

The Connected Brain

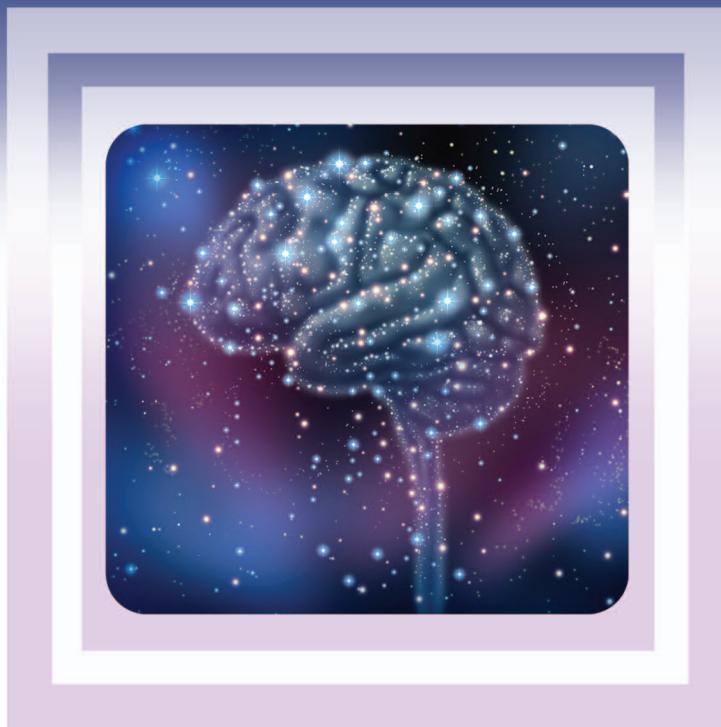


IMAGE LICENSED BY INGRAM PUBLISHING

Causality, models, and intrinsic dynamics

Adeel Razi and Karl J. Friston

Recently, there have been several concerted international efforts—the BRAIN Initiative, the European Human Brain Project, and the Human Connectome Project, to name a few—that hope to revolutionize our understanding of the connected brain. During the past two decades, functional neuroimaging has emerged as the predominant technique in systems neuroscience. This is foreshadowed by an ever-increasing number of publications on functional connectivity, causal modeling, connectomics, and multivariate analyses of distributed patterns of brain responses. In this article, we summarize pedagogically the (deep) history of brain mapping. We highlight the theoretical advances made in the (dynamic) causal modeling of brain function, which may have escaped the wider audience of this article, and provide a brief overview of recent developments and interesting clinical applications. We hope that this arti-

cle engages the signal processing community by showcasing the inherently multidisciplinary nature of this important topic and the intriguing questions that are being addressed.

Introduction

In this article, we use several key dichotomies to describe the evolution and emergence of modeling techniques used to characterize brain connectivity. We provide a historical overview of the brain connectivity literature, starting with the fundamental distinction between functional segregation and integration. In so doing, we introduce a key difference between functional and effective connectivity and emphasize their relationship via underlying models of distributed processing. Next, we consider various causal modeling techniques that are used to infer directed brain connectivity. With the help of a unified framework—based on (neuronal) state-space models—we show how (with a succession of simplifying approximations) standard models of connectivity

Digital Object Identifier 10.1109/MSP.2015.2482121
Date of publication: 27 April 2016

can be derived and how various measures of statistical dependencies arise from a generative (state-space) model of neuronal dynamics. Finally, we focus on the application of dynamic causal modeling (DCM) to endogenous neuronal activity and simulations of neuronal fluctuations based on the connectome. We describe a series of recent (and rapid) developments in modeling distributed neuronal fluctuations and how this modeling rests on functional connectivity. We contextualize these developments in terms of some historical distinctions that have shaped our approaches to connectivity in functional neuroimaging.

Notation

We use lowercase italics, x , for scalars and lowercase bold for vectors, \mathbf{x} , and vector functions, $\mathbf{x}(t)$, where each element represents a time-dependent state. Matrices are shown as uppercase bold, \mathbf{X} . In this article, $*$ corresponds to a convolution operator, \dagger denotes the complex conjugate transpose, $\langle \cdot \rangle$ denotes expectation, and \sim denotes discrete time-lagged variables. Fourier transforms of variables are in italic uppercase, such that $\mathbf{FT}(\mathbf{x}(t)) = \mathbf{X}(\omega)$. We use $F(\cdot)$ to denote a variational free-energy functional.

A historical perspective on brain connectivity

The notion of connectivity has a long history in brain imaging that can be traced back to the debates around classicism, modularity, and connectionism. In the recent past, a common notion among neuroscientists was that many functions of the brain were predetermined by its structure and that its structure was programmed by our genes. This view emphasized functional segregation and localizationism, tracing its history back to the days of phrenology (from Gall in the 18th century). Functional localization implies that a function can be localized in a cortical area. This is more general than functional segregation, which suggests that a cortical area is specialized for some aspect of neural processing and that this specialization is anatomically segregated within the cortex. This is similar to an understanding of how computers work, where each part has a preassigned function that cannot be substituted with other parts. However, in past decades, this view has changed, with clear evidence that the neural pathways in our brain are flexible, adaptable, connected, and moldable by changes in our environment or by injury or disease. In short, the brain is quintessentially plastic and can adapt and adopt new functionalities through necessity. This understanding rests on the notion of *connectionism* (a term first coined by Donald Hebb in the 1940s), with the central idea that brain function can be understood as the interaction among simple units, for example, neurons connected by synapses, that give rise to a connected whole that changes over time. Connectionism is closely related to (hierarchical) distributed processing, a perspective that has been substantiated by the work of Hubel and Wiesel (recipients of the Nobel Prize in Physiology or Medicine 1981) on how information

is processed in the visual cortex. They found that the visual system comprises simple and complex cells arranged in a hierarchical fashion. This finding underwrites the focus on neural network implementations based on hierarchical distributed constructs, leading to recent exciting developments in machine learning (e.g., hierarchical Bayesian inference [1] and deep learning algorithms [2]).

These ideas emerged in functional brain imaging as functional segregation and functional integration. Since their inception, there has been a sustained trend to move from functional segregation (and the study of regionally specific brain activation) toward functional integration (and the study of its connectivity). Functional localization implies that a function can be localized to a cortical area, whereas segregation suggests that a

cortical area is specialized for some aspects of perceptual or motor processing and that this specialization is anatomically segregated within the cortex. The cortical infrastructure supporting a single function may then involve many specialized areas whose union is mediated by the functional integration among them. In this view, functional segregation is meaningful only in the context of functional integration and vice versa. There are several descriptions of neuronal processing that accommodate the tendency for brain regions to engage in specialized functions (i.e., segregation) and the tendency to coordinate multiple functions (i.e., integration) through coupling specialized regions. This functional integration is a dynamic self-assembling process, with parts of the brain engaging and disengaging over time, and has been described by appealing to dynamical systems theory, for example, self-organized criticality [3], pattern formation, and metastability [4].

This review pursues another key theme—the distinction between functional and effective connectivity. This dichotomy relies on the definition of connectivity (i.e., functional integration) per se. The former uses a pragmatic definition of connectivity based on (Pearson) correlations and rests on statistical dependencies between remote neurophysiological events. However, this approach is problematic when dealing with distributed neuronal processes in the brain that are mediated by slender (axonal) neuronal connections or wires. A more mechanistic explanation of observed responses comes from the definition of effective connectivity that refers explicitly to the influence that one neural system exerts over another. In [5], it was proposed that “effective connectivity should be understood as the experiment and time-dependent, simplest possible circuit diagram that would replicate the observed timing relationships between the recorded neurons.” This speaks to two important points: effective connectivity is dynamic (activity dependent) and depends on a model of directed interactions or coupling, which we focus on in this review. Given this, an interesting development in functional connectivity now considers temporal dynamics, referred to as *dynamic functional connectivity* [6]. However, these developments fall short of furnishing a causal explanation of the sort provided by (model-based)

Several concerted international efforts hope to revolutionize our understanding of the connected brain.

(or difference equations in discrete time) and therefore explicitly take time into account. This simple example emphasizes the importance of temporal fluctuations in connectivity, even in undirected graphs. However, we do not want to give the impression that temporal precedence is necessary to infer causal relationships. Temporal precedence is an important aspect, and many definitions of causation require cause to precede effect [12], [13], for example, directed functional connectivity measures based on Yule–Walker formulations [vector autoregressive (VAR) models]. However, temporal precedence alone cannot distinguish effective connectivity from spurious dependencies caused by unknown factors. As an example, the barometer falls before the rain, but it does not cause the rain. The type of causality that we are concerned with is based on control theoretic concepts, where the causes (exogenous experimental inputs, endogenous random neural fluctuations, or both) produce effects (neural activity) that are observed empirically through hemodynamics as blood oxygen level-dependent (BOLD) signals. This form of causality is closely related to the probabilistic and graphical framework of causal calculus [14] (see “Simpson–Yule Paradox”), although there is a clear distinction between the two approaches, which we return to later.

We use state-space models to describe the basic concepts here and demonstrate that causality based on temporal precedence can be regarded as a special case of causality based on state-space graphs. In what follows, we look at several measures of causality in functional neuroimaging literature (which refer largely to fMRI but also hold for other modalities such as EEG, MEG, and local field potentials). These measures can be cast in terms of a generalization of state-space models based on stochastic differential equations.

State-space modeling of neuronal dynamics

The most natural framework for modeling distributed and coupled neural activity is to use state-space models. State-space modeling has its origin in control engineering, but the term *state-space* was first used by Kalman [24] and can be traced back to von Bertalanffy, who introduced general systems theory to biology in the 1940s and 1950s. We start with a generic description of coupled neuronal dynamics in terms of differential equations of the form

$$\dot{\mathbf{x}} = f(\mathbf{x}(t), \boldsymbol{\theta}, \mathbf{u}(t)) + \mathbf{w}(t) \quad (\text{state equation}), \quad (1)$$

$$\mathbf{y}(t) = h(\mathbf{x}(t), \boldsymbol{\theta}) + \mathbf{e}(t) \quad (\text{observation equation}), \quad (2)$$

where $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_n(t)]^T$ represents a vector of n hidden state variables (where each state could correspond to a vast number of neurons in a cortical area, source, or spatial mode); $\dot{\mathbf{x}}(t)$ represents the change in those state variables; $\boldsymbol{\theta}$ are the underlying (connectivity) parameter that are assumed to be time-invariant; $\mathbf{y}(t)$ is the observed BOLD signal; and $\mathbf{w}(t)$ and $\mathbf{e}(t)$ are state noise (observation or instrument noise,

respectively), which makes this differential equation random. (Note: Strictly speaking, the hidden states include both neuronal and hemodynamic states; however, for simplicity, we ignore hemodynamic states in this article.) The (random) endogenous fluctuations $\mathbf{w}(t)$ on the motion of the hidden neuronal states represent the unknown influences (e.g., spontaneous fluctuations) that can only be modeled probabilistically. (Note: A reviewer of this article rightly pointed out that, in this exposition, we limited ourselves to an additive form of endogenous fluctuations that precludes the more general treatment of state-dependent neuronal fluctuation of the sort $f(\mathbf{x}(t), \boldsymbol{\theta}, \mathbf{u}(t), \mathbf{w}(t))$, which are used in modeling many complex volatile systems [25], including the brain [26].) The neuronal states are hidden because they cannot be measured directly. The function f defines the motion of the coupled dynamical system that is determined by inputs $\mathbf{u}(t)$, which we consider to be deterministic (but could also have stochastic component) and known. Inputs usually pertain to experimentally controlled variables, such as change in stimuli (a visual cue or an auditory signal) or instructions during an fMRI

experiment (we see later that this exogenous input is absent in resting-state fMRI). This description of neuronal dynamics provides a convenient model of causal interactions among neuronal populations because it describes when and where exogenous experimental input $\mathbf{u}(t)$ perturbs the system and how (hidden) states influence changes in other states. Note that we have assumed that the form of the system dependencies f (and the connectivity parameters $\boldsymbol{\theta}$) are time-invariant, which means that we are assuming that the structural properties of the system will remain fixed over time (i.e., during the length of data acquisition).

We have not discussed the nature of the state and the observation noise process, which we consider in the section “Dynamic Casual Modeling of Intrinsic Networks.” For now, we assume that they possess usual noise properties, that is, they are independent and identically distributed. We describe a more general framework for analytic (non-Markovian) random fluctuations in the same section. A key simplification in this form of modeling is that we have lumped together many microscopic neuronal states to form hidden states $\mathbf{x}(t)$ that are abstract representations of neuronal activity (cf. a mean field approximation). In reality, the equations of motion—and the observer equation—describe very complicated interactions among millions of neurons. The formulation above corresponds to the amplitude of macroscopic variables or order parameters summarizing the dynamics of large neuronal populations. (Note: In statistical physics, the order parameter is a variable that indicates which phase you are in; for example, in a phase transition between liquid and gas, the order parameter may be the density.) Essentially, this means that the individual neurons become ordered, showing a coordinated dynamic pattern that can be described with the concept of order parameters. This sort of formulation can be motivated by basic

The notion of connectivity has a long history in brain imaging that can be traced back to debates around classicism, modularity, and connectionism.

Simpson–Yule Paradox

The Simpson–Yule paradox, or simply Simpson’s paradox [15]–[17], refers to the disconcerting situation in which statistical relationships between variables (e.g., x and y) are reversed or negated by the inclusion of an additional variable (z); for a more recent discussion, see [18]–[20]. A famous example of this paradox is when the University of California, Berkeley, came under investigation in 1975 for gender bias in graduate admissions. The graduate admissions statistics revealed that men applying were more likely to be admitted than women. However, when data were analyzed for each department separately, the reverse was true: no department was statistically significant in favor of men. The resolution of this paradox turned out to be that women applied for more competitive departments—with low success rates—in relation to men, who applied for fewer competitive majors with greater chances of acceptance. The main point is that conclusions based on data are sensitive to the variables we choose to hold constant, and that is why the “adjustment problem” is so critical in

the analysis of observational studies. Even now, no formal procedure has emerged that tells us whether adjusting for variable z is appropriate for the given study, setting aside intractable criteria [21] based on counterfactuals [22]. However, Simpson’s paradox is easily resolved with causal graphs. A simple graphical procedure provides a general solution to the adjustment problem [23]. This procedure is shown in Figure S1 and summarized as follows:

Objective

Check if z_1 and z_2 are sufficient measurements.

- 1) z_1 and z_2 should not be descendants of x .
- 2) Delete all nonancestors of $\{x, y, z\}$.
- 3) Delete all edges from x .
- 4) Connect any two parents sharing a child.
- 5) Strip arrowheads from all edges.
- 6) Delete z_1 and z_2 . Check if x is disconnected from y in the remaining graph; then z_1 and z_2 are appropriate measurements.

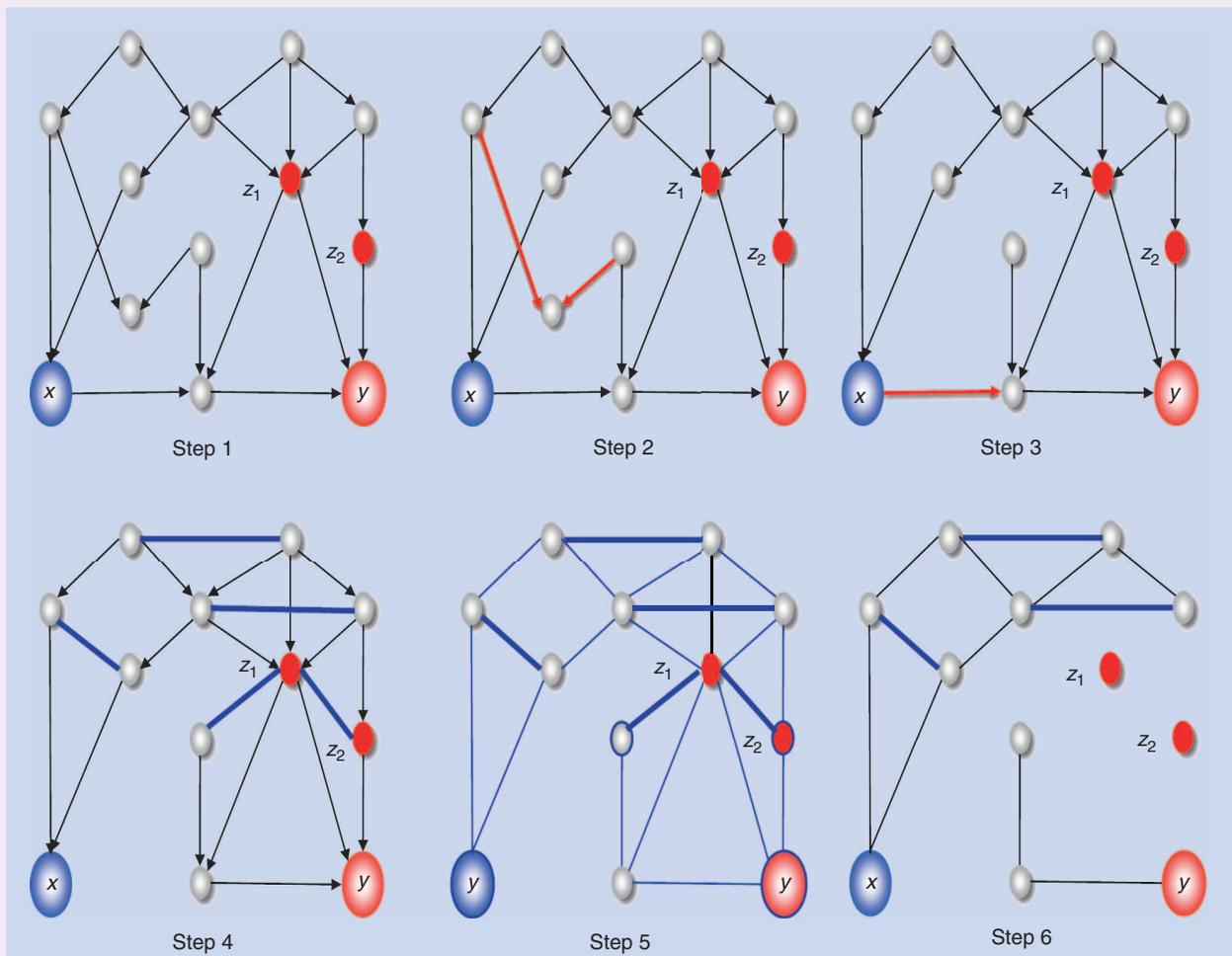


FIGURE S1. A simple graphical procedure provides a general solution to the adjustment problem. (Figure reproduced and redrawn with permission from [23].)

principles [27], for example, the center manifold theorem [28] and the slaving principle [29], [30], that apply generally to coupled dynamical systems.

State-space modeling and effective connectivity

The state and observation equations in (1) and (2) are generic representations; hence, there are several forms that the mappings or functions f and h can take. In turn, these define the sort of inference that can be made and the nature of causal relationships that can be identified from these models. We see in this section that almost all models in neuroimaging can be viewed as special cases of these equations.

Dynamic causal modeling

Although the application of general state-space models in neuroimaging has been around for decades, the explicit use of state-space models based on differential equations can be traced to [31], with the first introduction of a nonlinear neural mass model for EEG data. However, the most widely used and comprehensive framework, which uses Bayesian statistics to make model and parameter inferences, is DCM [32]. When first introduced, DCM used an ordinary differential equation (ODE) but was later extended to state-space models based on stochastic and random differential equations [33], [34]. The most widely used DCM is based on a Taylor expansion of (1) to its bilinear approximations:

$$\dot{\mathbf{x}}(t) = \left(\mathbf{A} + \sum_{j=0}^J \mathbf{B}^j \mathbf{u}_j \right) \mathbf{x}(t) + \mathbf{C} \mathbf{u}(t) + \mathbf{w}(t), \quad (3)$$

where $\mathbf{A} = \partial f / \partial \mathbf{x}$, $\mathbf{B} = \partial^2 f / \partial \mathbf{x} \partial \mathbf{u}$, and $\mathbf{C} = \partial f / \partial \mathbf{u}$ with $\boldsymbol{\theta}_n = \{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$. The matrix \mathbf{A} is known as the Jacobian (or Laplace–Beltrami operator) describing the behavior—that is, the effective connectivity—of the system near its fixed point ($f(\mathbf{x}_0) = 0$), in the absence of the fluctuations $\mathbf{w}(t)$ and the modulatory inputs $\mathbf{u}(t)$. The matrices \mathbf{B}^j encode the change in effective connectivity induced by the j th input $\mathbf{u}_j(t)$, and \mathbf{C} embodies the strength of the direct influences of inputs $\mathbf{u}(t)$ on neural activity. In fMRI, the mapping from hidden states to the observed BOLD data $y(t)$ is based on a hemodynamic model that transforms hidden neuronal states of each population or region into predicted BOLD responses using a previously established biophysical model [32], [35], [36]. This hemodynamic model is based on four ODEs and five hemodynamic parameters $\boldsymbol{\theta}_h$, such that $\boldsymbol{\theta} = \{\boldsymbol{\theta}_n, \boldsymbol{\theta}_h\}$. The hemodynamic model describes how neuronal activity engenders vasodilatory signals that lead to increases in blood flow, which in turn changes the blood volume and deoxyhemoglobin content, which subtend the measured signal.

The bilinear approximation to our general state-space model of neurophysiological dynamics furnishes a probabilistic model that specifies the probability of observing any time series given the parameters. This is known as a *likelihood model* and usually assumes that the observed data are a linear mixture of the model predictions and Gaussian observation noise. By combining this likelihood model with prior beliefs (specified in terms of probability distributions), we have

what is called, in Bayesian statistics, a *generative model*. This allows one to use standard (variational) procedures to estimate the posterior beliefs about the parameters and, crucially, the model itself. The real power of DCM lies in the ability to compare different models of the same data. This comparison rests on the model evidence, which is simply the probability of the observed data under the model in question (and given known or designed exogenous inputs). The evidence is also called the *marginal likelihood* because one marginalizes or removes dependencies on the unknown quantities (hidden states and parameters). The model evidence can simply be written as

$$p(\mathbf{y} | m, \mathbf{u}) = \int p(\mathbf{y}, \mathbf{x}, \boldsymbol{\theta} | m, \mathbf{u}) d\mathbf{x} d\boldsymbol{\theta}. \quad (4)$$

Model comparison rests on the evidence for one model relative to another (see [51] for a discussion in the context of fMRI). Model comparison based on the likelihood of different models provides the quantitative basis for all evidence-based hypothesis testing. Usually one selects the best model using Bayesian model comparison, in which different models are specified in terms of priors on the coupling parameters. These are used to switch off parameters by assuming a priori that they are zero (to create a new model). In DCM, priors used are so-called “shrinkage priors” because the posterior estimates shrink toward the prior mean. The size of the prior variance determines the amount of shrinkage. With a null model m_0 and an alternative model m_1 , the Bayesian model comparison rests on computing the logarithm of the evidence ratio

$$\begin{aligned} \ln \left(\frac{p(\mathbf{y} | m_1)}{p(\mathbf{y} | m_0)} \right) &= \ln p(\mathbf{y} | m_1) - \ln p(\mathbf{y} | m_0) \\ &\approx F(\mathbf{y}, \boldsymbol{\mu}_1) - F(\mathbf{y}, \boldsymbol{\mu}_0), \end{aligned} \quad (5)$$

where $F(\cdot)$ is the free energy that provides an (upper bound) approximation to Bayesian model evidence. Note that we have expressed the logarithm of the marginal likelihood ratio as a difference in log evidences. This is the preferred form because model comparison is not limited to two models but can cover a large number of models whose quality can be usefully quantified in terms of their log evidences. A relative log evidence of three corresponds to a marginal likelihood ratio (Bayes factor) of about 20 to one, which is considered strong evidence in favor of one model over another [37]. An important aspect of model evidence is that it includes a complexity cost (which is sensitive not only to the number of parameters but also to their interdependence). This means that a model with redundant parameters would have less evidence, even though it provided a better fit to the data (see [51]). In most current implementations of DCM, the log evidence is approximated with a (variational) free-energy bound that (by construction) is always less than the log evidence. As we see in (5), this bound is a function of the data and (under Gaussian assumptions about the posterior density) some proposed values for the states and parameters. When the free energy is maximized (using gradient ascent) with respect to the proposed values, they become the maximum posterior or conditional estimates, $\boldsymbol{\mu}$, and the free energy, $F(\mathbf{y}, \boldsymbol{\mu}_1) \leq \ln p(\mathbf{y} | m)$, approaches the log evidence.

We return to Bayesian model comparison and the inversion of DCMs in the section “Biophysical Modeling of Neuronal Dynamics.” Next, we consider some alternative models. The first is a discrete time linear approximation to (1), which is the basis of Granger causality.

Vector autoregressive modeling

In contrast to DCM, in which causality is based on control theoretic constructs, (multivariate) autoregressive models [38]–[40] use temporal precedence for inferring causality in BOLD time series [41]. This is known as *directed functional connectivity* in neuroscience. It is straightforward to see that one can convert a state-space model—or DCM—into a VAR model with a few simplifying assumptions. Using a linear approximation to the state-space model of (1) and assuming that we can measure the neuronal states directly [i.e., $\mathbf{y}(t) = \mathbf{x}(t)$], then we can write

$$\mathbf{y}(t) = \tilde{\mathbf{A}}\mathbf{x}(t - \delta) + \mathbf{z}(t), \quad (6)$$

which can be written as

$$\mathbf{Y} = \tilde{\mathbf{Y}}\tilde{\mathbf{A}}^T + \mathbf{Z},$$

where $\tilde{\mathbf{A}} = \exp(\delta\mathbf{A})$ and $\mathbf{z}(t) = \int_0^\delta \exp(\tau\mathbf{A})\mathbf{w}(t - \tau) d\tau$. The second equality expresses the resulting VAR model as a simple general linear model with explanatory variables $\tilde{\mathbf{Y}}$ that correspond to a time-lagged (time \times source) matrix of states. Here, the unknown parameters comprise the autoregression matrix $\tilde{\mathbf{A}}$. Note that the innovations, $\mathbf{z}(t)$, are now a mixture of past fluctuations in $\mathbf{w}(t)$ that are remembered by the system. There is a clear distinction between fluctuations $\mathbf{w}(t)$ that drive the hidden states (1) compared with the innovations $\mathbf{z}(t)$ in (6) that underlie autoregressive dependencies among observation $\mathbf{y}(t)$. There is an important point to note here. Because the reparameterization of the effective connectivity in (3) uses a matrix exponential, the autoregressive coefficients $\tilde{\mathbf{A}}$ in (6) are no longer the parameters of the underlying effective connectivity among neuronal states. This means that any model comparisons—based on classical likelihood ratio tests such as Bayesian information criterion—will be making inferences about the statistical dependencies modeled by the autoregressive process and not about the causal coupling as in DCM. This is why connectivity measures based on autoregressive coefficients, for example, Granger causality [42], are regarded as directed functional connectivity as opposed to effective connectivity. A further distinction is that most Granger causality applications either ignore hemodynamic convolution or assume that hemodynamics are identical and noiseless [147]. An important aspect of Granger causality measures based on autoregressive formulations (we provide analytic links between the two in Figure S2) is that they can become unreliable in the presence of measurement noise and more so when underlying dynamics are dominated by slow (unstable) modes, quantified by the principal Lyapunov exponent [43]. However, there are several recent advances in the Granger causality literature that speak to these limitations [44]–[46].

Structural equation modeling

Structural equation modeling (SEM) [47] is another generic approach developed primarily in economics and social sciences [48], [49] and was used in (structural) neuroimaging for the first time in [50]. We can again see that SEM is a special case of (1) by appealing to the (adiabatic) assumption that neuronal dynamics have reached equilibrium at each point of observation—or, in other words, the dynamics are assumed to occur over a timescale that is short relative to the fMRI sampling interval. In terms of implementation, we can force this condition by having very strong shrinkage priors in DCM. With this assumption, we can reduce the generative model of (3) so that it predicts the observed covariance among regional responses over time instead of predicting the time series itself. Mathematically, this means that we assume $\mathbf{y}(t) = \mathbf{x}(t)$, $\mathbf{u}(t) = 0$, and $\dot{\mathbf{x}}(t) = 0$. This simply means that $\mathbf{x}(t) = \mathbf{y}(t) = -\mathbf{A}^{-1}\mathbf{w}(t)$, which implies that

$$\Sigma_y = \mathbf{A}^{-1}\Sigma_w(\mathbf{A}^{-1})^T, \quad (7)$$

where $\Sigma_y = \langle \mathbf{y}(t)\mathbf{y}(t)^T \rangle$ and $\Sigma_w = \langle \mathbf{w}(t)\mathbf{w}(t)^T \rangle$. Note that we do not have to estimate hidden states because the generative model explains observed covariances in terms of random fluctuations and unknown coupling parameters. The form of (7) has been derived from the generic generative model. In this form, it can be regarded as a Gaussian process model, where the coupling parameters become, effectively, parameters of the covariance among observed signals due to the hidden states. We can also give an alternative formulation of SEM in terms of path coefficients, but we skip this for brevity (for details, see [51]).

Although SEM has been used in fMRI literature, it provides a description of static dependencies; hence, it is not suitable for fMRI (and EEG/MEG) time series, in which the characteristic time constants of the neuronal dynamics and hemodynamics are much larger than the exogenous inputs that drive them. This means that testing for context-sensitive changes in effective connectivity becomes problematic in event-related designs. For example, [52] used simulated fMRI time series from a realistic network model for two task conditions in which the anatomical connectivity was known and could be manipulated. The results suggested that caution is necessary in applying SEM to fMRI data and illustrate that functional interactions among distal network elements can appear abnormal even if only part of a network is damaged.

Another issue when using SEM to infer effective connectivity is that we can only use models of low complexity—usually, (acyclic) models that have no recurrent connections [53]. This is because fitting the sample covariance means that we have to throw away lots of information in the original time series. Heuristically, the ensuing loss of degrees of freedom means that conditional dependencies among the estimates of effective connectivity are less easy to resolve. In machine-learning literature, SEM can be regarded as a generalization of inference on linear Gaussian Bayesian networks that relaxes the acyclic constraint. As such, it is a generalization of structural causal modeling, which deals with directed acyclic graphics (DAGs) (see next

section). This generalization is important in the neurosciences because of the ubiquitous reciprocal connectivity in the brain that renders it cyclic or recursive. Next, we turn to the description of time series based on second-order statistics and show that they can be analytically derived from the state-space model of (1).

Coherence, cross spectra, and correlations

Until now, we have considered only procedures for identifying effective connectivity from fMRI time series. However, the following important question remains: Is there an analytical relationship between functional and effective connectivity? This question is addressed schematically in “Measures of Connectivity” by showing how various measures of statistical dependencies (functional connectivity) are interrelated and how they can be generated from a DCM. This schematic contextualizes different measures of functional connectivity and how they arise from (state-space) models of effective connectivity. In other words, measures that are typically used to characterize observed data can be regarded as samples from a probability distribution over functions whose expectation is known. This means that one can treat normalized measures, such as cross-correlation functions and spectral Granger causality, as explicit functions of the parameters of the underlying generative process.

In “Measures of Connectivity,” we include common (descriptive) measures of functional connectivity that have been used in fMRI, such as the correlation coefficient (the value of the cross-correlation function at zero lag), coherence, and (Geweke) Granger causality [54]. These measures can be regarded as standardized (second-order) statistics based on the cross-covariance function, the cross-spectral density, and the directed transfer functions, respectively. In turn, they are determined by the first-order (Volterra) kernels, their associated transfer functions, and VAR coefficients. For readers not familiar with Volterra kernels, their use provides an alternative to the conventional identification methods by expressing the output signal as high-order nonlinear convolution of the inputs. This can simply be thought of as a functional Taylor expansion and can be regarded as a power series with memory (see [55] for a detailed discussion). All of these representations can be generated from the underlying state-space model used by DCM. Let us examine these relationships further. First, there is a distinction between the state-space model (upper two panels of Figure S2), which refers to hidden or system states, and representations of dependencies among observations (lower panels), which do not. This is important because, although one can generate the dependencies among observations from the state-space model, one cannot do the converse. In other words, it is not possible to derive the parameters of the state-space model (e.g., effective connectivity) from transfer functions or autoregression coefficients. This is why one needs a state-space model to estimate effective connectivity or, equivalently, why effective connectivity is necessarily model-based. Second, we have seen in previous sections that SEM and autoregressive representations can be derived from (bilinear

and stochastic) DCM in a straightforward manner (under certain assumptions). The convolution kernel representation in Figure S2 provides a crucial link between covariance-based second-order measures, such as cross covariance and cross correlation, and their spectral equivalents, such as cross spectra and coherence. Figure S2 also highlights the distinction between second-order statistics (lower two rows) and models of the variables per se (upper three rows). For example, convolution and autoregressive representations can be used to generate time series (or their spectral counterparts), whereas cross-covariance functions and autoregression coefficients describe their second-order behavior. This is important because this second-order behavior can be evaluated directly from observed time series. Indeed, this is the common way of measuring functional connectivity in terms of

(second-order) statistical dependencies. We also highlight the dichotomy between time and frequency representations (measures in the top panel). For example, the (first-order Volterra) kernels in the convolution formulation are the Fourier transform of the transfer functions in frequency space (and vice versa). Similarly, the directed transfer functions of the autoregressive formulation are based on the Fourier transforms of the

autoregression coefficients. Another distinction is between representations that refer explicitly to random (state and observation) noise and autoregressive representations that do not. For example, notice that the cross-covariance functions of the data depend on the cross-covariance functions of state and observation noise. Conversely, the autoregression formulation only invokes (unit normal) innovations (although the autoregression coefficients are an implicit function of both state and observation noise covariance functions). In the current setting, autoregressive representations are not regarded as models but simply as ways of representing dependencies among observations. This is because (hemodynamic) responses do not cause responses—hidden (neuronal) states cause responses.

Crucially, all of the aforementioned formulations of statistical dependencies contain information about temporal lags (in time) or phase delays (in frequency). This means that, in principle, all measures are directed in the sense that the dependencies from one region to another are distinct from the dependencies in the other direction. However, only the autoregressive formulation provides directed measures of dependency—in terms of directed transfer functions or Granger causality. This is because the cross-covariance and cross-spectral density functions between two time series are antisymmetric. The autoregressive formulation can break this (anti)symmetry because it precludes instantaneous dependencies by conditioning the current response on past responses. Note that Granger causality is, in this setting, a measure of directed functional connectivity [56]. This means that Granger causality (or the underlying autoregression coefficients) reflects directed statistical dependencies such that two regions can have strong autoregression coefficients or Granger causality in the absence of a direct effective connection. Finally,

Key modeling assumptions underlying DCM are motivated by neuroanatomical and neurophysiological constraints.

Measures of Connectivity

The upper panel in Figure S2 illustrates the form of a state-space model that comprises differential equations coupling hidden states (first equation) and an observer equation mapping hidden states $\mathbf{x}(t)$ to observed responses $\mathbf{y}(t)$ (second equation). Dynamic causal models are summarized by a Taylor (bilinear) approximation. Crucially, both the motion of hidden states and the responses are subject to random fluctuations, also known as state $\mathbf{w}(t)$ and observation $\mathbf{e}(t)$ noise. The form of these fluctuations is modeled in terms of their cross-covariance functions $\Sigma(t)$ of time t or cross-spectral density functions $\mathbf{g}(\omega)$ of (radial) frequency ω , as shown in the lower equations. Given this state-space model and its parameters θ (which include effective connectivity), one can now parameterize a series of representations of statistical dependencies among successive responses as shown in the third row. These include convolution and autoregressive formulations shown on the left and right, respectively, in either time (pink and orange) or frequency (light green) space. The mapping between these representations rests on the Fourier transform, denoted by a dotted line, and its inverse. For example, given the equations of motion and observer function of the state-space model, one can compute the convolution kernels that, when applied to state noise, produce

the response variables. This allows one to express observed responses in terms of a convolution of hidden fluctuations and observation noise. The Fourier transform of these convolution kernels $\kappa(t)$ is called a transfer function $\mathbf{K}(\omega)$. Note that the transfer function in the convolution formulation maps from fluctuations in hidden states to response variables, whereas the directed transfer function in the autoregressive formulation $\mathbf{S}(\omega)$ maps directly among different response variables. These representations can be used to generate second-order statistics or measures that summarize the dependencies, as shown in the third row, for example, cross-covariance functions and cross spectra. The normalized or standardized variants of these measures are shown in the lower row and include the cross-correlation function (in time) or coherence (in frequency). The equations show how various representations can be derived from each other. All variables are either vector or matrix functions of time or frequency. For simplicity, the autoregressive formulations are shown in discrete form for the univariate case (the same algebra applies to the multivariate case, but the notation becomes more complicated). Here, $\mathbf{z}(t)$ is a unit normal innovation. Finally, note the Granger causality is only appropriate for a bivariate time series.

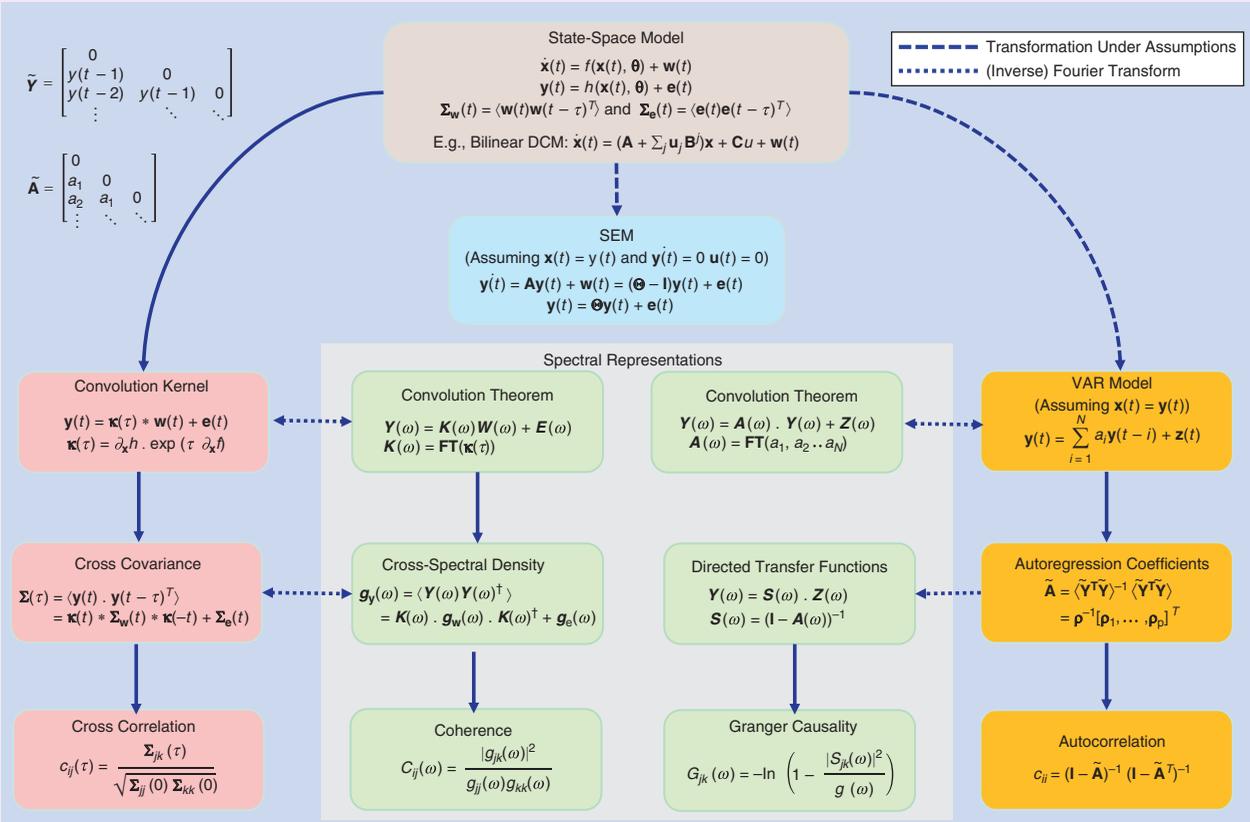


FIGURE S2. The relationship among different formulations of dependencies within a multivariate time series used in fMRI.

there is a distinction between (second-order) effects sizes in the upper row of dependency measures and their standardized equivalents in the lower row (Figure S2). For example, coherence is simply the amplitude of the cross-spectral density normalized by the autospectra of the two regions in question. Similarly, one can think of Granger causality as a standardized measure of the directed transfer function (normalized by the autospectra of the source region).

We also note another widely used measure of functional dependencies known as *mutual information* [57], which quantifies the shared information between two variables and can reflect both linear and nonlinear dependencies. For example, if two time series are independent, there is no shared information, and hence the mutual information is zero. Mutual information can be calculated relatively simply—under the assumption that time series are Gaussian—from coherence in the frequency domain as [58]–[60]

$$\vartheta_{ij} = \frac{1}{2\pi} \int_{\omega_1}^{\omega_2} \log(1 - C_{ij}(\omega)) d\omega, \quad (8)$$

where $C_{ij}(\omega)$ is the coherence (as defined in Figure S2) between the two time series i and j .

In summary, given a state-space model, one can predict or generate the functional connectivity that one would observe in terms of cross-covariance functions, complex cross spectra, or autoregression coefficients (where the latter can be derived in a straightforward way from the former using the Yule–Walker formulation). In principle, this means that one could use either the sampled cross-covariance functions or cross spectra as data features. It would also be possible to use the least-squares estimate of the autoregression coefficients—or, indeed, Granger causality—as data features to estimate the underlying effective connectivity. We describe such schemes in the next section.

Summary

In this section, we have tried to place different analyses of connectivity in relation to each other. The most prevalent approaches to effective connectivity are DCM, SEM, and Granger causality. We highlighted some of the implicit assumptions made when applying SEM and Granger causality to fMRI time series. Next we will focus on generative models of distributed brain responses and consider some of the exciting developments in this field.

Biophysical modeling of neuronal dynamics

Biophysical models of neuronal dynamics are usually used for one of two things: either to understand the emergent properties of neuronal systems or as observation models for measured neuronal responses. We discuss examples of both. In terms of emergent behaviors, we consider dynamics on structure [61]–[69] and how this behavior has been applied to characterizing autonomous or endogenous fluctuations in fMRI [70]–[73].

This section concludes with recent advances in DCM of directed neuronal interactions that support endogenous fluctuations. Some sections below are based on our previous review [10].

Intrinsic dynamics, criticality, and bifurcations

The use of resting-state fMRI [74], [75] or studies based on BOLD signal correlations while the brain is at rest are widespread [76]. These patterns are thought to reflect anatomical connectivity [77] and can be characterized in terms of remarkably reproducible spatial modes (resting-state or intrinsic networks). One of these modes recapitulates the pattern of deactivations observed across a range of activation studies (the default mode [78]). Resting-state fMRI studies show that even at rest, endogenous brain activity is self-organizing and highly

structured. The emerging picture is that endogenous fluctuations are a consequence of dynamics on anatomical connectivity structures with particular scale-invariant characteristics [70], [71], [79], [80]. These are well-studied and universal characteristics of complex systems and suggest that we may be able to understand the brain in terms of universal phenomena [81]. Universality is central to the hypothesis that the cerebral cortex is poised near a critical point where only one variable, a control parameter, determines the macroscopic behavior of the system [82], [83]. This is an important issue because systems near phase transitions show universal phenomena [84]–[88]. Near the critical point, correlations between neurons would occur across all scales, leading to optimized communication [89]. Experimental evidence for this notion has accumulated during the past decades, and power laws and scaling relationships have been found in human neuroimaging time series [90], [91]. However, it should be noted that with more attention on this new direction, there are a variety of distributions (e.g., stretched exponential, Rayleigh, double exponential, and lognormal) that are found in neurophysiological time series [26], [92], [93]. Hence, there may be a need to carefully disambiguate the causes of these heavy-tailed distributions found in the brain and behavior. From the dynamical system perspective, endogenous dynamics are thought to be generated by the dynamic instabilities that occur near bifurcations, that is, dynamics that accompany a loss of stability when certain control parameter(s) reach a critical value [26], [94]–[96]. The eigenmodes of neuronal (effective) connectivity that define the stability of the resting state give rise to scale-free fluctuations that emerge from the superposition of the few modes that decay slowly. These slowly fluctuating (unstable) modes have Lyapunov exponents that are close to zero. This occurs when systems approach transcritical bifurcations (or stochastic Hopf bifurcations when the eigenvalues are complex [97], [98] and show critical slowing [93]). Put simply, this means that the ensuing networks are defined by trajectories that have fixed points close to instability and that the neuronal fluctuations persist over longer timescales to generate the patterns responsible for the emergence of intrinsic brain networks. The

Spectral and stochastic DCMs furnish estimates of the effective connectivity that underlies intrinsic brain networks.

amplitudes of these eigenmodes or patterns correspond to the order parameters described in the “State-Space Modeling and Effective Connectivity” section. The (negative) inverse of the Lyapunov exponent corresponds to the characteristic time constant of each mode, where each mode with a small exponent (large time constant) corresponds to an intrinsic brain network or resting-state network.

Causal modeling of neuronal dynamics

The past decade has seen the introduction of graph theory to brain imaging. Graph theory provides an important formulation for understanding dynamics on structure. Developments in this area have progressed on two fronts: understanding connections between graphs and probability calculus and the use of probabilistic graphs to resolve causal interactions. The probabilistic graph framework goes beyond classical constructs by providing powerful symbolic machinery and notational convenience (e.g., the use of dependency graphs to resolve Simpson’s paradox; see “Simpson–Yule Paradox”).

Within this enterprise, one can differentiate at least two streams of work: one based on Bayesian dependency graphs or graphical models called structural causal modeling [99] and the other based on causal influences over time, which we consider under DCM. Structural causal modeling originated with SEM [47] and uses graphical models (Bayesian dependency graphs or Bayes nets) in which direct causal links are encoded by directed edges. These tools have been largely developed by Pearl [22] and are closely related to the ideas in [100]–[102]. An essential part of network discovery in structural causal modeling is the concept of intervention: eliminating connections in the graph and setting certain nodes to given values. Structural causal modeling lends a powerful and easy-to-use graphical method to show that a particular model specification identifies a causal effect of interest. Moreover, the results derived from structural causal modeling do not require specific distributional or functional assumptions, such as multivariate normality, linear relationships, and so on. However, it is not the most suitable framework to understand coupled dynamical systems because it is limited in certain respects. Crucially, it deals only with conditional independencies in DAGs. This is problematic because brains perform computations on a directed and cyclic graph. Every brain region is connected reciprocally (at least polysynaptically), and every computational theory of brain function rests on some form of reciprocal or reentrant message passing. Another drawback is that the causal calculus of structural causal modeling ignores time. Pearl argued that a causal model should rest on functional relationships between variables. However, these functional relationships cannot deal with (cyclic) feedback loops. Pearl [14] argued for DCMs when attempting to identify hysteresis effects, where causal influences depend on the history of the system. Interestingly, the DAG restriction can be finessed by considering dynamics and temporal precedence within

Neural pathways are flexible, adaptable, connected, and moldable by changes in our environment or by injury or disease.

structural causal modeling. This is because the arrow of time can be used to convert a directed cyclic graph into an acyclic graph when the nodes are deployed over successive time points. This leads to SEM with time-lagged data and related autoregression models, such as those employed by Granger causality described previously. As established in the previous section, these can be regarded as discrete time formulations of DCMs in continuous time.

Structural and dynamic causal modeling

As already established, in relation to the modeling of fMRI time series, DCM refers to the (Bayesian) inversion and comparison of models that cause observed data. These models are usually state-space models expressed as (ordinary, stochastic, or random) differential equations that govern the motion of hidden neurophysiological states. These models are generally equipped with an observer function that maps from hidden states to observed signals [see (1)]. The basic idea behind DCM is to formulate one or more models of how data are caused

in terms of a network of distributed sources. These sources talk to each other through parameterized connections and influence the dynamics of hidden states that are intrinsic to each source. Model inversion provides estimates of their parameters and the model evidence.

We have introduced DCM for fMRI using a simple state-space model based on a bilinear approximation (extensions to, for example, nonlinear [103] and two-state [104] DCM, among others, are also available and are in use) to the underlying equations of motion that couple neuronal states in different brain regions [32]. Most DCMs consider point sources for both fMRI and EEG/MEG data (cf. equivalent current dipoles) and are formally equivalent to the graphical models used in structural causal modeling. However, in DCM, they are used as explicit generative models of observed responses. Inference on the coupling within and between nodes (brain regions) is generally based on perturbing the system and trying to explain the observed responses by inverting the model. This inversion furnishes posterior or conditional probability distributions over unknown parameters (e.g., effective connectivity) and the model evidence for model comparison [105]. The power of the Bayesian model comparison in the context of DCM has become increasingly evident. This now represents one of the most important applications of DCM and allows different hypotheses to be tested, where each DCM corresponds to a specific hypothesis about functional brain architectures [106]–[112]. DCM has been used mostly for (task-based) fMRI and electrophysiological dynamics (EEG/MEG/LFPs), but the most recent advances have focused on the modeling of intrinsic brain networks in the absence of exogenous influence, known as resting-state fMRI [74]. In the remainder of this section, we briefly review these developments and discuss these new mathematical models. We also showcase some of their clinical applications to neurodegenerative diseases, such as Parkinson’s disease.

Dynamic casual modeling of intrinsic networks

There has been an explosion of research examining spontaneous fluctuations in fMRI signals (Figure 2). These fluctuations can be attributed to spontaneous neuronal activity, which is usually ignored in deterministic models of responses to (designed) experimental inputs. Deterministic DCMs are cast as multiple-input, multiple-output systems, in which exogenous inputs perturb the brain to produce an observed BOLD response. In the absence of external inputs, as in the case of resting-state fMRI, neuronal networks are driven by activity that is internal to the system [113]. The generative model for resting-state fMRI time series has the same form as (3) but discounts exogenous modulatory input. It should be noted that we can still include exogenous (or experimental) inputs, $u(t)$, in our model. These inputs drive the hidden states and are usually set to zero in resting-state models. It is perfectly possible to have external (nonmodulatory) stimuli, as in the case of conventional functional neuroimaging studies. For example, in [114] we used an attention-to-visual-motion paradigm to illustrate this point. Figure 3 provides a schematic of the resulting stochastic DCM. In contrast to the previous section, we adopt a generalized framework in which state noise $w(t)$ and observation noise $e(t)$ are analytic (i.e., non-Markovian). This simply means that generalized motion of the state noise $\mathbf{w}(t) = [\mathbf{w}(t), \dot{\mathbf{w}}(t), \ddot{\mathbf{w}}(t) \dots]$ is well defined in terms of its higher-order statistics. Similarly, the observation noise $\tilde{\mathbf{e}}(t) = [\mathbf{e}(t), \dot{\mathbf{e}}(t), \ddot{\mathbf{e}}(t) \dots]$ has a well-defined covariance (for a more detailed discussion, see [115]). Consequently, the stochastic part of the generative model in (1) can be conveniently parameterized in terms of its precision (inverse covariance). This allows us to cast (1) as a random differential equation instead of a stochastic differential equation, hence eschewing Itô calculus [34], [116]. Interested readers will find a theoretical motivation for using analytic state noise in [34]. Under linearity assumptions, (1) can be written compactly in generalized coordinates of motion as

$$\begin{aligned} \mathbf{D}\tilde{\mathbf{x}}(t) &= \tilde{\mathbf{f}}(\tilde{\mathbf{x}}, \tilde{\mathbf{u}}, \boldsymbol{\theta}) + \tilde{\mathbf{w}}(t) \\ \tilde{\mathbf{y}}(t) &= \tilde{\mathbf{h}}(\tilde{\mathbf{x}}, \boldsymbol{\theta}) + \tilde{\mathbf{e}}(t), \end{aligned} \quad (9)$$

where \mathbf{D} is the block diagonal temporal derivative operator, such that the vectors of generalized coordinates of motion are shifted as we move from lower orders of motion to higher orders [115]. For resting-state activity, (9) takes a very simple linear form:

$$\mathbf{D}\tilde{\mathbf{x}}(t) = \mathbf{A}\tilde{\mathbf{x}}(t) + \mathbf{C}\tilde{\mathbf{u}}(t) + \tilde{\mathbf{v}}(t). \quad (10)$$

This is an instance of a linear dynamical system with quasideterministic behavior [117], [118]. Put simply, the linear dynamical system described by (10) is insensitive to the initial conditions. For this reason, it can exhibit only a limited repertoire of behavior: linear systems can contain closed orbits, but they will not be isolated; no limit cycles—either stable or unstable—can exist, which precludes chaotic behavior. Technically speaking, if $\boldsymbol{\lambda}$ represents the eigenvalues of the Jacobian $\partial_{\tilde{\mathbf{x}}}\mathbf{f} = \mathbf{A}$, that is, $\boldsymbol{\lambda} = \mathbf{v}^\dagger \mathbf{A} \mathbf{v}$, where \dagger denotes the

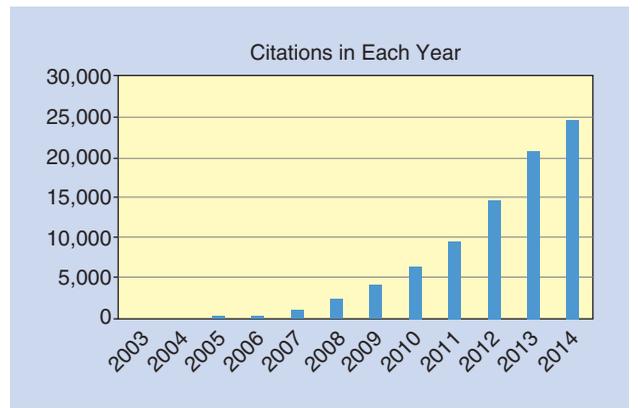


FIGURE 2. Citation rates for resting-state fMRI studies. These citations were identified by searching for “fMRI*” and “resting state.” (Source: Web of Science.)

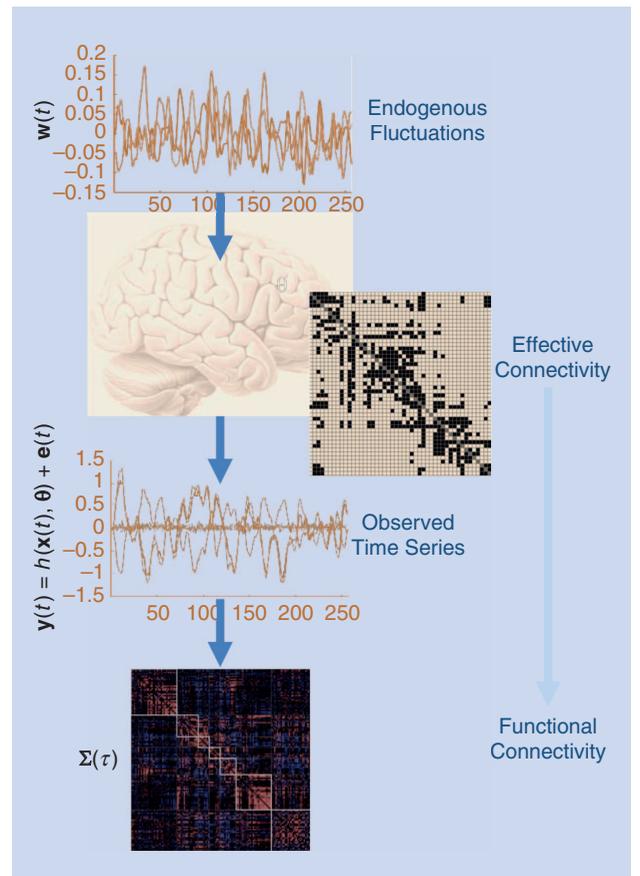


FIGURE 3. This schematic illustrates the forward (dynamic causal) model for modeling intrinsic or endogenous fluctuations. The endogenous fluctuations (state noise) are the driving input to the state-space model of effective connectivity, which is a function of the current neuronal states $\mathbf{x}(t)$ and the connectivity parameters $\boldsymbol{\theta}$ that define the underlying structure or functional architecture of the model and the random fluctuations $\mathbf{w}(t)$. The driving fluctuations cause change in neural activity that can, in turn, be observed using the observer function h after addition of observation noise $\mathbf{e}(t)$. The associated functional connectivity (e.g., cross-covariance function) can be calculated easily from this forward or generative model (see Figure S2 in “Measures of Connectivity”) for any given parameters. Note that the effective connectivity matrix shown is actually a structural connectivity matrix of the famous macaque/CoCoMac. We use it here as a schematic for effective connectivity.

generalized inverse, then the Lyapunov exponents $\Re(\lambda)$ of this linear dynamical system will always be negative. In general, the Jacobian is not symmetrical (causal effects are asymmetric); the modes and eigenvalues take complex values. See [119] for a detailed treatment of the special case of symmetrical connectivity, in which the eigenmodes of functional and effective connectivity become the same. It is worth noting that these eigenmodes are also closely related to (group) independent component analysis (ICA) except with a rotation based on higher-order statistics (for details, see [120]).

There are currently two schemes to invert models of the form (9). They differ in what data features they use for the parameter estimation. The first inverts data in the time domain, and the model is used to predict the time series per se.

This is referred to as *stochastic DCM* [116]. The second approach makes predictions in the frequency domain and is based on fitting second-order data features, such as cross spectra. This is referred to as *spectral DCM* [114], [121]. We briefly review both schemes and illustrate their clinical applications. For a schematic illustration of DCM of intrinsic dynamics, see Figure 4. Figure 5 presents a comparison of the two schemes.

Stochastic dynamic causal models

Stochastic DCM entails inverting a model of the form given by (10) in the time domain, which includes state noise. This requires estimation of not only the model parameters (and any hyperparameters that parameterize the precision of generalized

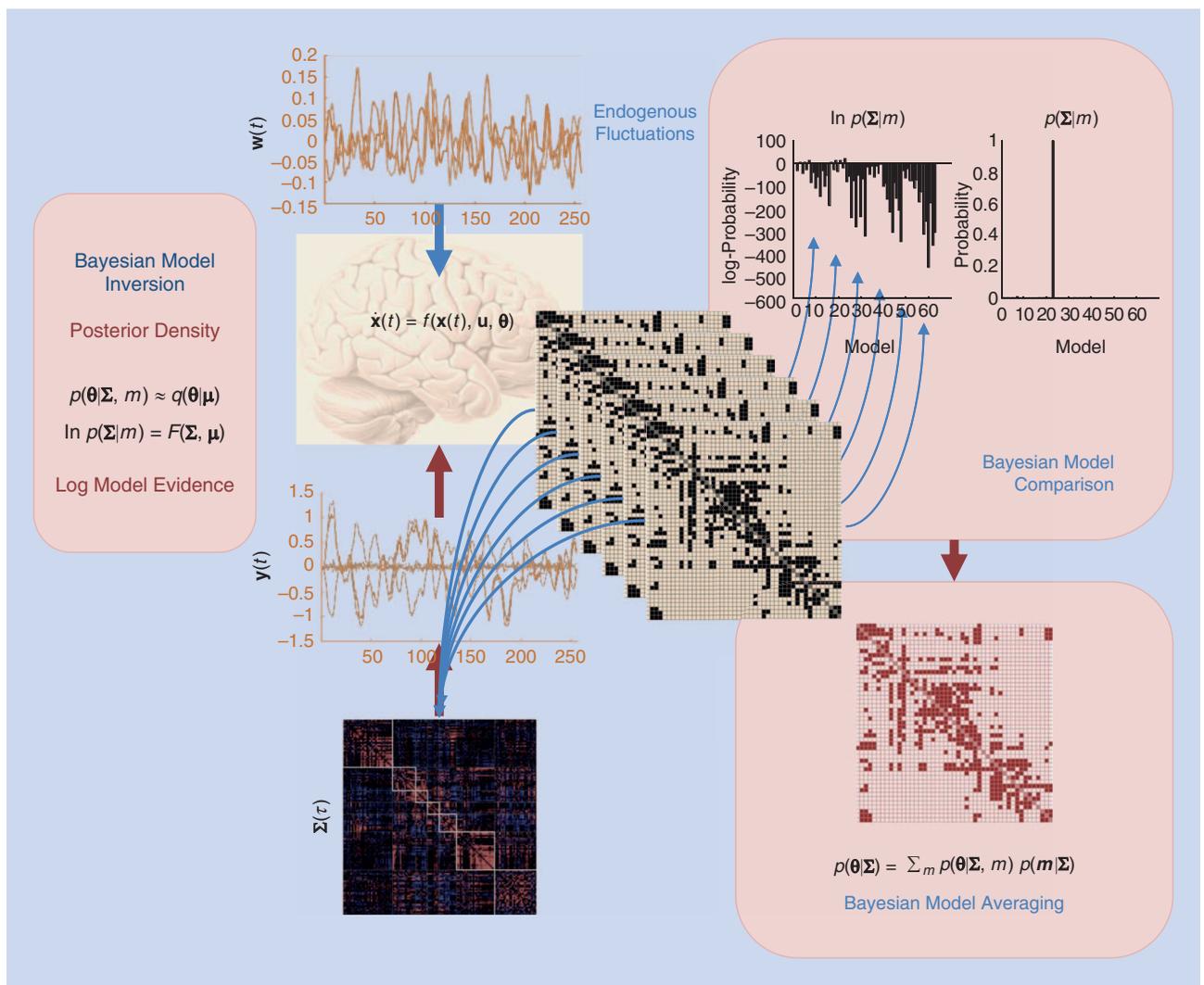


FIGURE 4. This schematic shows a DCM that embodies the best effective connectivity—identified using Bayesian model inversion (top left panel)—among hidden neuronal states that explains the observed functional connectivity, $\Sigma(t)$, among hemodynamic responses. This explanation is possible because the cross spectra contain all the information about (second-order) statistical dependencies among regional dynamics. Bayesian model inversion furnishes posterior estimates for the parameters of each model and provides the associated log model evidence in terms of a variational free-energy bound. Because the mapping from functional connectivity to effective connectivity is not objective (there may be many combinations of effective connectivity parameters that induce the same functional connectivity), one can use a Bayesian model comparison (top right panel) to score competing models. The model with the highest model evidence can then be selected. Alternatively, one can use Bayesian model averaging to average all possible models (bottom right panel).

random fluctuations) but also the hidden states, which become random (probabilistic) variables. The unknown quantities to be estimated under a stochastic DCM are $\boldsymbol{\psi} = \{\tilde{\mathbf{x}}(t), \boldsymbol{\theta}, \boldsymbol{\sigma}\}$, where $\boldsymbol{\sigma}$ refers to any hyperparameters describing random fluctuations. In terms of temporal characteristics, the hidden states are time-variant, whereas the model parameters (and hyperparameters) are time-invariant.

There are various variational schemes in the literature that can invert such models, for example, dynamic expectation maximization (DEM) [122] and generalized filtering (GF) [34]. There is a subtle but important distinction between DEM and GF. DEM calls on the mean field approximation described above, that is, it assumes $q(\boldsymbol{\psi}) = q(\tilde{\mathbf{x}}(t))q(\boldsymbol{\theta})q(\boldsymbol{\sigma})$, whereas GF, as the name suggests, is more general in that it does not make this assumption. However, both schemes assume a fixed-form Gaussian distribution for the approximate conditional

posterior densities (the Laplace approximation). GF considers all unknown quantities to be conditionally dependent variables, that is, $q(\boldsymbol{\psi}) = q(\tilde{\mathbf{x}}, \boldsymbol{\theta}, \boldsymbol{\sigma})$, and produces time-dependent conditional densities for all unknown quantities. The time-invariant parameters and hyperparameters are cast as time-variant with the prior constraint that their temporal variation is small. In brief, this online scheme assimilates log evidence at each time point in the form of variational free energy and provides time-dependent conditional densities for all unknown quantities. This is in contrast to schemes such as DEM (or deterministic model inversion using variational Laplace) with mean field approximations, which assimilates all the data before computing the free energy.

“Effective Connectivity in Parkinson’s Disease” shows an exemplar data analysis reported in [123] that used stochastic DCM to quantify effective connectivity changes in

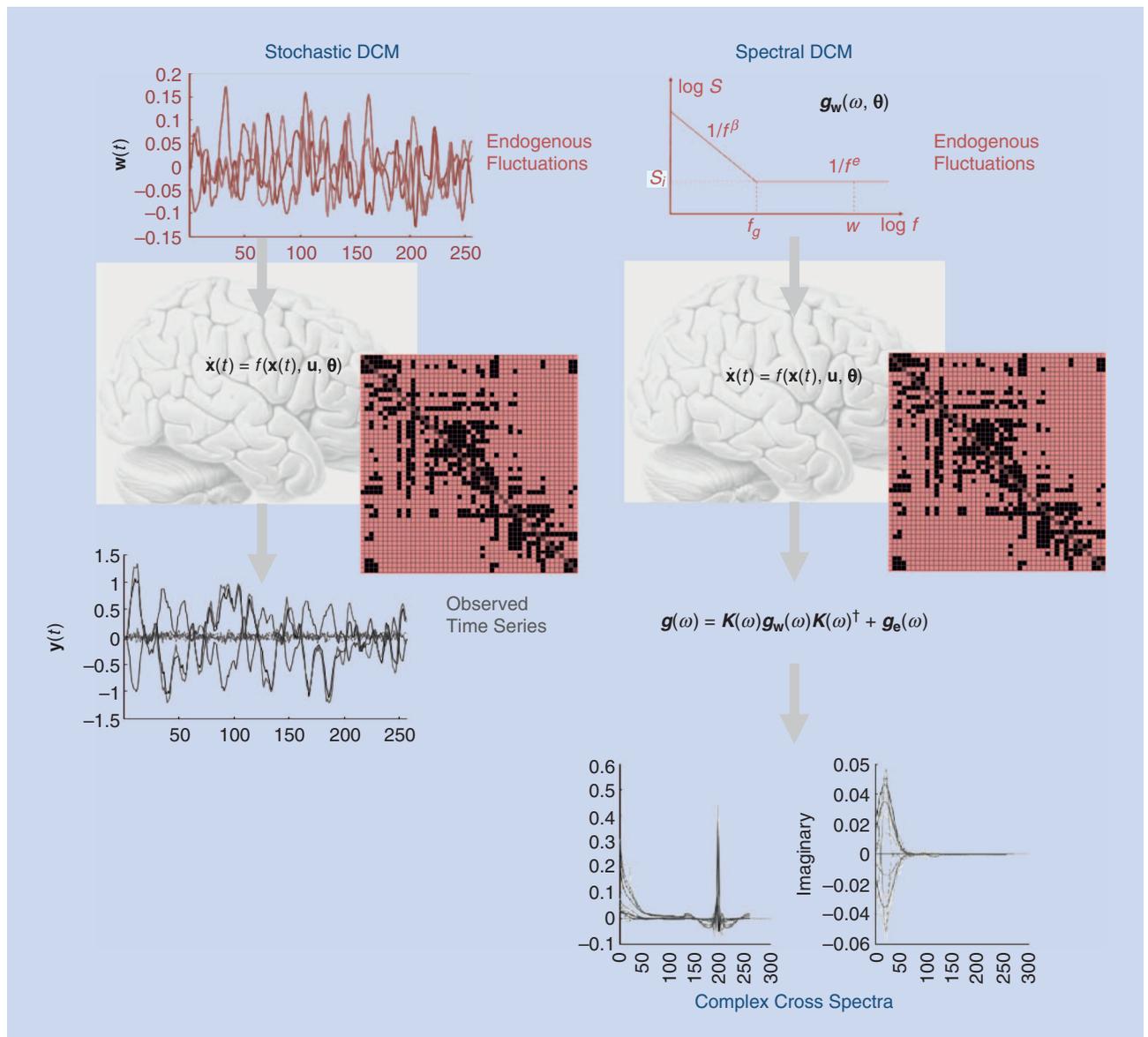


FIGURE 5. A schematic illustrating the distinction between stochastic and spectral DCM. See the “Biophysical Modeling of Neuronal Dynamics” section for a detailed description of how these schemes are used to model intrinsic network dynamics.

Effective Connectivity in Parkinson's Disease

Given the marked clinical effect of subthalamic nucleus (STN) deep brain stimulation (DBS) in patients with Parkinson's disease, Kahan et al. [123] used stochastic dynamic causal modeling (DCM) to estimate the coupling between key nodes of the basal ganglia network and to study whether this coupling was changed by DBS. In Figure S3(a), a network was specified based on human and animal literature, and priors were placed on the nature of the coupling (excitatory or inhibitory) based on the neurochemical systems known to mediate neuronal connections. The literature-based anatomical model of the motor cortico-striato-thalamic loop was further simplified by removing the pallidal nodes and summarizing polysynaptic connections [thick arrows joining the putamen (Put), STN, and thalamus (Tha)]. Red arrows indicate excitatory coupling, and blue arrows indicate inhibitory coupling.

Placing priors on the direction of coupling was enabled using the two-state DCM (left). In (b), it is shown that model inversion yielded coupling parameters on and off DBS, demonstrating significant DBS-related changes in extrinsic (between-node) coupling throughout the network. Paired t-tests revealed significant differences between extrinsic coupling on and off stimulation. Corticostriatal, direct pathway, and thalamocortical connections were potentiated by DBS, whereas STN afferents (lower panel) and efferents (upper panel) were attenuated. Note the difference in scale between the upper and lower panels. This is because the STN was modeled as a hidden node that was not measured with fMRI. Using a series of regression models, (c) shows the modulatory effects of DBS on connectivity to predict the clinical improvements seen in the patient cohort. (See [123] for more details.)

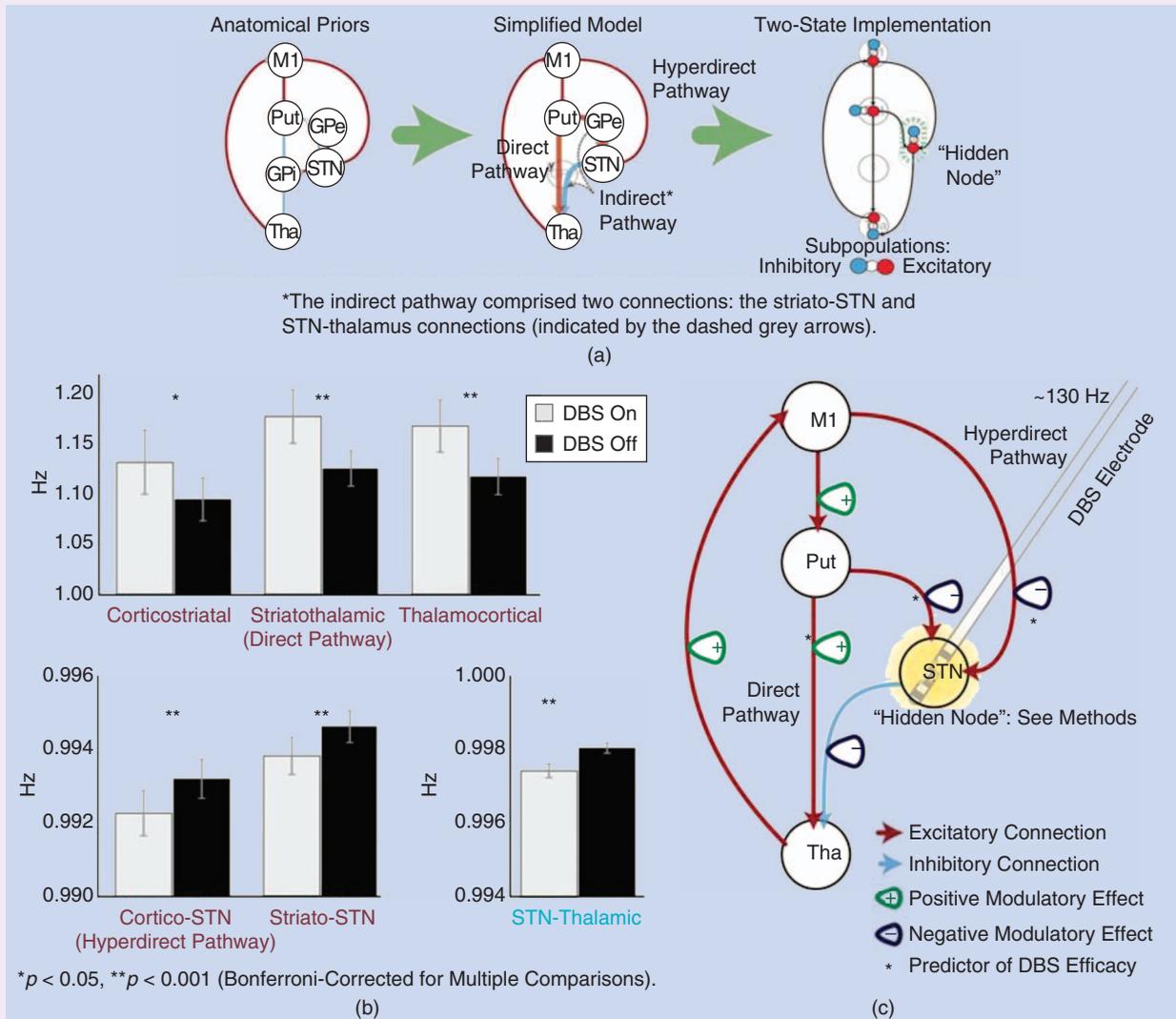


FIGURE S3. The summary of data analysis reported in [123] using stochastic DCM in Parkinson's disease under DBS.

Parkinson's disease. Depleted of dopamine, the dynamics of the Parkinsonian brain impact on both action and resting motor activity. Deep brain stimulation (DBS) has become an established means of managing these symptoms, although its mechanisms of action remain unclear. Using stochastic DCM, Kahan et al. [123] modeled the effective connectivity underlying low-frequency BOLD fluctuations in the resting Parkinsonian motor network. They were particularly interested in the distributed effects of DBS on cortico-subcortical connections. Specifically, they showed (see Figure S3 in "Effective Connectivity in Parkinson's Disease") that subthalamic nucleus (STN) DBS modulates all major components of the motor cortico-striato-thalamo-cortical loop, including the corticostriatal, thalamocortical, direct, and indirect basal ganglia pathways and the hyperdirect STN projections. The strength of effective STN afferents and efferents was reduced by stimulation, whereas corticostriatal, thalamocortical, and direct pathways were strengthened. Remarkably, regression analysis revealed that the hyperdirect, direct, and basal ganglia afferents to the STN predicted clinical status and therapeutic response to DBS; however, suppression of the sensitivity of the STN to its hyperdirect afferents by DBS may subvert the clinical efficacy of DBS. These findings highlight the distributed effects of stimulation on the resting motor network and provide a framework for analyzing effective connectivity in resting-state functional MRI with strong a priori hypotheses.

Spectral and stochastic DCMs furnish estimates of the effective connectivity that underlies intrinsic brain networks.

Spectral dynamic causal models

Although the stochastic models in (10) and their inversion in the time domain provide a useful means to estimate effective connectivity, they also entail the estimation of hidden states. This poses a difficult inverse problem that is computationally demanding, especially when the number of hidden states becomes large. To finesse this problem, a DCM based on a deterministic model that generates predicted cross spectra was explored [114], [121]. This scheme provides a constrained inversion of the stochastic model by parameterizing the spectral density neuronal fluctuations. This parameterization also provides an opportunity to compare parameters encoding neuronal fluctuations among groups. The parameterization of endogenous fluctuations means that the states are no longer probabilistic; hence, the inversion scheme is significantly simpler, requiring estimation of only the parameters (and hyperparameters) of the model. The ensuing model inversion in the spectral domain is similar in spirit to previous approaches described in [26], [98], and [124]. Put simply, although GF estimates time-dependent fluctuations in neuronal states producing observed data, spectral DCM simply estimates the time-invariant parameters of their cross spectra. Effectively, this is achieved by replacing the original time series with their second-order statistics (i.e., cross spectra). This means that instead of

estimating time-varying hidden states, we are estimating their covariance. In turn, this means that we need to estimate the covariance of the random fluctuations using a scale-free (power law) form for the state noise (resp. observation noise) that can be motivated from previous work on neuronal activity [125]–[127]:

$$\begin{aligned} \mathbf{g}_w(\omega, \boldsymbol{\theta}) &= \boldsymbol{\alpha}_w \omega^{-\beta_w} \\ \mathbf{g}_e(\omega, \boldsymbol{\theta}) &= \boldsymbol{\alpha}_e \omega^{-\beta_e}. \end{aligned} \quad (11)$$

Here, $\mathbf{g}_x(\omega) = X(\omega)X(\omega)^\dagger$ represents the complex cross spectra, where $X(\omega)$ is the Fourier transform of $\mathbf{x}(t)$, $\{\boldsymbol{\alpha}, \boldsymbol{\beta}\} \subset \boldsymbol{\theta}$ are the parameters controlling the amplitudes and exponents of the spectral density of the neural fluctuations, and $\omega = 2\pi f$ is the angular frequency. This models neuronal noise with generic $1/f^\gamma$ spectra that characterize fluctuations in systems that are at nonequilibrium steady state. A linear scaling regime of the spectral density in double logarithmic coordinates—implicit in (11)—is not by itself indicative of a scale-free critical process unless γ is less than 1.5 Hz (and the regime scales over

several orders of magnitude). For the human EEG, this is generally not the case: above 10 Hz, $\gamma = 2.5$, and above 70 Hz, γ is usually greater than 3.5, which is consistent with a Poisson process (see [128] and [129]). However, at low frequencies (less than 1.5 Hz), the slope is more shallow, and it is likely that the amplitude or power envelopes of faster frequencies are scale-free [130, 131] or another heavy-tailed distribution [132]. Using the model parameters $\boldsymbol{\theta} \supseteq \{\mathbf{A}, \mathbf{C}, \boldsymbol{\alpha}, \boldsymbol{\beta}\}$, one can simply generate the expected cross spectra as follows:

$$\begin{aligned} \mathbf{y}(t) &= \boldsymbol{\kappa}(t) \otimes \mathbf{w}(t) + \mathbf{e}(t) \\ \boldsymbol{\kappa}(t) &= \partial_x g \exp(t \partial_x f) \\ \mathbf{g}_y(\omega, \boldsymbol{\theta}) &= |\mathbf{K}(\omega)|^2 \mathbf{g}_w(\omega, \boldsymbol{\theta}) + \mathbf{g}_e(\omega, \boldsymbol{\theta}), \end{aligned} \quad (12)$$

where $\mathbf{K}(\omega)$ is the Fourier transform of the system's (first-order) Volterra kernels $\boldsymbol{\kappa}(t)$, which are a function of the Jacobian or effective connectivity (see Figure S2). The unknown quantities $\boldsymbol{\psi} = \{\boldsymbol{\varphi}, \boldsymbol{\theta}, \boldsymbol{\sigma}\}$ of this deterministic model can now be estimated using standard variational Laplace [133]. The resulting inversion provides the free energy bound on the log-evidence $\log p(\mathbf{g}_y(\omega)|m)$ and approximate conditional densities $q(\boldsymbol{\psi}) \approx p(\boldsymbol{\psi}|\mathbf{g}(\omega), m)$. Here $\mathbf{g}_y(\omega)$ represents the predicted cross spectra that can be estimated, for example, using an autoregressive model.

An example from aging

Finally, in "Aging and Spectral DCM," we show an example from recent work on aging [134] that used spectral DCM. Well-being across the lifespan depends on the preservation of cognitive function. It was hypothesized that successful cognitive aging is determined by the connectivity within and between large-scale brain networks at rest. Spectral DCM

Aging and Spectral DCM

In Figure S4(a), spatial distribution of three independent components using group independent component analysis (ICA) ($n = 602$) are identified as the default-mode network (DMN) (in blue), the dorsal attention network (DAN) (in red), the salience network (SN) (in yellow), and the peaks of their corresponding nodes (green circles). Temporal correlation is between the first eigenvariates of the ensuing time series across all nodes and networks. Coefficients for how well effective connectivity (white), neuronal (green), and hemodynamic (red) parameters predict age are shown in (b), and dynamic causal modeling parameters with bars (95% confidence intervals) that exclude zero are considered as significant predictors. A between-network canonical variate analysis is shown in (c). More specifically, shown is a heliograph of variate

loadings for the first canonical variate, where the relative sizes of correlations are indicated by the relative length of the bars (the dark is positive, and the white is negative). These reflect the statistical relationship between variables of effective connectivity (connectivity profile) and cognitive performance (cognitive profile) ($r = 0.440$, $p < 0.001$). Variables with low contribution ($r < 0.3$) are shown as bars with a dashed outline. Half-maximum strength of a correlation is indicated by dashed rings (outer, $r = +0.5$; inner, $r = -0.5$). The corresponding bivariate canonical correlations for three age groups are shown in (d). The relationships between connectivity and cognitive profiles are more pronounced for older patients, suggesting that performance in older adults reflects a preserved connectivity.

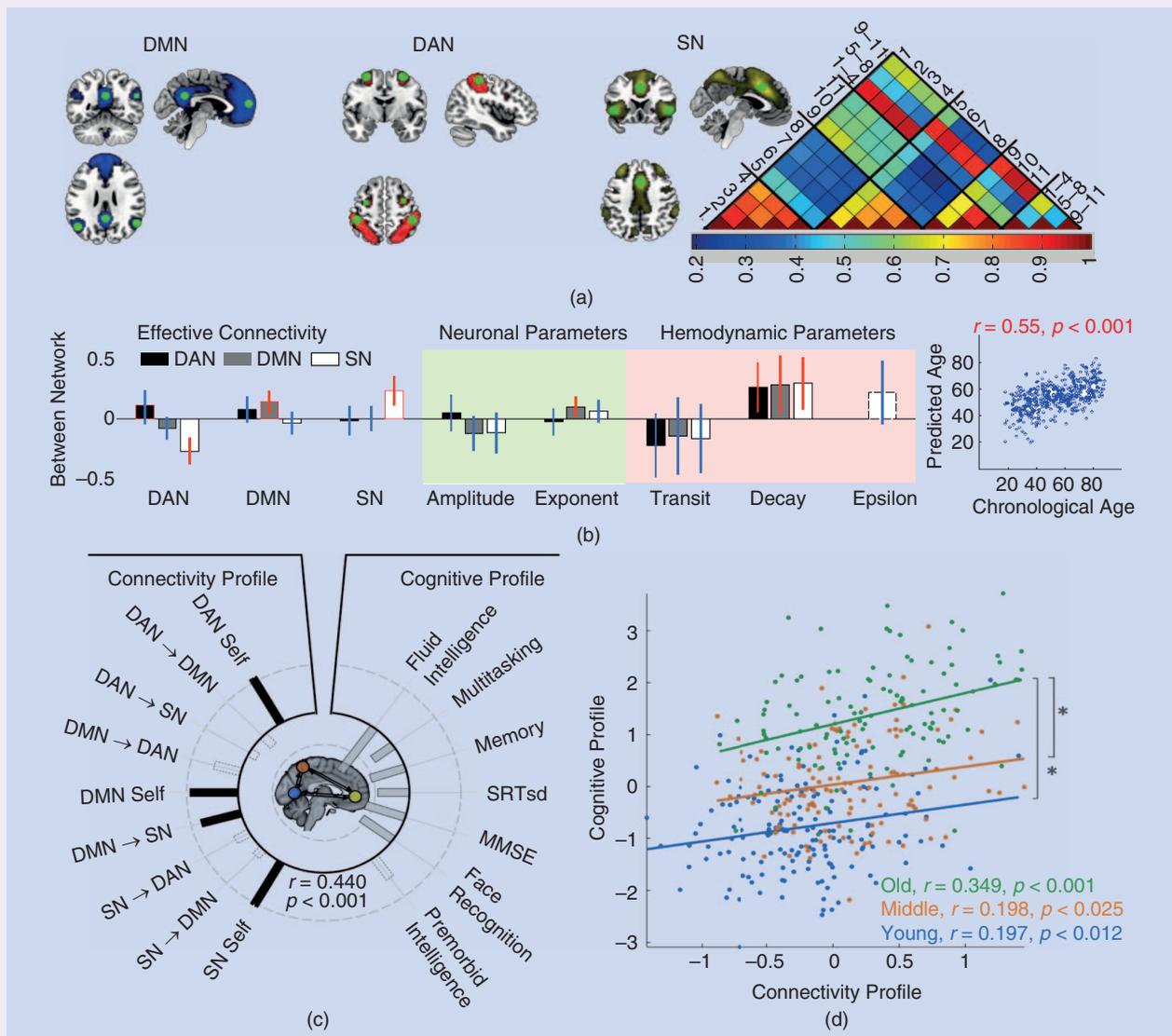


FIGURE S4. The summary of between-network connectivity changes over the adult lifespan.

was used to explain the spectral characteristics of resting-state fMRI data from 602 healthy adults in a cohort across ages 18–88 (www.cam-can.org). The location of the key cortical regions in each network was identified by spatial ICA using group ICA [120] to extract 20 low-dimensional components. The three well-established functional networks, the salience network (SN), dorsal attention network (DAN), and default-mode network (DMN), were then identified by spatial matching to preexisting templates [135]. Effective connectivity was assessed within and between these three key large-scale networks although, for brevity, we have included more interesting results only for the between-network connectivity in this review. In brief, a two-step process is used in which ICA identifies linearly coherent networks, and the (potentially nonlinear) relationship among these networks is then tested within a causal modeling framework using spectral DCM. This approach has been used several times in both task-based and resting-state fMRI data [136]–[138].

Using multiple linear regression, it was found that about 30% of age variance can be predicted ($r = 0.544$, $p < 0.001$) by 1) increased inhibitory self-connections in SN and DMN, 2) decreased effective connectivity from DAN to SN, and 3) increased hemodynamic decay times for all networks [Figure S4(b)]. Subsequently, a classical multivariate test (canonical variate analysis) was used to determine to what degree the DCM parameters predict cognitive performance, shown in Figure S4(c). For between-network analysis, the corresponding canonical vector suggested that high performance across a range of cognitive tasks [high scores of general intelligence (Cattell), face processing (Benton Faces), memory (story recall), multitasking (Hotel), and response consistency (inverse of response variability on simple motor task)] was associated with less self-inhibition of the networks and a smaller influence of the DMN on SN ($r = 0.447$, $p < 0.001$). In other words, about 20% of the variance in performance—across a range of cognitive tasks studied—could be predicted from changes in effective connectivity between networks. To further investigate whether the relationship between cognitive performance and connectivity was age-dependent, moderation analysis was used. It was found that the interaction between age and connectivity values (age \times connectivity profile) predicted a significant proportion of variance in cognitive performance ($T(398) = 3.115$, p (one-tailed) < 0.001). The direction of the interaction was such that increasing age strengthened the relationship between cognitive and connectivity profiles. This is shown in Figure S4(d), where the relationship between cognitive performance and connectivity profile becomes stronger for older age groups. This is an interesting study because it used spectral DCM to dissociate neuronal from vascular components of the fMRI signal to find age-dependent and behaviorally relevant differences in resting-state effective connectivity between large-scale brain networks. Taken

Resting-state fMRI studies show that intrinsic brain activity is self-organizing and highly structured.

together, the results suggest that maintaining a healthy resting-state connectivity becomes increasingly important for older adults to maintain high levels of domain-general cognitive function and may play a critical role in the mechanisms of healthy cognitive aging.

Summary

Both spectral and stochastic DCMs furnish estimates of the effective connectivity that underlies intrinsic brain networks.

These estimates are based on BOLD data acquired at rest using different inversion schemes. We suppose that these resting-state networks emerge from the dynamic instabilities and critical slowing near transcritical bifurcations. In this setting, neuronal activity is modeled with random differential equations, which can be estimated using stochastic inversion schemes

(such as GF in stochastic DCM) or by deterministic schemes modeling observed functional connectivity (specifically, the cross-spectral densities modeled by spectral DCM).

Discussion

The limitations and challenges of DCM and the implicit scoring of large numbers of models have been addressed in a number of critical reviews (e.g., [139] and [140]). Their key conclusions highlight several issues. First, although the modeling assumptions underlying DCM are motivated by neuroanatomical and neurophysiological constraints, their plausibility is difficult to fully establish. For example, in DCM for fMRI, physiological details of the neurovascular coupling are potentially important. Many DCMs neglect the potential influence of inhibitory activity on the hemodynamic response and call on a simplistic account of the metabolic cascade that relates synaptic activity and neuronal metabolism to the vasodilatation. In principle, these are issues that can be resolved using Bayesian model comparison. In other words, if a more complex and complete model is supported by the data, one can always optimize the DCM. Examples of this include recent trends toward more detailed physiological modeling. For example, several extensions are proposed in [141], such as an adaptive two-state neuronal model that accounts for a wide range of neuronal time courses during stimulation and post-stimulus deactivation, a neurovascular coupling model that links neuronal activity to blood flow in a strictly feedforward fashion, and a balloon model that can account for a vascular uncoupling between blood flow and blood volume due to viscoelastic properties of venous blood vessels.

There are also questions about the robustness of the statistical (approximate Bayesian) inference techniques employed in DCM. For example, it has been argued that 1) the number of parameters and the complexity of the models preclude robust parameter estimation [140], [142]; 2) Bayesian model comparison cannot compare DCMs in the sense that it cannot falsify them; and 3) selecting a model based on the model evidence does not ensure that it will generalize. All of these concerns

stem from frequentist thinking and are dissolved within a Bayesian framework (see [139] for a detailed discussion). There are also several well-founded technical concerns about the variational Bayes (VB) schemes employed in DCM. For example, the objective function based on the free-energy functional is prone to local maxima that can result in inconsistent parameter estimations and model comparisons (e.g., across trials or subjects). There are several experimental studies (e.g., [143]–[147]) that have addressed the reproducibility of DCM and provide reassuring experimental validation. There is an issue of overconfidence usually associated with VB schemes due to the potentially biased inference that results from mean field and Laplace approximations to the posterior density. This issue has been addressed by simulation studies that compare the results of VB with standard (e.g., Gibb's) sampling methods. The failures of approximate Bayesian inference are usually mitigated by formulating the inversion problem in a way that eschews brittle nonlinearities.

Given these issues, one obvious alternative is to use either exact inference schemes, such as Markov chain Monte Carlo (MCMC) or nonparametric methods based on Gaussian processes. Both have recently been explored for inverting Bayesian hierarchical models. For example, Gaussian process optimization was used for model inversion in [148], several gradient-free MCMC schemes (e.g., for random walk-based Hasting's sampling, adaptive MCMC sampling, and population-based MCMC sampling) were explored in [149], and more robust gradient-based MCMC schemes (e.g., for Hamiltonian and Langevin MCMC sampling) were extensively studied in [150]. However, these alternative and promising inference methods are still in an early phase of development and validation phase and will require exhaustive experimental studies to establish validity.

Clearly, most of these issues transcend DCM per se and speak to the challenges facing any modeling initiative that has to contend with big data and a large model or hypothesis space. These challenges have focused recent research on contextualizing the inversion of models of single subjects using (empirical or hierarchical) Bayesian models that distinguish between within- and between-subject effects on one hand and the scoring of large model spaces with techniques such as Bayesian model reduction on the other. This is an active research field with developments nearly every month.

In conclusion, we have used several distinctions to review the history and modeling of macroscopic brain connectivity. We started with the distinction between functional segregation and integration. In functional integration, we considered the key distinction between functional and effective connectivity and their relationship to underlying models of distributed processing. In effective connectivity, we looked at structural and dynamic causal modeling while highlighting recent advances in the DCM of resting-state fMRI data.

We close with a few words on recent large-scale projects in neurosciences, for example, the American BRAIN Initiative and the European Human Brain Project. These initiatives reflect an increasing appreciation of the impor-

tance of neuroscience and the challenges of understanding how brains work. Furthermore, they represent initiatives that exploit remarkable advances in computer science and neuroimaging at many different scales (from the molecular to multisubject) and the modeling (and mining) of the resulting data. The experience of the systems neuroscience community, with the big data obtained from neuroimaging, is reflected in this review. This experience highlights the importance of formal models of how data are generated and the computational schemes used to evaluate and invert these models. We are just embarking on a difficult journey to uncover the governing principles of how brains work and their functional (computational) architectures. Perhaps it is fitting to end with an encouraging quote from Abdus Salam (recipient of the Nobel Prize in Physics 1979): "Nature is not economical of structures—only of principles."

Acknowledgments

This work was funded by the Wellcome Trust and the CHDI Foundation. We thank our colleagues at the Wellcome Trust Centre for Neuroimaging, on whose work this article is based. We acknowledge very useful comments from reviewers that helped improve the readability of this article.

Authors

Adeel Razi (adeel.razi@ieee.org) received his B.E. degree in electrical engineering from the NED University of Engineering and Technology, Karachi, Pakistan; his M.S. degree in communications engineering from the University of Technology Aachen (RWTH), Germany; and his Ph.D. degree in electrical engineering from the University of New South Wales, Sydney, Australia, in 2012. Upon the completion of his undergraduate studies, he joined the Department of Electronic Engineering at the NED University of Engineering and Technology, where he is an associate professor. He is currently on leave from NED University to work at the Wellcome Trust Centre for Neuroimaging, University College London, United Kingdom. His general research interest is in demystifying (complex) networks. He currently works on modeling the changes in connectivity patterns of (large) human brain networks.

Karl J. Friston (k.friston@ucl.ac.uk) is a theoretical neuroscientist and authority on brain imaging. He invented statistical parametric mapping, voxel-based morphometry, and dynamic causal modeling. He received the first Young Investigators Award in Human Brain Mapping (1996) and was elected a fellow of the Academy of Medical Sciences (1999). In 2000, he was president of the international Organization of Human Brain Mapping. He was awarded the Minerva Golden Brain Award in 2003 and was elected a fellow of the Royal Society in 2006. In 2008, he received a medal from the Collège de France, and, in 2011, he received an honorary doctorate from the University of York. He became a fellow of the Society of Biology in 2012, received the Weldon Memorial Prize and Medal in 2013, and was elected as a member of Excellence in the Life Sciences in 2014.

References

- [1] D. George and J. Hawkins, "Towards a mathematical theory of cortical micro-circuits," *PLoS Comput. Biol.*, vol. 5, no. 10, e1000532, Oct. 2009.
- [2] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.
- [3] P. Bak, C. Tang, and K. Wiesenfeld, "Self-organized criticality: An explanation of $1/f$ noise," *Phys. Rev. Lett.*, vol. 59, no. 4, pp. 381–384, July 1987.
- [4] E. Tognoli and J. A. S. Kelso, "The metastable brain," *Neuron*, vol. 81, no. 1, pp. 35–48, Jan. 2014.
- [5] A. Aertsen and F. Preissl, "Dynamics of activity and connectivity in physiological neuronal networks," in *Nonlinear Dynamics and Neuronal Networks*, H. Schuster, Ed. Weinheim, Germany: VCH Verlag, 1991, pp. 281–301.
- [6] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, et al., "Dynamic functional connectivity: Promise, issues, and interpretations," *NeuroImage*, vol. 80, pp. 360–378, Oct. 2013.
- [7] G. K. Cooray, B. Sengupta, P. Douglas, and K. Friston, "Dynamic causal modeling of electrographic seizure activity using Bayesian belief updating," *NeuroImage*, vol. 125, pp. 1142–1154, Jan. 2016.
- [8] G. K. Cooray, B. Sengupta, P. Douglas, M. Englund, R. Wickstrom, and K. Friston, "Characterising seizures in anti-NMDA-receptor encephalitis with dynamic causal modelling," *NeuroImage*, vol. 118, pp. 508–519, May 2015.
- [9] P. A. Robinson, "Interrelating anatomical, effective, and functional brain connectivity using propagators and neural field theory," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 85, no. 1, 011912, Jan. 2012.
- [10] K. J. Friston, "Functional and effective connectivity: A review," *Brain Connect.*, vol. 1, no. 1, pp. 13–36, 2011.
- [11] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 4, pp. 623–656, 1948.
- [12] Y. Shoham, *Reasoning About Change: Time and Causation from the Standpoint of Artificial Intelligence*. Cambridge, MA: MIT Press, 1988.
- [13] P. Suppes, *A Probabilistic Theory of Causality*. Amsterdam: North-Holland Publishing, 1970.
- [14] J. Pearl, *Causality: Models, Reasoning and Inference*. Cambridge, U.K.: Cambridge Univ. Press, 2009.
- [15] G. U. Yule, "Notes on the theory of association of attributes in statistics," *Biometrika*, vol. 2, no. 2, pp. 121–134, 1903.
- [16] K. Pearson, A. Lee, and L. Bramley-Moore, "Genetic (reproductive) selection: Inheritance of fertility of man," *Philos. Trans. R. Soc. London A, Math. Phys. Sci.*, vol. 73, pp. 534–539, Jan. 1899.
- [17] E. H. Simpson, "The interpretation of interaction in contingency tables," *J. R. Stat. Soc. Series B Stat. Method.*, vol. 13, no. 2, pp. 238–241, 1951.
- [18] T. W. Armistead, "Resurrecting the third variable: A critique of Pearl's causal analysis of Simpson's paradox," *Am. Stat.*, vol. 68, no. 1, pp. 1–7, Feb. 2014.
- [19] J. Pearl, "Comment: Understanding Simpson's paradox," *Am. Stat.*, vol. 68, no. 1, pp. 8–13, Feb. 2014.
- [20] R. Christensen, "Comment," *Am. Stat.*, vol. 68, no. 1, pp. 13–17, Jan. 2014.
- [21] P. R. Rosenbaum and D. B. Rubin, "The central role of the propensity score in observational studies for causal effects," *Biometrika*, vol. 70, no. 1, pp. 41–55, 1983.
- [22] J. Pearl, "Causal diagrams for empirical research," *Biometrika*, vol. 82, no. 4, pp. 669–710, Dec. 1995.
- [23] J. Pearl, "The new challenge: From a century of statistics to the age of causation," *Comput. Sci. Stat.*, vol. 29, no. 2, pp. 415–423, 1997.
- [24] R. E. Kalman, "A new approach to linear filtering and prediction problems," *J. Basic Eng.*, vol. 82, no. 1, pp. 35–45, 1960.
- [25] C. Anteneodo and R. Riera, "Additive-multiplicative stochastic models of financial mean-reverting processes," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 72, no. 2, 026106, Aug. 2005.
- [26] F. Freyer, J. A. Roberts, P. Ritter, and M. Breakspear, "A canonical model of multistability and scale-invariance in biological systems," *PLoS Comput. Biol.*, vol. 8, no. 8, e1002634, 2012.
- [27] K. J. Friston, B. Li, J. Daunizeau, and K. Stephan, "Network discovery with DCM," *NeuroImage*, vol. 56, no. 3, pp. 1202–1221, June 2011.
- [28] J. Carr, *Applications of Centre Manifold Theory*. Berlin: Springer-Verlag, 1981.
- [29] H. Haken, *Synergetics: An Introduction. Non-Equilibrium Phase Transition and Self-Organisation in Physics, Chemistry and Biology*. Berlin: Springer-Verlag, 1983.
- [30] V. L. Ginzburg and L. D. Landau, "On the theory of superconductivity," *Zh. Eksp. Teor. Fiz.*, vol. 20, pp. 1064–1082, 1950.
- [31] P. A. Valdes, J. C. Jimenez, J. Riera, R. Biscay, and T. Ozaki, "Nonlinear EEG analysis based on a neural mass model," *Biol. Cybern.*, vol. 81, nos. 5–6, pp. 415–424, Nov. 1999.
- [32] K. J. Friston, L. Harrison, and W. Penny, "Dynamic causal modelling," *NeuroImage*, vol. 19, no. 4, pp. 1273–1302, Aug. 2003.
- [33] K. J. Friston, "Variational filtering," *NeuroImage*, vol. 41, no. 3, pp. 747–766, July 2008.
- [34] K. J. Friston, K. Stephan, B. J. Li, and J. Daunizeau, "Generalised filtering," *Math. Probl. Eng.*, vol. 2010, 621670, 2010.
- [35] R. B. Buxton, E. C. Wong, and L. R. Frank, "Dynamics of blood flow and oxygenation changes during brain activation: The balloon model," *Magn. Reson. Med.*, vol. 39, no. 6, pp. 855–864, June 1998.
- [36] K. E. Stephan, N. Weiskopf, P. M. Drysdale, P. A. Robinson, and K. J. Friston, "Comparing hemodynamic models with DCM," *NeuroImage*, vol. 38, no. 3, pp. 387–401, Nov. 2007.
- [37] R. E. Kass and A. E. Raftery, "Bayes factors," *J. Amer. Statist. Assoc.*, vol. 90, no. 430, pp. 773–795, June 1995.
- [38] R. Goebel, A. Roebroeck, D. S. Kim, and E. Formisano, "Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping," *Magn. Reson. Imaging*, vol. 21, no. 10, pp. 1251–1261, Dec. 2003.
- [39] L. Harrison, W. D. Penny, and K. Friston, "Multivariate autoregressive modeling of fMRI time series," *NeuroImage*, vol. 19, no. 4, pp. 1477–1491, Aug. 2003.
- [40] P. A. Valdes-Sosa, A. Roebroeck, J. Daunizeau, and K. Friston, "Effective connectivity: Influence, causality and biophysical modeling," *NeuroImage*, vol. 58, no. 2, pp. 339–361, Sept. 2011.
- [41] S. L. Bressler and A. K. Seth, "Wiener-Granger causality: A well established methodology," *NeuroImage*, vol. 58, no. 2, pp. 323–329, Sept. 2011.
- [42] C. W. J. Granger, "Investigating causal relations by econometric models and cross-spectral methods," *Econometrica*, vol. 37, no. 3, pp. 424–438, Aug. 1969.
- [43] K. J. Friston, A. M. Bastos, A. Oswal, B. van Wijk, C. Richter, and V. Litvak, "Granger causality revisited," *NeuroImage*, vol. 101, pp. 796–808, Nov. 2014.
- [44] G. Deshpande, K. Sathian, and X. Hu, "Assessing and compensating for zero-lag correlation effects in time-lagged Granger causality analysis of fMRI," *IEEE Trans. Biomed. Eng.*, vol. 57, no. 6, pp. 1446–1456, June 2010.
- [45] M. Havlicek, J. Jan, M. Brazdil, and V. D. Calhoun, "Dynamic Granger causality based on Kalman filter for evaluation of functional network connectivity in fMRI data," *NeuroImage*, vol. 53, no. 1, pp. 65–77, Oct. 2010.
- [46] L. Barnett and A. K. Seth, "Granger causality for state-space models," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 91, no. 4, Apr. 2015.
- [47] S. Wright, "Correlation and causation," *J. Agric. Res.*, vol. 20, no. 7, pp. 557–585, Jan. 1921.
- [48] O. Duncan, *Introduction to Structural Equation Models*. New York: Academic Press, 1975.
- [49] A. S. Goldberger, "Structural equation methods in social-sciences," *Econometrica*, vol. 40, no. 6, pp. 979–1001, Nov. 1972.
- [50] A. R. McIntosh and F. Gonzalez-Lima, "Structural modeling of functional neural pathways mapped with 2-deoxyglucose: Effects of acoustic startle habituation on the auditory system," *Brain Res.*, vol. 547, no. 2, pp. 295–302, May 1991.
- [51] W. D. Penny, K. E. Stephan, A. Mechelli, and K. J. Friston, "Modelling functional integration: A comparison of structural equation and dynamic causal models," *NeuroImage*, vol. 23, no. S1, pp. S264–S274, 2004.
- [52] J. Kim and B. Horwitz, "How well does structural equation modeling reveal abnormal brain anatomical connections? An fMRI simulation study," *NeuroImage*, vol. 45, no. 4, pp. 1190–1198, May 2009.
- [53] K. A. Bollen, *Structural Equations with Latent Variables*. Hoboken, NJ: Wiley, 1989.
- [54] J. Geweke, "Measurement of linear dependence and feedback between multiple time series," *J. Amer. Statist. Assoc.*, vol. 77, no. 378, pp. 304–313, 1982.

- [55] J. S. Bendat, *Nonlinear System Analysis and Identification from Random Data*. Hoboken, NJ: Wiley, 1990.
- [56] K. Friston, R. Moran, and A. K. Seth, "Analysing connectivity with Granger causality and dynamic causal modelling," *Curr. Opin. Neurobiol.*, vol. 23, no. 2, pp. 172–178, Apr. 2013.
- [57] C. E. Shannon, "A mathematical theory of communication," *Bell Syst. Tech. J.*, vol. 27, no. 3, pp. 379–423, 1948.
- [58] D. Zhou, W. K. Thompson, and G. Siegle, "MATLAB toolbox for functional connectivity," *NeuroImage*, vol. 47, no. 4, pp. 1590–1607, Oct. 2009.
- [59] R. Salvador, J. Suckling, C. Schwarzbauer, and E. Bullmore, "Undirected graphs of frequency-dependent functional connectivity in whole brain networks," *Philos. Trans. R. Soc. London B, Biol. Sci.*, vol. 360, no. 1457, pp. 937–946, May 2005.
- [60] S. M. Smith, K. L. Miller, G. Salimi-Khorshidi, M. Webster, C. F. Beckmann, T. E. Nichols, J. D. Ramsey, M. W. Woolrich, "Network modelling methods for fMRI," *NeuroImage*, vol. 54, no. 2, pp. 875–891, Jan. 2011.
- [61] W. J. Freeman, "Characterization of state transitions in spatially distributed, chaotic, nonlinear, dynamical systems in cerebral cortex," *Integr. Physiol. Behav. Sci.*, vol. 29, no. 3, pp. 294–306, July–Sept. 1994.
- [62] S. Coombes and S. H. Doole, "Neuronal populations with reciprocal inhibition and rebound currents: Effects of synaptic and threshold noise," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics*, vol. 54, no. 4, pp. 4054–4065, Oct. 1996.
- [63] P. A. Robinson, C. J. Rennie, and J. J. Wright, "Propagation and stability of waves of electrical activity in the cerebral cortex," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 56, no. 1, pp. 826–840, July 1997.
- [64] I. Tsuda, "Toward an interpretation of dynamic neural activity in terms of chaotic dynamical systems," *Behav. Brain Sci.*, vol. 24, no. 5, pp. 793–810, Oct. 2001.
- [65] W. J. Freeman, "A field-theoretic approach to understanding scale-free neocortical dynamics," *Biol. Cybern.*, vol. 92, no. 6, pp. 350–359, June 2005.
- [66] S. L. Bressler and E. Tognoli, "Operational principles of neurocognitive networks," *Int. J. Psychophysiol.*, vol. 60, no. 2, pp. 139–148, May 2006.
- [67] B. Kriener, T. Tetzlaff, A. Aertsen, M. Diesmann, and S. Rotter, "Correlations and population dynamics in cortical networks," *Neural Comput.*, vol. 20, no. 9, pp. 2185–2226, Sept. 2008.
- [68] M. Rubinov, O. Sporns, C. van Leeuwen, and M. Breakspear, "Symbiotic relationship between brain structure and dynamics," *BMC Neurosci.*, vol. 10, p. 55, June 2009.
- [69] M. A. Buice and J. D. Cowan, "Statistical mechanics of the neocortex," *Prog. Biophys. Mol. Biol.*, vol. 99, nos. 2–3, pp. 53–86, Feb.–May 2009.
- [70] C. J. Honey, R. Kotter, M. Breakspear, and O. Sporns, "Network structure of cerebral cortex shapes functional connectivity on multiple time scales," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 104, no. 24, pp. 10240–10245, June 2007.
- [71] G. Deco, V. Jirsa, A. R. McIntosh, O. Sporns, and R. Kotter, "Key role of coupling, delay, and noise in resting brain fluctuations," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 106, no. 25, pp. 10302–10307, June 2009.
- [72] C. J. Honey, O. Sporns, L. Cammoun, X. Gigandet, J. P. Thiran, R. Meuli, and P. Hagmann, "Predicting human resting-state functional connectivity from structural connectivity," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 106, no. 6, pp. 2035–2040, Feb. 2009.
- [73] G. Deco, V. K. Jirsa, and A. R. McIntosh, "Emerging concepts for the dynamical organization of resting-state activity in the brain," *Nat. Rev. Neurosci.*, vol. 12, no. 1, pp. 43–56, Jan. 2011.
- [74] B. Biswal, F. Z. Yetkin, V. M. Haughton, and J. S. Hyde, "Functional connectivity in the motor cortex of resting human brain using echo-planar MRI," *Magn. Reson. Med.*, vol. 34, no. 4, pp. 537–541, Oct. 1995.
- [75] B. B. Biswal, J. Van Kylen, and J. S. Hyde, "Simultaneous assessment of flow and BOLD signals in resting-state functional connectivity maps," *NMR Biomed.*, vol. 10, nos. 4–5, pp. 165–170, June–Aug. 1997.
- [76] J. S. Damoiseaux and M. D. Greicius, "Greater than the sum of its parts: A review of studies combining structural connectivity and resting-state functional connectivity," *Brain Struct. Funct.*, vol. 213, no. 6, pp. 525–533, Oct. 2009.
- [77] M. D. Greicius, K. Supekar, V. Menon, and R. F. Dougherty, "Resting-state functional connectivity reflects structural connectivity in the default mode network," *Cereb. Cortex*, vol. 19, no. 1, pp. 72–78, Jan. 2009.
- [78] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, "A default mode of brain function," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 98, no. 2, pp. 676–682, Jan. 2001.
- [79] S. Achard, R. Salvador, B. Whitcher, J. Suckling, and E. Bullmore, "A resilient, low-frequency, small-world human brain functional network with highly connected association cortical hubs," *J. Neurosci.*, vol. 26, no. 1, pp. 63–72, Jan. 2006.
- [80] D. S. Bassett, A. Meyer-Lindenberg, S. Achard, T. Duke, and E. Bullmore, "Adaptive reconfiguration of fractal small-world human brain functional networks," *Proc. Natl. Acad. Sci. U.S.A.*, vol. 103, no. 51, pp. 19518–19523, Dec. 2006.
- [81] O. Sporns, *Networks of the Brain*. Cambridge, MA: MIT Press, 2010.
- [82] J. Beggs, "Editorial: Can there be a physics of the brain?" *Phys. Rev. Lett.*, vol. 114, no. 22, 220001, June 2015.
- [83] J. M. Beggs, "The criticality hypothesis: How local cortical networks might optimize information processing," *Philos. Trans. A, Math. Phys. Eng. Sci.*, vol. 366, no. 1864, pp. 329–343, Feb. 2008.
- [84] V. K. Jirsa, R. Friedrich, H. Haken, and J. A. Kelso, "A theoretical model of phase transitions in the human brain," *Biol. Cybern.*, vol. 71, no. 1, pp. 27–35, 1994.
- [85] V. K. Jirsa and H. Haken, "Field theory of electromagnetic brain activity," *Phys. Rev. Lett.*, vol. 77, no. 5, pp. 960–963, July 1996.
- [86] V. K. Jirsa and J. A. Kelso, "Spatiotemporal pattern formation in neural systems with heterogeneous connection topologies," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Topics*, vol. 62, no. 6, pp. 8462–8465, Dec. 2000.
- [87] E. Tognoli and J. A. Kelso, "Brain coordination dynamics: True and false faces of phase synchrony and metastability," *Prog. Neurobiol.*, vol. 87, no. 1, pp. 31–40, Jan. 2009.
- [88] W. Tschacher and H. Haken, "Intentionality in non-equilibrium systems? The functional aspects of self-organized pattern formation," *New Ideas Psychol.*, vol. 25, no. 1, pp. 1–15, Apr. 2007.
- [89] W. L. Shew and D. Plenz, "The functional benefits of criticality in the cortex," *Neuroscientist*, vol. 19, no. 1, pp. 88–100, Feb. 2013.
- [90] O. Shriki, J. Alstott, F. Carver, T. Holroyd, R. N. Henson, M. L. Smith, R. Coppola, E. Bullmore, et al., "Neuronal avalanches in the resting MEG of the human brain," *J. Neurosci.*, vol. 33, no. 16, pp. 7079–7090, Apr. 2013.
- [91] A. Haimovici, E. Tagliazucchi, P. Balenzuela, and D. R. Chialvo, "Brain organization into resting state networks emerges at criticality on a model of the human connectome," *Phys. Rev. Lett.*, vol. 110, no. 17, 178101, Apr. 2013.
- [92] J. A. Roberts, T. W. Boonstra, and M. Breakspear, "The heavy tail of the human brain," *Curr. Opin. Neurobiol.*, vol. 31, pp. 164–172, Apr. 2015.
- [93] J. A. Kelso, "Instabilities and phase transitions in human brain and behavior," *Front. Hum. Neurosci.*, vol. 4, p. 23, 2010.
- [94] G. Deco and V. K. Jirsa, "Ongoing cortical activity at rest: Criticality, multistability, and ghost attractors," *J. Neurosci.*, vol. 32, no. 10, pp. 3366–3375, Mar. 2012.
- [95] F. Freyer, J. A. Roberts, R. Becker, P. A. Robinson, P. Ritter, and M. Breakspear, "Biophysical mechanisms of multistability in resting-state cortical rhythms," *J. Neurosci.*, vol. 31, no. 17, pp. 6353–6361, Apr. 2011.
- [96] P. A. Robinson, C. J. Rennie, and D. L. Rowe, "Dynamics of large-scale brain activity in normal arousal states and epileptic seizures," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 65, no. 4, 041924, Apr. 2002.
- [97] M. J. Aburn, C. A. Holmes, J. A. Roberts, T. W. Boonstra, and M. Breakspear, "Critical fluctuations in cortical models near instability," *Front. Physiol.*, vol. 3, p. 331, Aug. 2012.
- [98] P. A. Robinson, "Neurophysical theory of coherence and correlations of electroencephalographic and electrocorticographic signals," *J. Theor. Biol.*, vol. 222, no. 2, pp. 163–175, May 2003.
- [99] H. White and X. Lu, "Granger causality and dynamic structural systems," *J. Financ. Economet.*, vol. 8, no. 2, pp. 193–243, 2010.
- [100] P. Spirtes, C. Glymour, and R. Scheines, *Causation, Prediction, and Search*. New York: Springer, 1993.
- [101] S. L. Lauritzen, *Graphical models*. Gloucestershire, U.K.: Clarendon Press, 1996.
- [102] J. Robins, "A new approach to causal inference in mortality studies with a sustained exposure period—Application to control of the healthy worker survivor effect," *Math. Model.*, vol. 7, nos. 9–12, pp. 1393–1512, 1986.
- [103] K. E. Stephan, L. Kasper, L. M. Harrison, J. Daunizeau, H. E. den Ouden, M. Breakspear, K. J. Friston, "Nonlinear dynamic causal models for fMRI," *NeuroImage*, vol. 42, no. 2, pp. 649–662, Aug. 2008.

- [104] A. C. Marreiros, S. J. Kiebel, and K. J. Friston, "Dynamic causal modelling for fMRI: A two-state model," *NeuroImage*, vol. 39, no. 1, pp. 269–278, Jan. 2008.
- [105] W. D. Penny, K. E. Stephan, A. Mechelli, and K. J. Friston, "Comparing dynamic causal models," *NeuroImage*, vol. 22, no. 3, pp. 1157–1172, July 2004.
- [106] F. Acs and M. W. Greenlee, "Connectivity modulation of early visual processing areas during covert and overt tracking tasks," *NeuroImage*, vol. 41, no. 2, pp. 380–388, June 2008.
- [107] P. Allen, A. Mechelli, K. E. Stephan, F. Day, J. Dalton, S. Williams, and P. K. McGuire, "Fronto-temporal interactions during overt verbal initiation and suppression," *J. Cogn. Neurosci.*, vol. 20, no. 9, pp. 1656–1669, Sept. 2008.
- [108] M. Boly, M. I. Garrido, O. Gosseries, M. A. Bruno, P. Boveroux, C. Schnakers, M. Massimini, V. Litvak, et al., "Preserved feedforward but impaired top-down processes in the vegetative state," *Science*, vol. 332, no. 6031, pp. 858–862, May 2011.
- [109] M. J. Grol, J. Majdandzic, K. E. Stephan, L. Verhagen, H. C. Dijkerman, H. Bekkering, F. A. J. Verstraten, and I. Toni, "Parieto-frontal connectivity during visually guided grasping," *J. Neurosci.*, vol. 27, no. 44, pp. 11877–11887, Oct. 2007.
- [110] S. Heim, S. B. Eickhoff, A. K. Ischebeck, A. D. Friederici, K. E. Stephan, and K. Amunts, "Effective connectivity of the left BA 44, BA 45, and inferior temporal gyrus during lexical and phonological decisions identified with DCM," *Hum. Brain Mapp.*, vol. 30, no. 2, pp. 392–402, Feb. 2009.
- [111] A. P. Smith, K. E. Stephan, M. D. Rugg, and R. J. Dolan, "Task and content modulate amygdala-hippocampal connectivity in emotional retrieval," *Neuron*, vol. 49, no. 4, pp. 631–638, Feb. 2006.
- [112] C. Summerfield and E. Kochlin, "A neural representation of prior information during perceptual inference," *Neuron*, vol. 59, no. 2, pp. 336–347, July 2008.
- [113] A. A. Faisal, L. P. Selen, and D. M. Wolpert, "Noise in the nervous system," *Nat. Rev. Neurosci.*, vol. 9, no. 4, pp. 292–303, Apr. 2008.
- [114] K. J. Friston, J. Kahan, B. Biswal, and A. Razi, "A DCM for resting state fMRI," *NeuroImage*, vol. 94, no. 100, pp. 396–407, July 2014.
- [115] K. Friston, "Hierarchical models in the brain," *PLoS Comput. Biol.*, vol. 4, no. 11, e1000211, Nov. 2008.
- [116] B. Li, J. Daunizeau, K. E. Stephan, W. Penny, D. Hu, and K. Friston, "Generalised filtering and stochastic DCM for fMRI," *NeuroImage*, vol. 58, no. 2, pp. 442–457, Sept. 2011.
- [117] J. Daunizeau, K. E. Stephan, and K. J. Friston, "Stochastic dynamic causal modelling of fMRI data: Should we care about neural noise?" *NeuroImage*, vol. 62, no. 1, pp. 464–481, Aug. 2012.
- [118] M. M. Tropper, "Ergodic and quasi-deterministic properties of finite-dimensional stochastic systems," *J. Stat. Phys.*, vol. 17, no. 6, pp. 491–509, 1977.
- [119] K. J. Friston, J. Kahan, A. Razi, K. E. Stephan, and O. Sporns, "On nodes and modes in resting state fMRI," *NeuroImage*, vol. 99, pp. 533–547, May 2014.
- [120] V. D. Calhoun, T. Adali, G. D. Pearlson, and J. J. Pekar, "A method for making group inferences from functional MRI data using independent component analysis," *Hum. Brain Mapp.*, vol. 14, no. 3, pp. 140–151, Nov. 2001.
- [121] A. Razi, J. Kahan, G. Rees, and K. J. Friston, "Construct validation of a DCM for resting state fMRI," *NeuroImage*, vol. 106, pp. 1–14, Feb. 2015.
- [122] K. J. Friston, N. Trujillo-Barreto, and J. Daunizeau, "DEM: A variational treatment of dynamic systems," *NeuroImage*, vol. 41, no. 3, pp. 849–885, July 2008.
- [123] J. Kahan, M. Urner, R. Moran, G. Flandin, A. Marreiros, L. Mancini, M. White, J. Thornton, et al., "Resting state functional MRI in Parkinson's disease: The impact of deep brain stimulation on 'effective' connectivity," *Brain*, vol. 137, no. 4, pp. 1130–1144, 2014.
- [124] P. A. Robinson, C. J. Rennie, D. L. Rowe, and S. C. O'Connor, "Estimation of multiscale neurophysiological parameters by electroencephalographic means," *Hum. Brain Mapp.*, vol. 23, no. 1, pp. 53–72, Sept. 2004.
- [125] C. J. Stam and E. A. de Bruin, "Scale-free dynamics of global functional connectivity in the human brain," *Hum. Brain Mapp.*, vol. 22, no. 2, pp. 97–109, June 2004.
- [126] C. W. Shin and S. Kim, "Self-organized criticality and scale-free properties in emergent functional neural networks," *Phys. Rev. E, Stat. Nonlin. Soft Matter Phys.*, vol. 74, no. 4, 045101, Oct. 2006.
- [127] J. M. Beggs and D. Plenz, "Neuronal avalanches in neocortical circuits," *J. Neurosci.*, vol. 23, no. 35, pp. 11167–11177, Dec. 2003.
- [128] C. Bedard, H. Kroger, and A. Destexhe, "Does the 1/f frequency scaling of brain signals reflect self-organized critical states?" *Phys. Rev. Lett.*, vol. 97, no. 11, 118102, Sept. 2006.
- [129] K. J. Miller, L. B. Sorensen, J. G. Ojemann, and M. den Nijs, "Power-law scaling in the brain surface electric potential," *PLoS Comput. Biol.*, vol. 5, no. 12, e1000609, Dec. 2009.
- [130] K. Linkenkaer-Hansen, V. V. Nikouline, J. M. Palva, and R. J. Ilmoniemi, "Long-range temporal correlations and scaling behavior in human brain oscillations," *J. Neurosci.*, vol. 21, no. 4, pp. 1370–1377, Feb. 2001.
- [131] M. G. Kitzbichler, M. L. Smith, S. R. Christensen, and E. Bullmore, "Broad-band criticality of human brain network synchronization," *PLoS Comput. Biol.*, vol. 5, no. 3, e1000314, Mar. 2009.
- [132] F. Freyer, K. Aquino, P. A. Robinson, P. Ritter, and M. Breakspear, "Bistability and non-Gaussian fluctuations in spontaneous cortical activity," *J. Neurosci.*, vol. 29, no. 26, pp. 8512–8524, July 2009.
- [133] K. Friston, J. Mattout, N. Trujillo-Barreto, J. Ashburner, and W. Penny, "Variational free energy and the Laplace approximation," *NeuroImage*, vol. 34, no. 1, pp. 220–234, Jan. 2007.
- [134] K. A. Tsvetanov, R. N. A. Henson, L. K. Tyler, A. Razi, L. Geerligns, T. Ham, and J. Rowe, "Higher extrinsic and intrinsic brain network connectivity maintains cognition across the lifespan despite age-related decay of regional neuronal activation," *J. Neurosci.*, to be published.
- [135] W. R. Shirer, S. Ryali, E. Rykhlevskaia, V. Menon, and M. D. Greicius, "Decoding subject-driven cognitive states with whole-brain connectivity patterns," *Cereb. Cortex*, vol. 22, no. 1, pp. 158–165, Jan. 2012.
- [136] P. L. St. Jacques, P. A. Kragel, and D. C. Rubin, "Dynamic neural networks supporting memory retrieval," *NeuroImage*, vol. 57, no. 2, pp. 608–616, July 2011.
- [137] M. C. Stevens, K. A. Kiehl, G. D. Pearlson, and V. D. Calhoun, "Functional neural networks underlying response inhibition in adolescents and adults," *Behav. Brain Res.*, vol. 181, no. 1, pp. 12–22, July 2007.
- [138] X. Di and B. B. Biswal, "Identifying the default mode network structure using dynamic causal modeling on resting-state functional magnetic resonance imaging," *NeuroImage*, vol. 86, pp. 53–59, Feb. 2014.
- [139] J. Daunizeau, O. David, and K. E. Stephan, "Dynamic causal modelling: A critical review of the biophysical and statistical foundations," *NeuroImage*, vol. 58, no. 2, pp. 312–322, Sept. 2011.
- [140] G. Lohmann, K. E. Stephan, K. Müller, and R. Turner, "Critical comments on dynamic causal modelling," *NeuroImage*, vol. 59, no. 3, pp. 2322–2329, Feb. 2012.
- [141] M. Havlicek, A. Roebroeck, K. Friston, A. Gardumi, D. Ivanov, and K. Uludag, "Physiologically informed dynamic causal modeling of fMRI data," *NeuroImage*, vol. 122, pp. 355–372, Nov. 2015.
- [142] K. Friston, J. Daunizeau, and K. E. Stephan, "Model selection and gobbledygook: Response to Lohmann et al," *NeuroImage*, vol. 75, pp. 275–278, discussion 279–281, July 2013.
- [143] S. Frassle, K. E. Stephan, K. J. Friston, M. Steup, S. Krach, F. M. Paulus, and A. Jansen, "Test-retest reliability of dynamic causal modeling for fMRI," *NeuroImage*, vol. 117, pp. 56–66, Aug. 2015.
- [144] J. B. Rowe, L. E. Hughes, R. A. Barker, and A. M. Owen, "Dynamic causal modelling of effective connectivity from fMRI: Are results reproducible and sensitive to Parkinson's disease and its treatment?" *NeuroImage*, vol. 52, no. 3, pp. 1015–1026, Sept. 2010.
- [145] B. Schuyler, J. M. Ollinger, T. R. Oakes, T. Johnstone, and R. J. Davidson, "Dynamic causal modeling applied to fMRI data shows high reliability," *NeuroImage*, vol. 49, no. 1, pp. 603–611, Jan. 2010.
- [146] M. I. Garrido, J. M. Kilner, S. J. Kiebel, K. E. Stephan, and K. J. Friston, "Dynamic causal modelling of evoked potentials: A reproducibility study," *NeuroImage*, vol. 36, no. 3, pp. 571–580, July 2007.
- [147] O. David, I. Guillemain, S. Saittet, S. Reynt, C. Deransart, C. Segebarth, and A. Depaulis, "Identifying neural drivers with functional MRI: An electrophysiological validation," *PLoS Biol.*, vol. 6, no. 12, pp. 2683–2697, Dec. 2008.
- [148] E. I. Lomakina, S. Paliwal, A. O. Diaconescu, K. H. Brodersen, E. A. Aponte, J. M. Buhmann, and K. E. Stephan, "Inversion of hierarchical Bayesian models using Gaussian processes," *NeuroImage*, vol. 118, pp. 133–145, Sept. 2015.
- [149] B. Sengupta, K. J. Friston, and W. D. Penny, "Gradient-free MCMC methods for dynamic causal modelling," *NeuroImage*, vol. 112, pp. 375–381, May 2015.
- [150] B. Sengupta, K. J. Friston, and W. D. Penny, "Gradient-based MCMC samplers for dynamic causal modelling," *NeuroImage*, vol. 125, pp. 1107–1118, Jan. 2016.