Network models in neuroscience

Danielle S. Bassett1,2,3,4,7, Perry Zurn5, and Joshua I. Gold6

1Department of Bioengineering, School of Engineering and Applied Sciences, University of Pennsylvania, Philadelphia, PA, 19104
2Department of Physics & Astronomy, College of Arts and Sciences, University of Pennsylvania, Philadelphia, PA, 19104
3Department of Electrical & Systems Engineering, School of Engineering and Applied Sciences, University of Pennsylvania, Philadelphia, PA, 19104
4Department of Neurology, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, 19104
5Department of Philosophy, American University, Washington, DC, 20016
6Department of Neuroscience, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, 19104 and
7To whom correspondence should be addressed: dsb@seas.upenn.edu

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, which are known collectively as network science, have been used with increasing success to build models of neural systems across spatial scales and species. Here we discuss the nature of network models in neuroscience. We begin 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. We then summarize 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. We then consider ways to validate these models, focusing on approaches that perturb a system to probe its function. We close with a description 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. For example, 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 [1]. At intermediate spatial scales, laminar architecture in cortex is accompanied by stereotyped inter-laminar connectivity thought to support ensemble dynamics and resultant computations [2]. 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 [3,4].

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 Ramon 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.

One fruitful set of approaches for understanding relationships between the brain’s networked architecture and its many functions that has emerged in recent years is network science. This discipline deals with the study of systems whose structure, function, or dynamics depend upon the pattern of interconnections between units [5]. Network science is inherently interdisciplinary, drawing on and integrating among recent advances in mathematics, physics, computer science, and engineering [6,7]. 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 [8,10]. 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 [11].

In this chapter, we review recent work in network neuroscience that has been applied to understanding the brain, emphasizing the diversity of approaches that now fall under this general framework [12]. Because network neuroscience is fundamentally a modelling endeavor, we begin with a broad perspective on model theory from philosophy. 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 bring quite different pictures to 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 [13]. A person one wishes to emulate, or an ideal one hopes to become. IBM’s Watson. A cerebral organoid (miniature organ in vitro [14]) or a human blinking eye-on-a-chip [15]. 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 [16]. At its most basic level, a model is a representation of one or more aspects of the world. It aims to better understand what something is by measuring and imaging what something does. As such, models inherit a basic philosophical conundrum [17]: 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, and, if so, how is that meaningfulness determined?

In science, at least four types of models have been recognized that each provide different answers to these questions. These types are: scalar, idealized, analogical, and phenomenological [18][19]. 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 [20][22]. The models-as-fictions theory reconceptualizes models as fictional entities that aim to narrativize certain features of a target system [23][31]. According to this framework, 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 [32]. 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, so 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 [33]. Finally, different sorts of models can be used together to build a multi-scalar 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 [34]. 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 (“What are the most meaningful characteristics?”) and a commitment to epistemic cleanliness (“What details of biology can we defensively 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 [35]. 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 and not encapsulate 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 not just the level of interconnected neurons but also networks of subcellular components, multicellular systems, or both. As detailed below, these models also can 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, which represent units of the system, and edges, which represent interactions or links between those units [36, 37]. Studies of graphs can be neatly separated into two categories: those that consider artificial graphs with arbitrary wiring principles [38], and those that consider graphs that reflect the architecture of a real system [39]. 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 [40–43].

In simple graphs, all units are represented by identical nodes, and all edges are represented as either existing or not existing (Fig. 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 $j$, then an edge also exists between node $j$ and node $i$. The very first formal network models of neural systems employed such binary, undirected graphs [44–49]. 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 [3, 50–53].

Network models are simple constructs. They can be used effectively to study social, biological, technological, and physical systems [60]. Yet this flexibility is a marked reminder that the intuitions that 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 friendships 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 [54].

**TYPES OF NETWORK MODELS IN NEUROSCIENCE**

Network neuroscience as a field is devoted to building, exercising, and validating network models of neural systems with the explicit goal of better understanding brain structure and function, as well as cognition, behavior, and disease [55–59]. 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. However, these models differ from one another in many important ways that 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 that reflect the model categories described above [12]: first, their phenomenological basis, ranging from representations of measured phenomena to first-principles theory; second, their target of idealization, from biophysical to functional features; and third, their scalar focus, ranging from elementary descriptions to coarse-grained approximations (Fig. 2).

The dimension from data representation to first-principles theory is arguably the most fundamental to network modeling efforts in neuroscience [60]. 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 edge. Intuitively, the data representation provides an abstract, non-visual depiction or description of the system (for a few early examples, see [44, 49, 61–66]). 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, for example, [67, 75]). 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, for example [48, 49, 53, 76, 78]). 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 [47, 55, 79–82]. 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 [83–85].
FIG. 1. Schematic of network models. **(Top Left Corner)** 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, and dynamic networks can be used to understand the reconfiguration of network systems over time. **(Bottom Right Corner)** Common measures of interest include: degree, which is the number of edges emanating from a node; clustering, which is related to the prevalence of triangles; cavities, which describe the absence of edges; hubness, which is related to a node’s influence; paths, which determine the potential for information transmission; communities, or local groups of densely interconnected nodes; shortcuts, which are one possible marker of global efficiency of information transmission; and core-periphery structure, which facilitates local integration of information gathered from or sent to more sparsely connected areas. Figure adapted with permission from [12].

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. Yet in many cases, it is worthwhile or at least practical to consider a single scale for a given study and 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 network nodes represent well-defined, discrete, non-overlapping units, and that network edges represent organic, irreducible relations [1,2]. Whereas models at the final spatial scale consider elementary building blocks [4,86–92], models at the coarse spatial scale consider emergent functions [67].

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 (Fig. 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 biophysical realism [60,93,94]. 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 capacities 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 multi-faceted 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. 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: descriptive, explanatory, and predictive. Intuitively, descriptive validity requires that the model resemble the system under study, a concept akin to face validity in animal models. For example, a model with descriptive validity might accu-

FIG. 2. Three dimensions of network model types. (a) We posit that efforts to understand mechanisms of brain structure, function, development, and evolution in network neuroscience can be organized along three key dimensions of model types. (b) The first dimension extends from elementary descriptions to coarse-grained approximations. (c) The second dimension extends from biophysical realism to functional phenomenology. (d) The third dimension extends from data representation to first-principles theory. Figure adapted with permission from [12].
Functional data approaches focused on point perturbations in the form of node or perhaps the simplest type to study. Initial modeling approaches to better understand both endogenous and exogenous perturbations may be accessible to different types of perturbations. In the context of network neuroscience, these perturbations represent a useful starting point in considering separately perturbations to multiple points in the network model. The center of the 3-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.

Predictive validity is often the final goal of any scientific domain of inquiry. 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 how 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: (i) to a single node or to a single edge (“point perturbations”), (ii) to a set of nodes or to a set of edges, and (iii) 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.

From a modeling perspective, point perturbations are perhaps the simplest type to study. Initial modeling approaches focused point perturbations in the form of node or edge removal. At times referred to as virtual lesioning, 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 (Fig. 1). 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. These approaches have recently been used to better understand the impact of regional dysfunction in schizophrenia and stroke.

Other approaches address how a perturbation 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. 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. By modeling both the connectivity and the dynamics, one can identify “driver” nodes with time-dependent control that can guide the system’s activity. A recent application of these techniques to the connectome of C. elegans demonstrated that a particular network model had striking validity in predicting the effects of single cell ablations on the organism. The approach can also be extended beyond the identification of drivers controlling all dynamics to the identification of drivers controlling specific dynamics. 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.

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 cleanly perturb 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, multi-point control is a natural reflection of circuit activity, where several areas may be activated simultaneously to orchestrate a change in communication or dynamics. Multi-point control could also be fruitfully applied to the development of stimulation therapies to quiet seizure dynamics using implantable devices.
Fortunately, the general network control framework is readily extended to account for the activity of multiple control points simultaneously, and can be used to directly model the propagation of stimulation along white matter tracts to predict distant effects on regional activity. 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.

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. 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. In fully developed adult neurons, Hebb’s rule essentially postulates that perturbations to neuronal firing can alter cellular-level network architecture. Even at the large scale in humans, long-term training can induce changes in white matter architecture evident in noninvasive neuroimaging. In fully developed adult neurons, Hebb’s rule essentially postulates that perturbations to neuronal firing can alter cellular-level network architecture. Even at the large scale in humans, long-term training can induce changes in white matter architecture evident in noninvasive neuroimaging. Understanding how perturbations of the organism or part of the organism over both short and long time scales affect network growth and evolution is an important open area of research.

Some progress has been made over the last few years in constructing so-called 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. The basic idea is that a wiring rule that 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, and developmental models. A single-shot model specifies a form for the connection probabilities, from which all edges and their weights are then drawn. A growth model specifies a time-dependent wiring rule that indicates how nodes and possibly even edges are added over time. Developmental models extend the biological realism of the effort even farther by specifying wiring rules in which the time scales of the model match the time scales of development in the organism under study. Together, these three types of generative network models vary in the time scale 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. 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. 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. 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, 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. 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, and where the architecture of the network is allowed to vary over time in concert with
FIG. 4. Distinct classes of generative network models. Generative network models exist in three main classes that differ in the timescales over which they operate. (Left) Single-shot models specify a functional form for the probability with which any two nodes are linked with one another. (Middle) Growth models specify rules by which nodes and/or edges are added to the network over time, which is commonly discretized in arbitrary units. (Right) Developmental models specify wiring rules with fixed timescales in actual units of seconds, minutes, days, months, years, etc. in an effort to better match the true growth mechanisms of an organism [77]. Figure reproduced with permission from [121].

system function [67, 138–141]. 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 example, recent efforts have provided initial evidence that some higher-order, non-pairwise interactions occur between neurons and between large-scale brain areas [142, 143]. Critically, all of the network models we have discussed here are based on pairwise interactions and cannot directly account for non-pairwise interactions [144, 145]. 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) [146, 147]. These generalizations of graphs may be critical for an accurate understanding of neuronal codes and associated computations both at micro- and macro-scales [148–150]. 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 upon the expansion of these models in mathematical sophistication, and the development of richer, network-based hypothesis of brain structure, function, and dynamics.
ACKNOWLEDGEMENTS

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 National Institutes of Health (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 truth.

COMPETING INTERESTS

The authors declare no competing interests.

[1] Francis, M. M., Mellem, J. E. & Maricq, A. V. Bridging the gap between genes and behavior: recent advances in the electrophysiological analysis of neural function in Caenorhabditis elegans. *Trends Neurosci* **26**, 90–99 (2003).

[2] Sherman, M. A. et al. Neural mechanisms of transient neocortical beta rhythms: Converging evidence from humans, computational modeling, monkeys, and mice. *Proc Natl Acad Sci U S A* **113**, E4885–E4894 (2016).

[3] Betzel, R. F. & Bassett, D. S. The specificity and robustness of long-distance connections in weighted, interareal connectomes. *Proc Natl Acad Sci U S A* **Epub Ahead of print** (2018).

[4] Betzel, R. F., Medaglia, J. D. & Bassett, D. S. Diversity of meso-scale architecture in human and nonhuman connectomes. *Nature Communications In Press* (2018).

[5] Albert, E. & Barabasi, A.-L. Statistical mechanics of complex networks. *Reviews of Modern Physics* **74** (2002).

[6] Newman, M. E. J. *Networks: An Introduction* (MIT Press, 2010).

[7] Newman, M. E. J. Complex systems: A survey. *Am. J. Phys.* **79**, 800–810 (2011).

[8] Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat Rev Neurosci* **10**, 186–198 (2009).

[9] Fornito, A., Zalesky, A. & Bullmore, E. T. *Brain Network Analysis* (Oxford University Press, 2016).

[10] van den Heuvel, M. P., Bullmore, E. T. & Sporns, O. Comparative connectomics. *Trends Cogn Sci* **20**, 345–361 (2016).

[11] Bassett, D. S. & Sporns, O. Network neuroscience. *Nat Neurosci* **20**, 353–364 (2017).

[12] Bassett, D. S., Zurn, P. & Gold, J. I. On the nature and use of models in network neuroscience. *Nature Reviews Neuroscience In Press* (2018).

[13] Eugenides, J. *Middlesex* (Farrar, Straus and Giroux, 2002).

[14] Huang, J. et al. Tranylcypromine causes neurotoxicity and represses BHC110/LSD1 in human-induced pluripotent stem cell-derived cerebral organoids model. *Front Neurol* **8**, 626 (2017).

[15] Chan, Y. K. et al. In vitro modeling of emulsification of silicone oil as intraocular tamponade using microengineered eye-on-a-chip. *Invest Ophthalmol Vis Sci* **56**, 3314–3319 (2015).

[16] Gelfert, A. *How to Do Science with Models: A Philosophical Primer* (2016).

[17] Papineau, D. *Reality and Representation* (Blackwell, 1987).

[18] Hartmann, S. & Frigg, R. Scientific models. In et al., S. S. (ed.) *The Philosophy of Science: An Encyclopedia*, vol. 2, 740–749 (Routledge, 2012).

[19] Frigg, R. & Hartmann, S. Models in science (2012). URL https://plato.stanford.edu/entries/models-science

[20] Toon, A. *Models as Make-Believe: Imagination, Fiction, and Scientific Representation* (Palgrave, 2012).

[21] Frigg, R. & Hunter, M. C. *Beyond Mimesis and Convention: Representation in Art and Science* (Springer, 2010).

[22] Suarez, M. *Fictions in Science: Philosophical Essays on Modelling and Idealization* (Routledge, 2009).

[23] Barbrousse, A. & Ludwig, P. Fictions and models. In Suarez, M. (ed.) *Fictions in Science: Philosophical Essays on Modelling and Idealization*, 56–75 (Routledge, 2009).

[24] Frigg, R. Models and fiction. *Synthese* **172**, 251–268 (2010).

[25] Frigg, R. & Nguyen, J. The fiction view of models reloaded. *The Monist* **99**, 225–242 (2016).

[26] Godfrey-Smith, P. Models and fictions in science. *Philosophical Studies* **143**, 101–116 (2009).

[27] Toon, A. *Models as Make-Believe: Imagination, Fiction, and Scientific Representation* (Palgrave, 2012).

[28] Frigg, R. Fiction and scientific representation. In Frigg, R. & Hunter, M. C. (eds.) *Beyond Mimesis and Convention: Representation in Art and Science*, 97–138 (2010).

[29] Frigg, R. Fiction in science. In Woods, J. (ed.) *Fictions and Models: New Essays*, 247–287 (Philosophia Verlag, 2010).

[30] Garcia-Carpintero, M. Fictional entities, theoretical models, and figurative truth. In Frigg, R. & Hunter,
Gomez-Gardenes, J., Moreno, Y. & Arenas, A. Paths to synchronization on complex networks. Phys Rev Lett 12, 44–64 (2015).

Tufto, E. R. The visual display of quantitative information (2001).

Bollobas, B. Graph Theory: An Introductory Course (Springer-Verlag, 1979).

Bellomo, N., Elaiw, A., Althiabi, A. M. & Alghamdi, M. A. On the interplay between mathematics and biology: hallmarks toward a new systems biology. Phys Life Rev 12, 171–195 (2007).

Gomez-Gardenes, J., Moreno, Y. & Arenas, A. Paths to synchronization on complex networks. Phys Rev Lett 98, 034101 (2007).

Cisneros, L., Jimenez, J., Cosenza, M. G. & Parravano, A. Information transfer and nontrivial collective behavior in chaotic coupled map networks. Phys Rev E Stat Nonlin Soft Matter Phys 65, 045204 (2002).

Simonson, I., Buzna, L., Peters, K., Bornholdt, S. & Helbing, D. Transient dynamics increasing network vulnerability to cascading failures. Phys Rev Lett 100, 218701 (2008).

Albert, R., Jeong, H. & Barabasi, A. L. Error and attack tolerance of complex networks. Nature 406, 378–382 (2000).

Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex 1, 1–47 (1991).

De Vico Fallani, F., Sporns, O., Tononi, G. & Kötter, R. The human connectome: a structural description of the human brain. PLoS Comput Biol 1, e42 (2005).

Kaiser, M. & Hilgetag, C. C. Nonoptimal component placement, but short processing paths, due to long-distance projections in neural systems. PloS Computational Biology 2, e95 (2006).

Bassett, D. S. & Bullmore, E. T. Small-world brain networks revisited. Neuroscientist Sep 21, 1073585416667720 (2016).

Rubinov, M. & Sporns, O. Weight-conserving characterization of complex functional brain networks. Neuroimage 56, 2068–2079 (2011).

Markov, N. T. et al. The role of long-range connections on the specificity of the macaque interareal cortical network. Proceedings of the National Academy of Sciences 110, 5187–5192 (2013).

Qin, S. W. et al. A mesoscale connectome of the mouse brain. Nature 508, 207–214 (2014).

Butts, C. T. Revisiting the foundations of network analysis. Science 325, 414–416 (2009).

Medaglia, J. D., Lynall, M. E. & Bassett, D. S. Cognitive network neuroscience. J Cogn Neurosci 27, 1471–1491 (2015).

Sporns, O. Contributions and challenges for network models in cognitive neuroscience. Nat Neurosci 17, 652–660 (2014).

Stam, C. J. Modern network science of neurological disorders. Nat Rev Neurosci 15, 683–695 (2014).

Braun, U. et al. From maps to multi-dimensional network mechanisms of mental disorders. Neuron 97, 14–31 (2018).

Fornito, A., Bullmore, E. T. & Zalesky, A. Opportunities and challenges for psychiatry in the connectomic era. Biol Psychiatry Cogn Neurosci Neuroimaging 2, 9–19 (2017).

Abbott, L. F. Theoretical neuroscience rising. Neuron 60, 489–495 (2008).

Sporns, O. Contributions and challenges for network models in cognitive neuroscience. Nat Rev Neurosci 17, 652–660 (2014).

Stam, C. J. Functional connectivity patterns of human magnetoencephalographic recordings: a ‘small-world’ network? Neurosci Lett 355, 25–28 (2004).

De Vico Fallani, F. et al. Brain connectivity structure in spinal cord injured: evaluation by graph analysis. Conf Proc IEEE Eng Med Biol Soc, 988–991 (2006).

Micheloyannis, S. et al. Small-world networks and disturbed functional connectivity in schizophrenia. Schizophr Res 87, 60–66 (2006).

Bettencourt, L. M., Stephens, G. J., Ham, M. I. & Gross, G. W. Functional structure of cortical neuronal networks grown in vitro. Phys Rev E 75, 021915 (2007).

Breakspear, M. Dynamic models of large-scale brain activity. Nat Neurosci 20, 340–352 (2017).

Melozzi, F., Woodman, M. M., Jirsa, V. K. & Bernard, C. The virtual mouse brain: A computational neuroinformatics platform to study whole mouse brain dynamics. eNeuro 4, ENEURO.0111-17.2017 (2017).

Bezgin, G., Solodkin, A., Bakker, R., Ritter, P. & McIntosh, A. R. Mapping complementary features of cross-species structural connectivity to construct realistic “Virtual Brains”. Hum Brain Mapp 38, 2080–2093 (2017).

Ritter, P., Schirmer, M., McIntosh, A. R. & Jirsa, V. K. The virtual brain integrates computational modeling and multimodal neuroimaging. Brain Connect 3, 121–145 (2013).

Roy, D. et al. Using the virtual brain to reveal the role of oscillations and plasticity in shaping brain’s dynamical landscape. Brain Connect 4, 791–811 (2014).

Falcon, M. I., Jirsa, V. & Solodkin, A. A new neuroinformatics approach to personalized medicine in neurology: The Virtual Brain. Curr Opin Neurol 29, 429–436
Controllability of structural brain networks.

Network control principles predict neuron function in the Caenorhabditis elegans connectome. Nature 550, 519–523 (2017).

Structural properties of the Caenorhabditis elegans neuronal network. PLoS Comput Biol 7, e1001066 (2011).

Phase transition in the economically modeled growth of a cellular nervous system. Proceedings of the National Academy of Sciences USA 110, 7880–7885 (2013).

Efficient physical embedding of topologically complex information processing networks in brains and computer circuits. PLoS Comput Biol 6, e1000748 (2010).

Opportunities and methodological challenges in EEG and MEG resting state functional brain network research. Clin Neurophysiol 126, 1468–1481 (2015).

Emergence of stable functional networks in long-term human electroencephalography. J Neurosci 32, 2703–2713 (2012).

Network dynamics of the brain and influence of the epileptic seizure onset zone. Proc Natl Acad Sci U S A 111, E5321–E5330 (2014).

Virtual cortical resection reveals push-pull network control preceding seizure evolution. Neuron 91, 1170–1182 (2016).

Functional and effective connectivity: a review. Brain Connect 1, 13–36 (2011).

Correlations without synchrony. Neural Comput 11, 1537–1551 (1999).

Disambiguating different covariation types. Neural Comput 11, 1527–1535 (1999).

Dissecting functional connectivity of neuronal microcircuits: experimental and theoretical insights. Trends Neurosci 34, 225–236 (2011).

Fast sparsely synchronized brain rhythms in a scale-free neural network. Phys Rev E 92, 022717 (2015).

Robustness of network of networks under targeted attack. Phys Rev E Stat Nonlin Soft Matter Phys 87, 052804 (2013).

Functional connectivity and brain networks in schizophrenia. J Neurosci 30, 9477–9487 (2010).

Randomization and resilience of brain functional networks as systems-level endophenotypes of schizophrenia. Proc Natl Acad Sci U S A 112, 9123–9128 (2015).

The impact of lesions in the human brain. PLoS Comput Biol 5, e1000408 (2009).

Networkcontrolology. Chaos 25, 097621 (2015).

Linear Systems (Prentice-Hall, Inc., 1980).

Controllability of complex networks. Nature 473, 167–173 (2011).

Controlability metrics, limitations and algorithms for complex networks. IEEE Transactions on Control of Network Systems 1, 40–52 (2014).

Fast flexible information routing by transient synchrony. Nat Neurosci 20, 1014–1022 (2017).

Novel applications for cortical stimulation. Neuro modulation 20, 206–214 (2017).

Closed-loop control of a fragile network: application to seizure-like dynamics of an epilepsy model. Front Neurosci 9, 58 (2015).

Optimal control based seizure abatement using patient derived connectivity. Front Neurosci 9, 202 (2015).

Brain-responsive neurostimulation in patients with medically intractable seizures arising from eloquent and other neocortical areas. Epilepsia 58, 1005–1014 (2017).
[115] Muldoon, S. F. et al. Stimulation-based control of dynamic brain networks. *PLoS Comput Biol* 12, e1005076 (2016).

[116] Stiso, J. et al. White matter network architecture guides direct electrical stimulation through optimal state transitions. *arXiv* 1805.01260 (2018).

[117] Borroto-Escuela, D. O. et al. The role of transmitter diffusion and flow versus extracellular vesicles in volume transmission in the brain neural-glial networks. *Philos Trans R Soc Lond B Biol Sci* 370 (2015).

[118] Savtchouk, I. & Volterra, A. Gliotransmission: Beyond black-and-white. *J Neurosci* 38, 14–25 (2018).

[119] Kaiser, M., Hilgetag, C. C. & van Ooyen, A. A simple architecture and distance. *Brain Struct Funct* 220, 3167–3184 (2015).

[120] Bruinsma, T. J. et al. The relationship between dopamine neurotransmitter dynamics and the blood-oxygen-level-dependent (BOLD) signal: A review of pharmacological functional magnetic resonance imaging. *Front Neurosci* 12, 238 (2018).

[121] Betzel, R. F. & Bassett, D. S. Generative models for network neuroscience: prospects and promise. *J R Soc Interface* 14, 20170623 (2017).

[122] Wrobel, M. R. & Sundararaghavan, H. G. Directed migration in neural tissue engineering. *Tissue Eng Part B Rev* 20, 93–105 (2014).

[123] Bi, G. & Poo, M. Synaptic modification by correlated activity: Hebb’s postulate revisited 24, 139–166 (2001).

[124] Scholz, J., Klein, M. C., Behrens, T. E. & Johansen-Berg, H. Training induces changes in white-matter architecture. *Nat Neurosci* 12, 1370–1371 (2009).

[125] Vertes, P. E. et al. Simple models of human brain functional networks. *Proc Natl Acad Sci U S A* 109, 5868–5873 (2012).

[126] Vertes, P. E., Alexander-Bloch, A. & Bullmore, E. T. Generative models of rich clubs in Hebbian neuronal networks and large-scale human brain networks. *Philos Trans R Soc Lond B Biol Sci* 369, 1653 (2014).

[127] Betzel, R. F. et al. Generative models of the human connectome. *Neuroimage* 124, 1054–1064 (2016).

[128] Beul, S. F., Grant, S. & Hilgetag, C. C. A predictive model of the cat cortical connectome based on cytoarchitecture and distance. *Brain Struct Funct* 220, 3167–3184 (2015).

[129] Beul, S. F., Barbas, H. & Hilgetag, C. C. A predictive structural model of the primate connectome. *Sci Rep* 7, 43176 (2017).

[130] Hilgetag, C. C., Medalla, M., Beul, S. F. & Barbas, H. The primate connectome in context: Principles of connections of the cortical visual system. *Neuroimage* 134, 685–702 (2016).

[131] Klimm, F., Bassett, D. S., Carlson, J. M. & Mucha, P. J. Resolving structural variability in network models and the brain. *PLoS Computational Biology* 10, e1003491 (2014).

[132] Kaiser, M., Hilgetag, C. C. & van Ooyen, A. A simple rule for axon outgrowth and synaptic competition generates realistic connection lengths and filling fractions. *Cereb Cortex* 19, 3001–3010 (2009).

[133] Newman, M. E. J. & Clauset, A. Structure and inference in annotated networks. *Nature Communications* 7, 11863 (2016).

[134] Murphy, A. C. et al. Explicitly linking regional activation and function connectivity: Community structure of weighted networks with continuous annotation. *arXiv* 1611.07962 (2016).

[135] Kivel, M. et al. Multilayer networks. *J Complex Netw.* 2, 203–271 (2014).

[136] Tewarie, P. et al. Structural degree predicts functional network connectivity: a multimodal resting-state fMRI and MEG study. *Neuroimage* 97, 296–307 (2014).

[137] Yu, Q. et al. Building an EEG-fMRI multi-modal brain graph: A concurrent EEG-fMRI study. *Front Hum Neurosci* 10, 476 (2016).

[138] Holme, P. & Saramäki, J. Temporal networks. *Phys. Rep.* 519, 97–125 (2012).

[139] Khamhanti, A. N., Szimore, A. E., Betzel, R. F. & Bassett, D. S. Modeling and interpreting mesoscale network dynamics. *Neuroimage* 5193–5199, 30500–1 (2017).

[140] Szimore, A. E. & Bassett, D. S. Dynamic graph metrics: Tutorial, toolbox, and tale. *Neuroimage* 5193–5199, 30564–5 (2017).

[141] Kopell, N. J., Gritton, H. J., Whittington, M. A. & Kramer, M. A. Beyond the connectome: the dynome. *Neuron* 83, 1319–1328 (2014).

[142] Gammor, E., Segev, R. & Schneidman, E. Sparse low-order interaction network underlies a highly correlated and learnable neural population code. *Proc Natl Acad Sci U S A* 108, 9679–9684 (2011).

[143] Lord, L. D. et al. Insights into brain architectures from the homological scaffolds of functional connectivity networks. *Front Syst Neurosci* 10, 85 (2016).

[144] Szimore, A. E., Phillips-Cremins, J., Ghrist, R. & Bassett, D. S. The importance of the whole: topological data analysis for the network neuroscientist. *arXiv* 1806.01311 (2018).

[145] Petri, G. et al. Homological scaffolds of brain functional networks. *J R Soc Interface* 11, 20140873 (2014).

[146] Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J. & Grafton, S. T. Cross-linked structure of network evolution. *Chaos* 24, 013112 (2014).

[147] Giusti, C., Ghrist, R. & Bassett, D. S. Two’s company, three (or more) is a simplex: Algebraic-topological tools for understanding higher-order structure in neural data. *J Comput Neurosci* 41, 1–14 (2016).

[148] Szimore, A. E. et al. Cliques and cavities in the human connectome. *J Comput Neurosci* Epub Ahead of Print (2017).

[149] Reimann, M. W. et al. Cliques of neurons bound into cavities provide a missing link between structure and function. *Front Comput Neurosci* 11, 48 (2017).

[150] Curto, C., Itskov, V., Veliz-Cuba, A. & Youngs, N. The neural ring: an algebraic tool for analyzing the intrinsic structure of neural codes. *Bull Math Biol* 75, 1571–1611 (2013).