have provided initial evidence that some higher-order, non-pairwise interactions occur between neurons and between large-scale brain areas (Ganmor et al. 2011; Lord et al. 2016). Critically, all of the network models we have discussed here are based on pairwise interactions and cannot directly account for non-pairwise interactions (Petri et al. 2014; Sizemore et al. 2018). Tools that have been developed in the applied mathematics community that can encode and characterize higher-order relations include hypergraphs (an edge can link any number of vertices) and simplicial complexes (higher-order interaction terms become fundamental units) (Bassett et al. 2014; Giusti et al. 2016). These generalizations of graphs may be critical for an accurate understanding of neuronal codes and associated computations both at micro- and macroscales (Curto et al. 2013; Reimann et al. 2017; Sizemore et al. 2017). As the field

moves beyond univariate accounts to postulate more network-based hypotheses, richer network models may be required.

Conclusion

From cellular to regional scales, neural circuitry is an interconnected system. In such a system, network modeling is a particularly useful approach for distilling interconnection patterns into tractable mathematical objects that are amenable to theory. Here we discussed the nature of network models, which share a similar networked architecture that is justified in terms of its analogies to brain structure but then differ along several dimensions: from data representations to first-principles theory, abstractions that emphasize biophysical or functional features, and different scales from elementary descriptions to coarse-grained approximations. We paid particular attention to models that have been developed to better understand the response of networked systems to perturbation enacted at a single point, at multiple points, or across extended areas or volumes of the organism. We also offered an extended discussion of generative network models that seek to identify candidate wiring mechanisms for circuit evolution or development. We suggest that network models are particularly appropriate for neural systems. Accordingly, future advances in our understanding of computation and cognition will depend on the expansion of these models in mathematical sophistication and the development of richer, network-based hypothesis of brain structure, function, and dynamics.

Acknowledgments

We thank Blevmore Labs and Ann E. Sizemore for efforts in graphic design. We also thank D. Lydon-Staley, A. E. Sizemore, E. Cornblath, and D. Zhou for helpful comments on an earlier version of this manuscript. D. S. B. acknowledges support from the John D. and Catherine T. MacArthur Foundation, the Alfred P. Sloan Foundation, the ISI Foundation, the Paul Allen Foundation, the Army Research Laboratory (W911NF-10-2-0022), the Army Research Office (Bassett-W911NF-14-1-0679, Grafton-W911NF-16-1-0474, DCIST-W911NF-17-2-0181), the Office of Naval Research, the National Institute of Mental Health (2-R01-DC-009209-11, R01-MH112847, R01-MH107235, R21-M MH-106799), the National Institute of Child Health and Human Development (1R01-HD086888-01), National Institute of Neurological Disorders and Stroke (R01-NS099348), and the National Science Foundation (BCS-1441502, BCS-1430087, NSF PHY-1554488 and BCS-1631550). J. I. G. acknowledges support from the National Science Foundation (NSF-NCS 1533623), the National Eye Institute (R01-EY015260), and the National Institute of Mental Health (R01-MH115557). The content is solely the responsibility of the authors and does not necessarily represent the views of any of the funding agencies.