



## OPEN ACCESS

## EDITED BY

Marta Bianciardi,  
Massachusetts General Hospital  
and Harvard Medical School,  
United States

## REVIEWED BY

Zeus Gracia-Tabuenca,  
McGill University, Canada  
Peter Zeidman,  
University College London,  
United Kingdom

## \*CORRESPONDENCE

Guoshi Li  
guoshi\_li@med.unc.edu

## SPECIALTY SECTION

This article was submitted to  
Cognitive Neuroscience,  
a section of the journal  
Frontiers in Human Neuroscience

RECEIVED 10 May 2022

ACCEPTED 28 July 2022

PUBLISHED 17 August 2022

## CITATION

Li G and Yap P-T (2022) From  
descriptive connectome  
to mechanistic connectome:  
Generative modeling in functional  
magnetic resonance imaging analysis.  
*Front. Hum. Neurosci.* 16:940842.  
doi: 10.3389/fnhum.2022.940842

## COPYRIGHT

© 2022 Li and Yap. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which  
does not comply with these terms.

# From descriptive connectome to mechanistic connectome: Generative modeling in functional magnetic resonance imaging analysis

Guoshi Li<sup>1,2\*</sup> and Pew-Thian Yap<sup>1,2</sup>

<sup>1</sup>Department of Radiology, University of North Carolina, Chapel Hill, NC, United States, <sup>2</sup>Biomedical Research Imaging Center, University of North Carolina, Chapel Hill, NC, United States

As a newly emerging field, connectomics has greatly advanced our understanding of the wiring diagram and organizational features of the human brain. Generative modeling-based connectome analysis, in particular, plays a vital role in deciphering the neural mechanisms of cognitive functions in health and dysfunction in diseases. Here we review the foundation and development of major generative modeling approaches for functional magnetic resonance imaging (fMRI) and survey their applications to cognitive or clinical neuroscience problems. We argue that conventional structural and functional connectivity (FC) analysis alone is not sufficient to reveal the complex circuit interactions underlying observed neuroimaging data and should be supplemented with generative modeling-based effective connectivity and simulation, a fruitful practice that we term “mechanistic connectome.” The transformation from descriptive connectome to mechanistic connectome will open up promising avenues to gain mechanistic insights into the delicate operating principles of the human brain and their potential impairments in diseases, which facilitates the development of effective personalized treatments to curb neurological and psychiatric disorders.

## KEYWORDS

computational modeling, brain network, dynamical causal model, biophysical network model, Dynamic Neural Model, neural mass model, connectome, fMRI

## Introduction

The human brain is a fascinating machine in which the interactions of vast numbers of distributed circuits and networks give rise to complex cognitive functions such as perception, attention, decision making and memory. Delineating the brain's anatomical wiring diagram and its functional operation principles constitutes an important first

step to decipher the underlying mechanisms of cognition. Connectomics is a newly developed field dedicated to providing a complete map of neuronal connections in the brain (Sporns et al., 2005). Fueled by the rapid advances of non-invasive neuroimaging techniques, connectomics has become one of the most vibrant disciplines in neuroscience (Craddock et al., 2013). As a major international neuroscience initiative, the Human Connectome Project (HCP) aims to build a network map of neural systems by systematically characterizing structural and functional connectivity (FC) of the human brain (Van Essen and Glasser, 2016). The project has made groundbreaking achievements in both method development and scientific discoveries, which acquired and analyzed a wealth of multimodal MRI and magnetoencephalography (MEG) data of unprecedented quality for further exploration of brain cognitive functions (Elam et al., 2021). Connectomics is not only important for studying normal brain functions but also highly promising to understand the pathological basis of neurological and psychiatric disorders for better diagnosis and treatments (Deco and Kringelbach, 2014).

As the essential components of human connectomics, structural and functional connectomes are driven by two major imaging modalities, diffusion MRI and functional MRI, respectively. Diffusion MRI is a method based on measuring the random Brownian motion of water molecules within the white matter (Baliyan et al., 2016), which can be used to infer long-distance white matter tracts (tractography) that connect distant brain regions informing structural connectivity (SC) (Van Essen and Glasser, 2016). Such structural information describes the physical substrate underlying brain functions. On the other hand, functional MRI (fMRI) is a class of imaging methods designed to measure regional, time-varying changes in brain metabolism in response to either task stimuli (task-fMRI) or spontaneous modulation of neural process (resting-state fMRI) (Glover, 2011). The most common form of fMRI employs blood oxygen level-dependent (BOLD) contrast imaging which measures variations in deoxyhemoglobin concentrations, an indirect measure of neuronal activity (Soares et al., 2016; Chow et al., 2017). BOLD-fMRI has been widely used for large-scale brain mapping (Gore, 2003; Just and Varma, 2007; Glover, 2011; Power et al., 2014) and the fMRI data is predominantly analyzed using FC, which is defined as the statistical dependencies among fluctuating BOLD timeseries from distributed brain regions (Friston, 2011). FC can be computed using either simple correlation analysis (e.g., Pearson's correlation) or more sophisticated statistical methods such as mutual information, independent component analysis and Hidden Markov models (Karahanoğlu and Van De Ville, 2017; Vidaurre et al., 2017; Bolton et al., 2018).

Despite the great success and widespread use of SC and FC in characterizing the organizational principles of large-scale brain networks, their application to fundamental cognitive neuroscience problems and clinical treatments is still limited

due to several notable limitations. *First of all*, such macroscopic connectivity analysis is largely descriptive and superficial (Stephan et al., 2015; Braun et al., 2018), thus unable to offer a mechanistic account of the neural process underlying cognitive function or dysfunction. *Second*, both SC and FC are undirected and unsigned. Consequently, they cannot model the inherent asymmetry observed in both anatomical and functional connections (Felleman and Van Essen, 1991; Frässle et al., 2016). In addition, the unsigned connections imply that SC and FC are not able to infer excitatory or inhibitory coupling strengths for excitation-inhibition (E-I) balance estimation. This disadvantage is non-trivial because E-I balance plays an important role in neural coding, synaptic plasticity, and neurogenesis (Froemke, 2015; Deneve and Machens, 2016; Lopatina et al., 2019). *Lastly*, neither SC or FC captures intrinsic or intra-regional connections. As cortical neurons are more subject to the influence of short-range (local) than long-range (inter-regional) connections (Unal et al., 2013), an alternative connectivity measure or modeling approach is needed to account for intrinsic neural interactions for a more comprehensive characterization of connectome.

To overcome the limitations of SC and FC, Friston and colleagues introduced the concept of effective connectivity (EC), which is defined as the directed causal influence among neuronal populations (Friston et al., 2003; Friston, 2011). Different from SC and FC, EC builds on a generative model of neural interactions and corresponds to the coupling strengths of the neuronal model estimated from the observed neuroimaging data (Friston et al., 2003; Friston, 2011). As the generative model describes how the latent (hidden) neuronal states and their interactions give rise to the observed BOLD measurements (for fMRI), EC can potentially provide mechanistic neuronal accounts of the fMRI data under both normal cognitive process and abnormal disease state. In addition, EC is directed and signed, allowing estimation of excitatory and inhibitory coupling strengths for E-I balance inference. Furthermore, by incorporating more fine-grained microscopic or mesoscopic neuronal models, intra-regional connection strengths can also be estimated. Thus, effective connectivity based on generative modeling provides a highly significant complementary approach to conventional SC and FC analysis, which enables deeper mechanistic understanding of functional connectome.

This review paper attempts to give a concise overview of existing generative modeling approaches in fMRI analysis and how they contribute to the emerging and rapidly growing field of human connectomics. We first define generative modeling in a general sense, discuss their important roles in neuroscience, and introduce three major generative modeling frameworks in fMRI connectome analysis. Next we review the state-of-the-art of each of the three major generative modeling approaches and sample their applications to cognitive or clinical neuroscience problems. Lastly, we summarize the developments and achievements of generative models in fMRI connectome analysis and discuss

future research directions for this dynamic and promising field in neuroscience.

## Generative modeling of human connectomics

Generative models are computational models that are built to simulate the response profile of a complex system based on the physical mechanisms of interactive components (Li and Cleland, 2018). Such models are designed to generate proper values for all embedded variables with interpretable physical meaning rather than deriving certain properties from a “black box” to study certain higher-level phenomena (Li and Cleland, 2018). Generative modeling has many useful applications, such as interpreting the underlying mechanisms of emergent properties in complex systems, testing the sufficiency of working hypothesis, and making testable predictions to guide experimental design. Since the introduction of the classical Hodgkin and Huxley model (Hodgkin and Huxley, 1952), generative modeling has become a core component in computational neuroscience. Traditional generative neuronal models (based on non-human data) have been focused on how collective properties of neurons, synapses and structures give rise to network activity and dynamics to sustain neural circuit functions (e.g., Hasselmo et al., 1995; Durstewitz et al., 2000; Li et al., 2009; Markram et al., 2015; Hass et al., 2016). Such detailed biophysical neuronal models involve collecting sufficient anatomical and electrophysiological data at the cellular and synaptic levels mostly from non-human subjects. Generative models can then be constructed by integrating the anatomical and physiological data with known neuronal biophysics within a microscale circuit. A notable example for such detailed modeling is the Blue Brain project, an international initiative that aims to create a digital reconstruction of the mouse brain to identify the fundamental principles of brain structure and function (Markram, 2006).

The generative models for human connectomics have several distinct features compared with detailed microcircuit models for non-human data. *First*, they focus on the estimation of effective connectivity or coupling strengths among neural populations. *Second*, they are driven and constrained by neuroimaging data such as diffusion MRI, fMRI and MEG, and seek to reveal the neuronal mechanisms underlying the observed data. *Third*, they generally apply population-level (neural mass) models to describe network dynamics as such dynamic models relate more closely to macroscopic neuroimaging data by describing the collective activity of neuronal populations (Breakspear, 2017). *Lastly*, generative models for connectomics commonly involve an optimization strategy to estimate the coupling strengths among neural populations based on neuroimaging data. This contrasts with detailed microcircuit neuronal models that usually infer model parameters from

anatomical and electrophysiological data. Thus, generative models for connectomics consist of three essential components: (1) a neuronal model that generates population-level neural activity; (2) an observation model that transforms the neural activity to simulated neuroimaging data; and (3) an optimization scheme to estimate connection parameters.

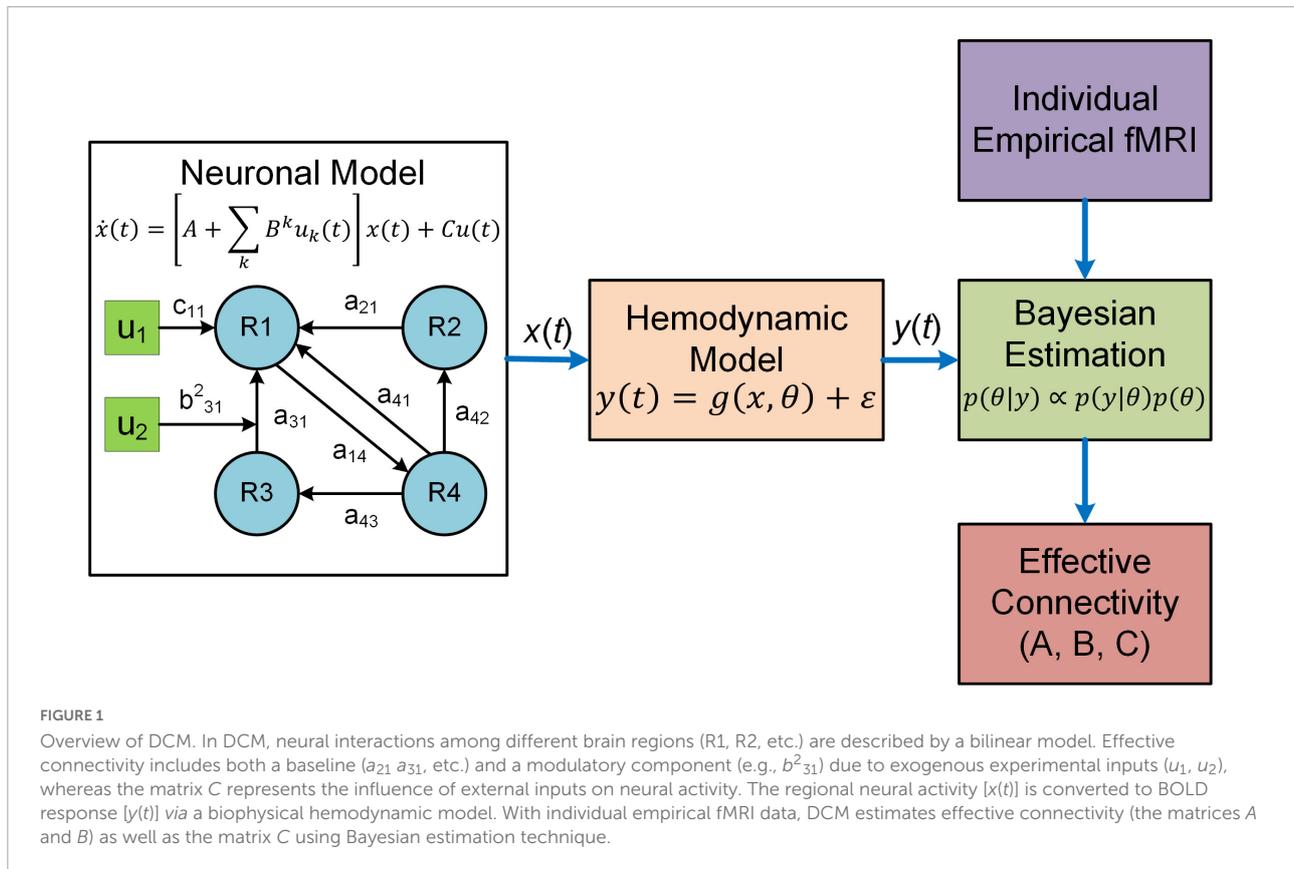
Depending on the differences in neuronal models, the scope of parameter estimation and the optimization scheme, existing generative models for connectomics can be broadly classified into three major types: (1) Dynamic Causal Model (DCM); (2) Biophysical Network Model (BNM); and (3) Dynamic Neural Model (DNM) with direct parameterization. Below we review the foundation and development of these three types of generative models and survey their applications to cognitive or clinical neuroscience.

## Dynamic causal modeling

### Overview

Dynamic causal modeling (DCM) is a predominant analysis framework to infer effectivity connectivity of individual connections at single subject level using standard Variational Laplace procedures (Friston et al., 2003, 2007). A typical DCM contains a forward (generative) model that describes the dynamics of interacting neuronal populations and a measurement model that converts the neuronal activity into measurable observations such as fMRI, MEG and electroencephalogram (EEG) (Figure 1). During model inference, DCM simulates the BOLD responses for models with different configurations of connectivity and determines the model that best characterizes the empirical fMRI data. DCM adopts a two-stage process to estimate EC (Zeidman et al., 2019a,b). The first stage is Bayesian model inversion (estimation), a process that finds the parameters which provide the best trade-off between model accuracy (how good the model fits the data) and model complexity (how far the parameters need to deviate from their prior values to fit the data); such trade-off is quantified as model evidence. The second stage is termed Bayesian model comparison where models with different network connectivity are compared based on evidence either at the single-subject or group level. For detailed procedures about the two-stage process, see two recent tutorial papers on DCM (Zeidman et al., 2019a,b).

The original deterministic DCM applies to task-fMRI only and is restricted to relatively small networks (< 10 brain regions) due to computational burden (Daunizeau et al., 2011). To account for resting-state activity, two variations of DCM, stochastic DCM (Li et al., 2011) and spectral DCM (Friston et al., 2014), have been developed. Stochastic DCM estimates EC as well as random fluctuations on both neural states and measurements, which is computationally intensive and



thus can handle only a few numbers of brain regions. By comparison, spectral DCM estimates the parameters of cross-spectral density of neuronal fluctuations instead of time-varying fluctuations in neuronal states, which is much more stable and computationally efficient. By using FC as prior constraints, the computational efficiency of spectral DCM is further improved, enabling modeling of relatively large networks (36 nodes, Razi et al., 2017).

The original DCM is limited to modeling one neural population for each brain region, which is later extended to two-state DCM to account for the intra-regional interactions between excitatory and inhibitory neural populations (Marreiros et al., 2008). Despite the added biological realism, the two-state DCM is still a linear model, and consequently may not accurately represent the brain neural dynamics in the long term (Singh et al., 2020). A major advancement to DCM for fMRI was introduced recently by replacing the bilinear model with a neural mass model (NMM) of the canonical microcircuit (Friston et al., 2019). Specifically, four neural populations are modeled in each brain region coupled with both inter- and intra-laminar connections of the cortical microcircuitry; each neural population is further represented by two hidden states whose dynamics are governed by second-order differential equations. By incorporating a sophisticated and physiologically informed NMM, DCM for fMRI parallels the development

of DCM for electrophysiological data (Moran et al., 2011; Friston et al., 2012), which is able to provide deeper mechanistic insights for observed fMRI data. It should be noted that both the earlier two-state DCM or the latest NMM-based DCM are designed for task-fMRI, although one could potentially apply them to the resting-state condition by modeling the exogenous inputs as Fourier series (Di and Biswal, 2014).

Despite the significant technical advance of DCM in both model scope (i.e., extension to resting-state fMRI) and model complexity (i.e., use of sophisticated NMM), computational efficiency remains a major limitation. Even for the most efficient spectral DCM, inversion of a medium size network with 36 nodes takes about 20–40 h (Razi et al., 2017), limiting its application to whole-brain network and big dataset. To address this limitation, a novel variant of DCM (i.e., regression DCM) has been developed (Frässle et al., 2017, 2018, 2021a,b). Regression DCM (rDCM) converts the linear DCM equations from the time domain to the frequency domain, enabling efficient solution of differential equations using Fourier transformation. Together with other assumptions such as fixed hemodynamic response function, rDCM treats model inversion in DCM as a special case of Bayesian linear regression problem. These technical innovations have equipped rDCM with extremely high computational efficiency, which requires just a few seconds to estimate EC of a whole-brain

network with 66 nodes (Frässle et al., 2017). Later technical improvements such as sparsity constraints in rDCM enable EC estimation of larger brain-wide networks with denser connections (> 200 regions, > 40,000 connections) within only a few minutes (Frässle et al., 2018, 2021b). Together with the latest extension to resting-state fMRI (Frässle et al., 2021a), rDCM opens promising new opportunities for human connectomics. Nevertheless, current rDCM is inherently linear and has not incorporated more realistic generative NMMs, which hinders its application to fundamental neuroscience problems. Other variants of DCM include sparse DCM (Prando et al., 2019) and DCM with Wilson-Cowan-based neuronal equations (Sadeghi et al., 2020).

## The neural and measurement models

The original DCM uses a bilinear state space model (Friston et al., 2003):

$$\dot{x}(t) = [A + \sum_k B^k u_k(t)]x(t) + Cu(t) \quad (1)$$

where  $x(t)$  denotes the hidden neuronal states for multiple brain regions,  $u(t)$  represents exogenous experimental inputs and the matrix  $C$  models the influence of external inputs on neuronal activity.  $A$  is the baseline effective connectivity and  $B^k$  represents the modulation on effective connectivity due to the input  $u_k(t)$ . DCM estimates the parameters  $A$ ,  $B^k$  and  $C$  based on fMRI data. For the generative neural models of other variants of DCM, refer to related publications (Marreiros et al., 2008; Li et al., 2011; Friston et al., 2014, 2019; Frässle et al., 2017, 2018).

DCM employs a biophysical hemodynamic model to translate the regional neural activity  $x_i(t)$  to observed BOLD response  $y_i(t)$  (Friston et al., 2003). Specifically, for each region  $i$ , the fluctuating neuronal activity  $x_i(t)$  leads to a vasodilatory signal  $s_i(t)$  which is subject to self-regulation. The vasodilatory signal induces change in the blood flow  $f_i(t)$  resulting in subsequent change in blood volume  $v_i(t)$  and deoxyhemoglobin content  $q_i(t)$ . The hemodynamic model equations are given as follows (Friston et al., 2003):

$$\frac{ds_i(t)}{dt} = x_i(t) - \kappa s_i(t) - \gamma(f_i(t) - 1) \quad (2)$$

$$\frac{df_i(t)}{dt} = s_i(t) \quad (3)$$

$$\tau \frac{dv_i(t)}{dt} = f_i(t) - v_i^\alpha(t) \quad (4)$$

$$\tau \frac{dq_i(t)}{dt} = \frac{f_i(t)}{\rho} \left[ 1 - (1 - \rho)^{1/f(t)} \right] - \frac{q_i(t)}{v_i(t)} v_i^{\frac{1}{\alpha}}(t) \quad (5)$$

where  $\kappa$  is the rate of decay,  $\gamma$  is the rate of flow-dependent elimination,  $\tau$  is the hemodynamic transit time,  $\alpha$  is the Grubb's

exponent and  $\rho$  is the resting oxygen extraction fraction. The predicted BOLD response is calculated as a static non-linear function of blood volume and deoxyhemoglobin content that depends on the relative contribution of intravascular and extravascular components:

$$y_i(t) = v_0(k_1(1 - q_i(t)) + k_2(1 - q_i(t)/v_i(t)) + k_3(1 - v_i(t))) \quad (6)$$

where  $v_0$  is the resting blood volume fraction, and  $k_1$ ,  $k_2$  and  $k_3$  are the intravascular, concentration and extravascular coefficients, respectively.

## Development and implementation of dynamic causal models

Development of DCM models involves several major steps: (1) experimental design and hypothesis formulation; (2) selection of regions and extraction of fMRI-BOLD timeseries; (3) selection of DCM model depending on the fMRI modality (task or resting-state), network size and the problem of interest (i.e., one state, two-state or NMM-based DCM); (4) specification of the neural model including network connectivity and experimental inputs; (5) model estimation at subject-level; and (6) group-level analysis with Parametric Empirical Bayes (PEB; Friston et al., 2016). DCM is implemented using the SPM software package<sup>1</sup> running under MATLAB. The detailed implementation procedures can be found in two recent DCM guide papers (Zeidman et al., 2019a,b).

## Applications to cognitive neuroscience

As the predominant analysis method to compute EC, DCM has been used extensively to study cognitive problems in neuroscience. Below, we review a few examples from the huge literature that showcase the applications of DCM to understand cognitive processes including attention, perception, emotion and decision making. Cognitive information processing is regulated by two fundamental principles including functional separation and functional integration (Sporns, 2013; Deco et al., 2015; Wang et al., 2021). To investigate the basic connectivity architecture of neural structures in goal-directed attentional processing, Brázdil et al. (2007) conducted an event-related fMRI study employing the visual oddball task, one of the most widely used experimental paradigms in cognitive neuroscience. The deterministic DCM and bayes factors were applied to infer the coupling strengths among different brain regions and the parameters embodying the influence of experimental inputs on connectivity, and to compare competing neurophysiological

<sup>1</sup> <https://www.fil.ion.ucl.ac.uk/spm/software/>

models with different intrinsic connectivity. The study revealed that a bidirectional frontoparietal information flow exists [from intraparietal sulcus (IPS) to prefrontal cortex (PFC) and from anterior cingulate cortex (ACC) to PFC and the IPS] during target stimulus processing, which may indicate simultaneous activation of two distinct attentional neural systems. Compared with exteroceptive attention (i.e., inputs from external environment), the neural mechanism of interoceptive attention (i.e., the awareness and conscious focus toward physiological signals arising from the body) is much less known. To evaluate the functional role of the anterior insular cortex (AIC) in interoceptive attention, Wang et al. (2019b) applied a novel cognitive task that directed attention toward breathing rhythm and utilized DCM analysis of fMRI data to explain the potential mechanisms of interaction between AIC and other brain regions. After model inversion, random-effects Bayesian Model Selection (BMS; Stephan et al., 2009) was applied to determine the best model from 52 candidate models based on the observed data from all participants. The authors reported that interoceptive attention was associated with elevated AIC activation, increased coupling strength between AIC and somatosensory areas, and weaker coupling between the AIC and visual sensory areas. Notably, the differences in individual interoceptive accuracy can be predicted by AIC activation, suggesting the essential role of AIC in interoceptive attention.

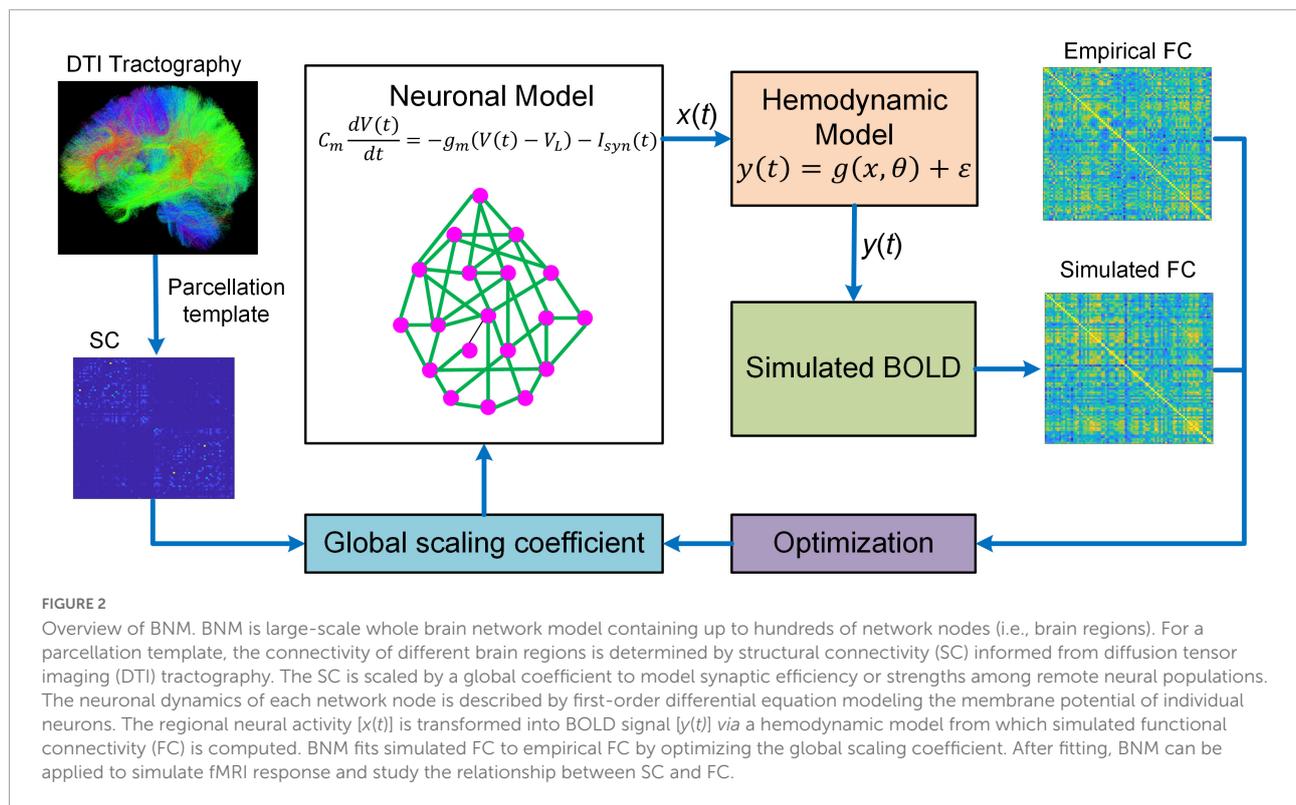
DCM has also been applied to study perceptual learning, a process of improved perceptual performance after intensive training. To examine how perceptual learning modulates the responses of decision making-related regions, Jia et al. (2018) combined psychophysics, fMRI and model-based approach, and trained participants on a motion direction discrimination task. DCM models were constructed to examine whether learning changes the EC among V3A, middle temporal area (MT), intraparietal sulcus (IPS), frontal eye field (FEF) and ventral premotor cortex (PMv). The network receives external motion inputs from both V3A and MT with motion direction (trained vs. untrained). BMS with random effect analysis (Stephan et al., 2009) was applied to test nine candidate models with different modulation assumptions. Results indicated that learning strengthened the EC on the feedforward connections from V3A to PMv and from IPS to FEF, suggesting that perceptual learning leads to decision refinement. In a recent study, Lumaca et al. (2021) employed DCM in conjunction with PEB analysis to identify modulation of brain EC during perceptual learning of complex tone patterns based on fMRI of a complex oddball paradigm. The authors found that errant responses were associated with excitation increase within the left Heschl's gyrus (HG) and left-lateralized increase in feedforward EC from the HG to the planum temporale (PT), which suggests that the prediction errors of complex auditory learning are encoded by connectivity changes in the feedforward and intrinsic pathways within the superior temporal gyrus.

In addition to task-fMRI, resting-state fMRI has been analyzed by DCM models to study the neural substrate of cognition. Esménio et al. (2019) combined functional and effective connectivity analysis to characterize the neurofunctional architecture of empathy in the default mode network (DMN). They performed resting-state fMRI scan on 42 participants who completed a questionnaire of dyadic empathy. Using spectral DCM on resting-state fMRI data, they observed that subjects with higher scores in empathy showed stronger EC from the posterior cingulate cortex (PCC) to bilateral inferior parietal lobule (IPL), and from the right IPL to the medial prefrontal cortex (mPFC). Such findings support the hypothesis that individual difference in self-perceived empathy is mediated by systematic variations in effective connectivity within the DMN, which underlie differences in FC.

## Biophysical network model

### Overview

Besides DCM, BNM is another popular generative model that has been proved to be useful in studying fMRI connectome. BNM for fMRI is a modeling framework that incorporates structural and physiological properties of brain networks, represents each network node with populations of neurons, and connects distinct nodes with long-range fibers estimated from diffusion MRI data to simulate fMRI responses (Figure 2; Honey et al., 2007, 2009; Deco and Jirsa, 2012; Deco et al., 2013a,b). BNM is typically large-scale whole brain network model comprising of up to 1,000 network nodes (Honey et al., 2007, 2009; Deco et al., 2013a,b; Sanz-Leon et al., 2015). There are two major approaches to model the neuronal dynamics of local network node. The first approach is “direct simulations” where a large number of individual neurons linked by local synaptic rules are simulated (Stephan et al., 2015). This approach is similar to detailed biophysical microcircuit modeling, though in BNM individual neurons are usually modeled by simplified spiking models instead of full-scale conductance-based compartmental models (i.e., Hodgkin and Huxley type model). The major drawbacks of “direct simulations” include heavy computational burden and large number of loosely constrained model parameters often requiring inference from animal electrophysiological data, which makes systematic exploration of parameter space and conclusive analysis infeasible (Stephan et al., 2015). An alternative method is to represent each network node with a neural mass or mean-field model of local neuronal populations, often referred to as a mean-field reduction approach (Deco and Jirsa, 2012; Deco et al., 2013b). Due to the tractability and balance between biophysical realism and model complexity, the mean-field modeling approach has become the mainstream of current BNMs of neuroimaging data.



Despite increased tractability of the mean-field reduction approach, BNM is still highly complex because of large network size and inherent non-linearity. As a result, it is extremely challenging to estimate the coupling strengths of all individual connections. Consequently, BNMs have been focusing on simulating fMRI data using SC as a proxy for synaptic weights (Honey et al., 2007, 2009; Deco et al., 2013a; Jirsa et al., 2017), estimating only one single global scaling coefficient for all inter-regional connections (Deco et al., 2013b; Wang et al., 2019a), or estimating a small subset of connection parameters typically at group-average level (Deco et al., 2014a,b; Demirtaş et al., 2019). In addition, most BNMs are optimized to fit the high-level statistics such as FC instead of raw BOLD timeseries (Deco et al., 2013a,b; Wang et al., 2019a), which may not capture the dynamic features of fMRI signals accurately.

Thankfully, substantial progress has been made in term of parameter estimation for BNMs in recent years. Using expectation-maximization (EM) approach, Wang et al. (2019a) estimated a total of 138 parameters of a large-scale dynamic mean-field model including region-specific recurrent excitation strength and subcortical input level, though only one global scaling constant was estimated for all inter-regional connections. Besides, a new SC-guided computational approach to estimate whole-brain EC has been proposed and applied to language development (Hahn et al., 2019). This new approach uses SC as initial guess for EC which is iteratively updated according to a gradient descent algorithm to maximize the

similarity between modeled FC and empirical FC. Lastly, a novel variant of BNM has been developed recently that is capable of individual EC inference in whole-brain network (Gilson et al., 2016, 2018, 2020). Different from traditional BNMs, this new framework models local neuronal dynamics with the multivariate Ornstein-Uhlenbeck (MOU) process and estimates EC by maximum likelihood, which is referred to as MOU-EC. From a dynamic systems point of view, MOU-EC corresponds to a network with linear feedback which is equivalent to the multivariate autoregressive (MAR) process, or the linearization of the non-linear Wilson-Cowan neuronal model (Wilson and Cowan, 1972; Gilson et al., 2020). Characterized by a balance between tractability and rich dynamics on parameter variations, MOU-EC offers a comprehensive new set of tools to study distributed cognition and neuropathology (Gilson et al., 2020). Nevertheless, the MOU-EC approach does not model within-region excitatory and inhibitory interactions.

In summary, BNM and DCM share similarities but also bear significant differences. At the common side, they both have a generative model for neuronal dynamics and a measurement model to convert the neuronal activity to BOLD signal, and usually integrate an optimization scheme to estimate EC. At the different note, DCM focuses on estimating individual EC at single subject level, while BNM focuses on simulating fMRI data. Based on such distinct design objectives, DCM utilizes a relatively simple bilinear neural model while BNM employs

more realistic spiking or neural mass models. Also, BNM usually utilizes SC as a proxy for synaptic efficacy or a backbone for EC, while DCM does not necessarily require structural information, though it can be used as a prior constraint (Sokolov et al., 2020). Besides, DCM is usually confined to small networks (up to tens of nodes) while BNM applies to large-scale whole brain networks (up to hundreds of nodes). Moreover, DCM uses a Bayesian framework to estimate EC which provides the full posterior probability distributions of model parameters and enables Bayesian model comparison. In contrast, the optimization schemes in BNMs are diverse and generally provide point estimates of model parameters, which do not possess the capability of automatic pruning of connections. Lastly, DCM generally fits to raw BOLD timeseries (but see spectral and regression DCM) while BNM usually fits to high-level summary statistics such as FC. It should be recognized that the latest developments of DCM for fMRI (e.g., NMM-based DCM, regression DCM) and BNM (e.g., MOU-EC) have strived to overcome the limitations of each framework, representing a convergence between DCM and BNM (Stephan et al., 2015).

## The neural model

### Spiking neuron model

For the “direct simulation” approach in BNM, individual neurons are commonly described by the classical integrate-and-fire (IF) spiking neurons (Tuckwell, 1988) connected with biophysical synapses. The membrane potential ( $V_m$ ) is governed by the following equation:

$$C_m \frac{dV_m}{dt} = -g_m (V_m - V_L) - I_{syn} + I_{Ext} \quad (7)$$

where  $C_m$  is the membrane capacitance,  $g_m$  the leak conductance,  $V_L$  the resting potential, and  $I_{Ext}$  the external input. When  $V_m$  crosses a threshold  $V_{th}$  in the upward direction, a spike is generated and the membrane potential is reset to a value  $V_{res}$  for a refractory time period  $\tau_{res}$ . The total synaptic current  $I_{syn}$  is a summation of recurrent excitatory/inhibitory inputs from local region and excitatory inputs from other brain areas; Excitatory inputs are mediated by both AMPA and NMDA receptors, while inhibitory inputs are mediated by GABA<sub>A</sub> receptors. Synaptic currents follow the following biophysical model (Deco and Jirsa, 2012; Deco et al., 2014b):

$$I_{syn}(t) = Wg_{syn}B(V_m)s(V_m - V_{syn}) \quad (8)$$

where  $W$  is the synaptic weight,  $g_{syn}$  the maximal synaptic conductance,  $s$  the gating variable, and  $V_{syn}$  is the synaptic reversal potential. The magnesium block function  $B(V_m) = 1/(1 + \exp(-0.062V_m)/3.57)$  for NMDA currents and  $B(V_m) = 1$  for AMPA and GABA<sub>A</sub> currents. The gating variable  $s$  follows the following dynamics:

For AMPA or GABA<sub>A</sub> current:

$$\frac{ds}{dt} = \frac{-s}{\tau_{AMPA/GABA}} + \sum_i \delta(t - t_i) \quad (9)$$

For NMDA current:

$$\frac{ds}{dt} = \frac{-s}{\tau_{NMDA}} + \alpha_s x(1 - s) \quad (10)$$

$$\frac{dx}{dt} = \frac{-x}{\tau_x} + \sum_i \delta(t - t_i) \quad (11)$$

where  $x$  models the neurotransmitter concentration with the rise time constant  $\tau_x$ .  $\tau_{AMPA}$ ,  $\tau_{GABA}$  and  $\tau_{NMDA}$  are the decay time constants for AMPA, GABA<sub>A</sub> and NMDA synapses, respectively,  $t_i$  is the presynaptic spike times, and  $\alpha_s$  controls the saturation properties of NMDA.

### Dynamic mean field model

The dynamic mean field model (MFM) has been frequently used in BNMs (Deco et al., 2013b, 2014a,b; Wang et al., 2019a), which is derived by mean-field reduction of the spiking neuronal network model. Each brain region is described by the following set of non-linear stochastic differential equations (Deco et al., 2013b; Wang et al., 2019a):

$$\frac{dS_i}{dt} = -\frac{S_i}{\tau_s} + r(1 - S_i)H(x_i) + \sigma v_i \quad (12)$$

$$H(x_i) = \frac{ax_i - b}{1 - \exp(-d(ax_i - b))} \quad (13)$$

$$x_i = wJ S_i + GJ \sum_j C_{ij} S_j + I \quad (14)$$

where  $S_i$ ,  $x_i$  and  $H(x_i)$  represent the average synaptic gating variable, the total input current, and the population firing rate at brain region  $i$ , respectively.  $G$  is a global scaling factor,  $J$  is the value of synaptic coupling,  $w$  is the recurrent connection strength, and  $C_{ij}$  denotes the SC between region  $i$  and region  $j$ .  $\tau_s$  and  $r$  are kinetic parameters, and  $a$ ,  $b$  and  $d$  are parameters of the function  $H(x_i)$ .  $v_i$  represents uncorrelated Gaussian noise with standard deviation  $\sigma$ . The simulated neural (synaptic) activity  $S_i$  is fed to the same hemodynamic model as DCM (Eqn. 2–6) to generate simulated BOLD timeseries.

## Development and implementation of biophysical network models

Development of BNMs includes the following major steps: (1) parcellating the brain into discrete regions; (2) extracting whole-brain fMRI-BOLD timeseries and calculating FC; (3) computing SC based on diffusion MRI data; (4) representing each network node with populations of spiking neurons or

neural-field/NMM of local neuronal populations; (5) linking individual network nodes with long-range connections based on SC; (6) transforming the network activities to simulated BOLD signals *via* a hemodynamic model; and (7) fitting model parameters to FC *via* an optimization scheme.

There are a number of computational software packages to facilitate the development and implementation of BNMs including NEURON (Carnevale and Hines, 2006), BRIAN (Goodman and Brette, 2009; Stimberg et al., 2019), NEST (Gewaltig and Diesmann, 2007), and The Virtual Brain (TVB; Jirsa et al., 2010; Sanz-Leon et al., 2013). Specifically, NEURON is the preferred simulation environment for the construction of morphologically and biophysically realistic neuronal models and networks. BRIAN and NEST are tailored for spiking network models that focus on the dynamics and structures of neural systems rather than the exact morphology and physiology of individual neurons. Different from NEURON, BRIAN and NEST that concentrate on simulation of individual neurons within small brain regions, TVB is a platform specifically designed for constructing and simulating personalized brain networks based on multimodal neuroimaging data such as fMRI, diffusion MRI, MEG and EEG. Conveniently, one could choose different neural mass or neural field models of local dynamics from TVB's predefined model classes and apply different measures of anatomical connectivity [CoCoMac or human diffusion-weighted imaging (DWI) data].

## Application to cognitive neuroscience

One fundamental question in cognitive neuroscience is how different cognitive states such as attention, sleep and wakefulness are defined mechanistically and switch from one state to the other. Existing definitions focus on resting networks and statistical description of functional activity patterns (Barttfeld et al., 2015; Tagliazucchi et al., 2016), which do not provide a mechanistic understanding of the dynamical coordination between neural systems. Recent breakthrough from computational neuroscience has filled this important gap (Deco et al., 2019; Kringelbach and Deco, 2020). Based on the concept of metastability (i.e., the ability of a system to maintain its equilibrium for an extended time in the presence of small perturbations), Deco et al. (2019) defined brain state as an ensemble of “metastable substates” each characterized by a probabilistic stability and occurrence frequency. This novel definition allows for systematic investigation of brain state transition. Using a unique fMRI and EEG dataset recorded from healthy subjects during awake and sleep conditions, Deco et al. (2019) fit a whole-brain generative BNM to the probabilistic metastable substates (PMS) space of the empirical data corresponding to the awake and sleep conditions. It was demonstrated that *in silico* stimulation predicted by the BNM can accurately force transitions between different brain states, and in particular, from deep sleep to wakefulness and vice versa.

These findings provide valuable insights how and where to induce a brain state transition using whole-brain BNM, which may potentially apply to restore the pathological brain state to normal state using external stimulation. The new brain state definition and modeling framework were recently applied to study the effect of external stimulation on functional networks in the aging healthy human brain (Escrichs et al., 2022). The authors first characterized the brain states as PMS space in two groups of subjects [middle-aged adults (age < 65) and older adults (age ≥ 65)], based on a large-cohort resting-state fMRI dataset ( $N = 310$  for each group). A whole-brain BNM was then developed and fit to the PMS and *in silico* stimulation with region-specific intensity was applied to induce transitions from the brain states of the older group to those of the middle-age group. The authors discovered that the precuneus, a brain region belonging to DMN and involved in a variety of complex functions such as episodic memory and visuospatial processing, is the best stimulation target for brain state transition. This elegant study suggests that generative models for neuroimaging data could potentially serve as gateway for the design of novel brain perturbation techniques to reverse the adverse effects of aging on cognitive functions.

Another important contribution of BNM to cognitive neuroscience is the manifestation of the dynamical origin of slowness of thought during task-based cognition (Kringelbach et al., 2015). This discovery relies on a major finding from whole-brain BNMs that the brain is not only metastable (Tognoli and Kelso, 2014), but also maximally metastable (Cabral et al., 2014; Kringelbach et al., 2015). Such dynamical property has far-reaching implications as it leads to a characteristic slowness of spontaneous dynamics when the brain network enters the state of transition, a phenomenon termed “critical slowing down” (Kringelbach et al., 2015; Meisel et al., 2015). Notably, a previous whole-brain BNM has also revealed a critical slowing down on the edge of a criticality (Deco et al., 2013b). Together these modeling studies suggest that optimal task processing requires extensive examination of the dynamical repositories of brain networks, which explains the nature of slowness of cognition. It should be stressed that such fundamental insights can only be made possible by generative neuronal modeling (Deco and Jirsa, 2012) because traditional FC analysis cannot reveal whether the brain is maximally metastable or not.

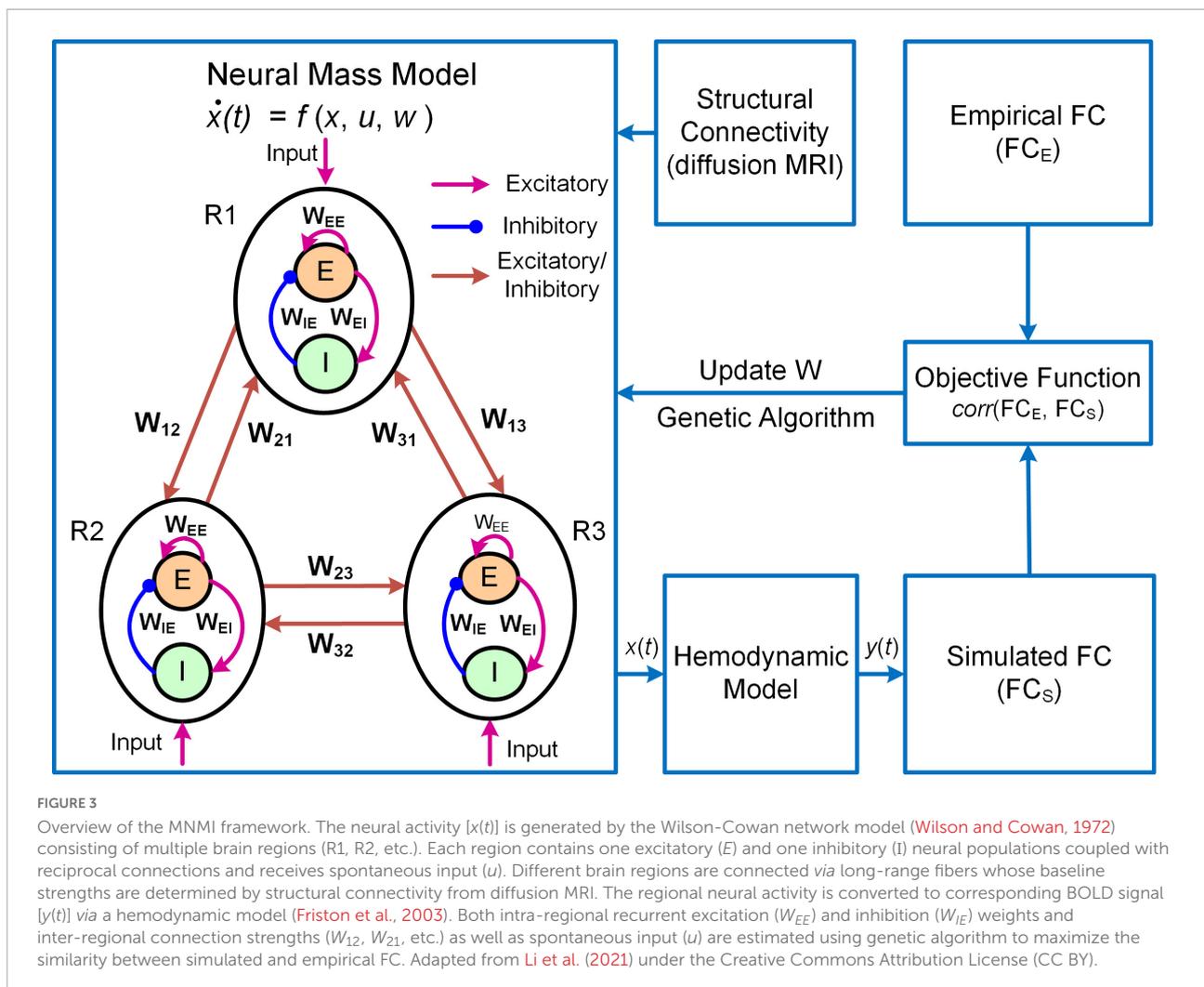
Since its introduction, the MOU-EC model (Gilson et al., 2016) has been applied to offer insights into the neural mechanisms of cognition. It has been well documented that perceptual categorization, the mapping of sensory stimuli to category labels, involves a two-stage processing hierarchy in both the visual and auditory systems (Riesenhuber and Poggio, 2000; Ashby and Spiering, 2004; Jiang et al., 2007, 2018). The first stage is a “bottom-up” stage where neurons in the sensory cortices learn to respond to stimulus features while the second stage is a “top-down” stage where neurons in higher cortical areas learn to classify the stimulus-selective inputs from the first stage. To investigate whether the two-stage

processing hierarchy also exists in the somatosensory system, Malone et al. (2019) designed an experiment where human participants, after training to label vibrotactile stimuli presented to their right forearm, underwent an fMRI scan while actively engaging in categorizing the stimuli. The authors first applied representational similarity analysis to identify the stimulus- and category- selective areas. Next they utilized the MOU-EC method (Gilson et al., 2016) to estimate whole-brain EC among 200 regions. It was observed that the influence (i.e., effective drive) from most of the category-selective areas to the stimulus-selective areas was much higher than that in the opposite direction. These findings support the two-stage processing hierarchy in the somatosensory system, providing a unified computational principle for perceptual categorization. The MOU-EC model has also been used to study the neural mechanisms underlying the engagement of functionally specialized brain regions, and in particular, how brain network connectivity is modulated under different cognitive conditions to give rise to differential regional dynamics (Gilson et al., 2018).

## Dynamic Neural Model with direct parameterization

### Overview

In addition to DCM and BNM, the two major types of generative models for fMRI, there exist other modeling frameworks that estimate EC and study connectome in a generative fashion (Havlicek et al., 2011; Olier et al., 2013; Fukushima et al., 2015; Becker et al., 2018; Singh et al., 2020; Li et al., 2021). One framework that shows promising application to cognitive and clinical neuroscience is Dynamic Neural Model (DNM) with direct parameterization. This type of modeling approach attempts to balance the complexity of BNM with the identifiability of DCM so that the model is both sufficiently realistic and equipped with the capability to efficiently estimate individual connections at single subject level. Notably, DNM utilizes non-linear dynamic models or NMMs to model local



neuronal population activity, thus able to capture the extended patterns of the spatiotemporal dynamics of brain networks.

One recent development of DNM is termed Mesoscale Individualized Neurodynamic (MINDy) modeling that fits non-linear dynamical neural models directly to fMRI data (Singh et al., 2020). To achieve computational efficiency, MINDy models regional neural population activity in the fMRI scale (instead of neuronal firing scale) using an abstracted NMM. After converting the continuous-time neural model to a discrete-time analog, an objective function is formed by summing up the one-step prediction errors of empirical BOLD timeseries, which allows the calculation of error gradients analytically. Using a computationally efficient optimization algorithm [Nesterov-Accelerated Adaptive Moment Estimation (NADAM, Dozat, 2016)], model parameters are estimated to minimize the objective function. The high computational efficiency enables identification of EC parameters in large-scale networks with hundreds of nodes in just 1–3 min per subject, making it ideally suitable for big dataset application. Notably, MINDy requires no prior anatomical constraints or long-term summary statistics for model inversion and provides estimation of individual connection parameters in an individualized and efficient manner, representing a significant departure from existing methods.

Another example of DNM with direct parameterization is a framework termed Multiscale Neural Model Inversion (MNMI, Figure 3; Li et al., 2021). In MNMI, the neuronal activity is generated by a neural mass network model comprised of multiple brain regions. Each region contains one excitatory and one inhibitory neural populations coupled with reciprocal connections, and the intrinsic dynamics are described by the classical Wilson-Cowan model (Wilson and Cowan, 1972). The excitatory neural populations within different brain regions are connected *via* long-range fibers whose baseline connection strengths are determined by SC from diffusion MRI; weak inter-regional connections are removed to construct sparse networks and avoid the problem of over-parameterization. The regional neural activity is converted to corresponding BOLD signal *via* a hemodynamic model (Friston et al., 2003) and FC is computed using Pearson's correlation. MNMI then applies genetic algorithm, a biologically inspired optimization algorithm, to estimate both intra-regional and inter-regional coupling strengths to minimize the difference between simulated and empirical FC. Similar to MINDy, MNMI estimates EC for individual connections at single subject level, thus offering personalized assessment. MNMI differs from MINDy in several aspects including using a more biologically informed NMM with excitatory and inhibitory interactions, modeling neural activity in neuronal firing scale (vs. fMRI scale), utilizing structural information to constrain EC estimation, and applying genetic algorithm (vs. gradient descent algorithm) to estimate model parameters. One limitation of MNMI is its heavy computational burden, which prevents

large-scale whole-brain network applications. Nevertheless, MNMI provides a multiscale modeling framework to fathom deeper brain connectivity features in health and disease.

## The neural model

Neural mass models are commonly used to model local neuronal dynamics in DNM. For instance, MINDy and MNMI use one-state and two-state NMMs, respectively. The NMM in MNMI is described by the following differential equations (Wilson and Cowan, 1972; Li et al., 2021):

$$\tau_e \frac{dE_j(t)}{dt} = -E_j(t) + S \left( \sum_k W_{kj} C_{kj} E_k(t) + W_{EE}^j E_j(t) - W_{IE}^j I_j(t) + u + \varepsilon(t) \right) \quad (15)$$

$$\tau_i \frac{dI_j(t)}{dt} = -I_j(t) + S \left( W_{EI}^j E_j(t) + \varepsilon(t) \right) \quad (16)$$

where  $E_j$  and  $I_j$  are the mean firing rates of excitatory and inhibitory neural populations in brain region  $j$ ,  $\tau_e$  and  $\tau_i$  are the excitatory and inhibitory time constants,  $W_{EE}^j$ ,  $W_{EI}^j$  and  $W_{IE}^j$  are the local coupling strengths from excitatory to excitatory neural population, from excitatory to inhibitory neural population and from inhibitory to excitatory neural population, respectively. The variable  $u$  is a constant external input representing average extrinsic drive from other un-modeled brain regions, and  $\varepsilon(t)$  is random additive noise following a normal distribution. The long-range connectivity strength from region  $k$  to region  $j$  is represented by  $W_{kj}$  scaled by empirical SC ( $C_{kj}$ ). The non-linear response function  $S$  is modeled as a sigmoid function  $S = 1/(1 + e^{-(\frac{x-\mu}{\sigma})})$ . MNMI estimates connection parameters  $W_{kj}$ ,  $W_{EE}^j$ ,  $W_{EI}^j$ , and input  $u$  based on individual SC and FC.

## Development and implementation of Dynamic Neural Models

The development steps of DNMs are generally similar to BNMs: (1) parcellating the brain into discrete regions; (2) selecting brain regions to model and extracting fMRI-BOLD timeseries; (3) calculating FC and/or SC if necessary; (4) representing each network node with a NMM of local neuronal populations; (5) linking individual network nodes with long-range connections with or without SC constraint; (6) deconvolving the empirical BOLD signals with a hemodynamic response function (HRF, Singh et al., 2020) or transforming the network activities to simulated BOLD signals *via* a hemodynamic model (Li et al., 2021); and (7) fitting model parameters to deconvolved BOLD signals or FC *via* an

optimization scheme. Due to the flexibility in neuronal model, SC utilization, choice of objective function and optimization algorithm, DNMs are generally implemented using customized scripts with MATLAB or other computing languages (e.g., C++).

## Application to clinical neuroscience

DNM has been applied to study the circuit mechanisms of major depressive disorder (MDD). MDD is a leading cause of chronic disability worldwide with a lifetime prevalence of up to 17% (Kessler et al., 2005), but the underlying pathophysiological mechanisms remain elusive. Functional connectome analysis indicates that MDD can be characterized as a disorder with dysfunctional connectivity and regulation among multiple resting-state networks including the DMN, salience network, executive control network and limbic network (Menon, 2011; Dutta et al., 2014; Mulders et al., 2015; Drysdale et al., 2017). However, two important questions remain unresolved. First, it is not clear which functional networks play a central role and which functional networks play a subordinate role in MDD pathogenesis. Second, it is unclear whether the dysconnectivity or dysregulation originates from limbic or cortical system and whether such dysregulation results from intrinsic (intra-regional) or extrinsic (inter-regional) mechanisms. Answering these two questions is not only important for deeper mechanistic understanding of MDD pathology but also necessary for more targeted treatments.

The MNMI framework is well suited to address these questions due to its biological realism utilizing a physiologically informed NMM and its multiscale nature incorporating both intra-regional and inter-regional neural interactions. By applying the MNMI framework to a large sample-size resting-state fMRI dataset consisting of 100 MDDs and 100 normal control (NC) healthy subjects, Li et al. (2021) demonstrated that MDD pathology is more likely caused by aberrant circuit interactions and dynamics within a core “executive-limbic” network rather than the “default mode-salience” network, consistent with the long-standing hypothesis of limbic-cortical dysregulation in MDD (Mayberg, 1997, 2002, 2003; Davidson et al., 2002; Disner et al., 2011). Notably, model results indicated that both limbic and cortical systems and both intra-regional and inter-regional connectivities could play a role in MDD pathology. Specifically, MNMI analysis showed that recurrent inhibition within the amygdala was abnormally decreased and the excitatory EC from the superior parietal cortex (SPC) to the amygdala was abnormally increased, which may underlie hyperactivity of the amygdala in MDD (Drevets, 2001; Siegle et al., 2002), leading to increased anxiety and cognitive bias over negative stimuli (Figure 4; Disner et al., 2011). In addition, the EC from the SPC to the dorsolateral prefrontal cortex (dlPFC) switched from excitation

in NC to inhibition in MDD, which well explains dlPFC hypoactivity (Fales et al., 2008; Hamilton et al., 2012), resulting in deficit cognitive control (Figure 4; Disner et al., 2011). The model also revealed other abnormal connectivity patterns in MDD including elevated recurrent excitation in the SPC, reduced SPC inhibition on the thalamus and decreased dlPFC excitation on the hippocampus, which may underlie biased attention for negative stimuli, abnormal brain oscillations and impaired memory function, respectively (Figure 4; Disner et al., 2011; Li et al., 2021). Overall, by employing a biologically plausible NMM, the MNMI framework provides a mechanistic account of circuit dysfunction in MDD which highlights the importance of targeting the executive-limbic system for maximal therapeutic benefits.

## Summary and future direction

Driven by the rapid advances in non-invasive neuroimaging techniques, the young emerging field of human connectomics has made significant accomplishments in characterizing the large-scale organizational features of both structural and functional brain networks. Notwithstanding, the potential of connectomics to answer fundamental neuroscience and clinical questions has yet to bring into full play. To achieve such important goal, computational connectomics has moved beyond the anatomical and statistical description of connectivity to more mechanistic formulation of the neural processes underlying neuroimaging data (i.e., mechanistic connectome). Mechanistic connectome based on generative modeling of fMRI offers a natural and principled tool to link microscopic or mesoscopic neural process with macroscopic BOLD dynamics, which enables mechanistic understanding of brain cognitive functions in health and dysfunction in diseases. It is important to note that, unlike static structural connectome, mechanistic connectome based on effective connectivity is parameterized by the state of the brain. That is, one would obtain a very different mechanistic connectome when the task demands change, or when the stimuli challenge, or when endogenous activity switches to a different state (e.g., inward vs. outwardly directed attention) (Friston et al., 2003; Jung et al., 2018; Park et al., 2021). Thus, there will not be a conventional atlas for the mechanistic connectome like one could obtain for the structural connectome.

Recent years have seen tremendous development and expansion of generative model-based connectome analysis toolsets. Two well-established and widely used modeling frameworks include DCM and BNM, which represent two approaches at the opposite end of biological realism and estimation tractability. Specifically, while DCM allows estimation of full connection parameters at individual subject level, the physiological interpretability for model parameters is limited due to the abstract bilinear state model. On the other

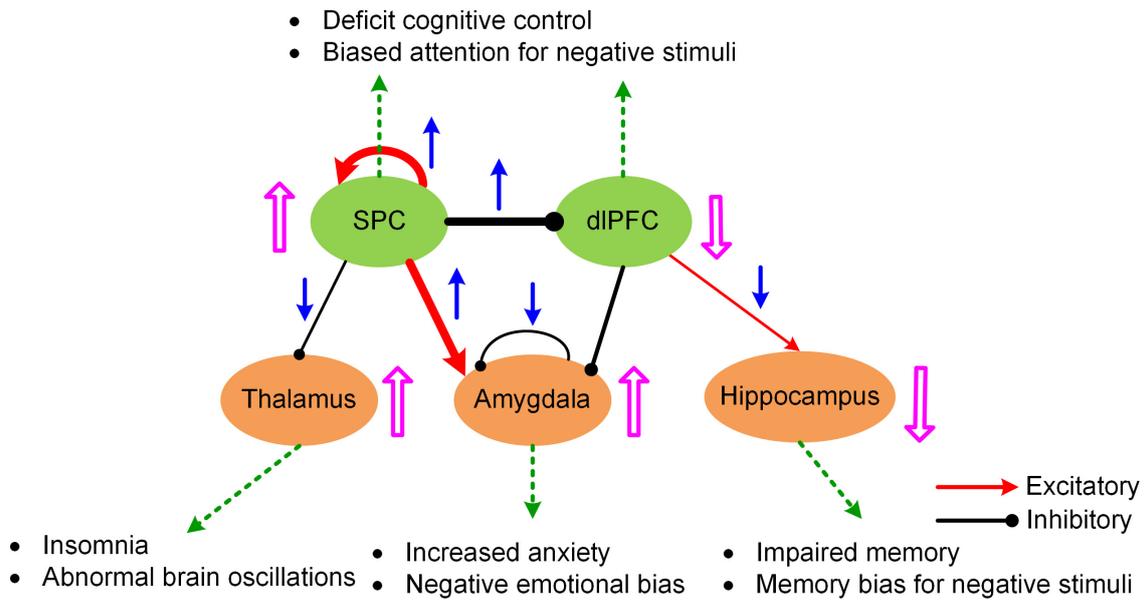


FIGURE 4

A hypothetical model of executive-limbic malfunction in MDD. MDD is mediated by increased recurrent excitation in the superior parietal cortex (SPC) and greater inhibition from the SPC to the dorsolateral prefrontal cortex (dlPFC), leading to increased SPC activity and decreased dlPFC response, which may underlie deficit cognitive control and biased attention for negative stimuli. In addition, the excitatory drive from the SPC to the amygdala is abnormally elevated in MDD. Combined with reduced recurrent inhibition, the amygdala shows hyperactivity which causes increased anxiety and biased processing of negative stimuli. Besides, the inhibitory drive from the SPC to the thalamus is reduced while the excitatory projection from the dlPFC to the hippocampus is abnormally decreased in MDD. The former change could result in abnormal brain oscillations and insomnia while the latter change could account for impaired memory function and biased memory for negative stimuli. The blue arrows indicate the change of the connection strengths in MDD from normal control. The pink UP/DOWN arrows next to the brain regions indicate the change in neural responses in MDD compared to normal control. Adapted from Li et al. (2021) under the Creative Commons Attribution License (CC BY).

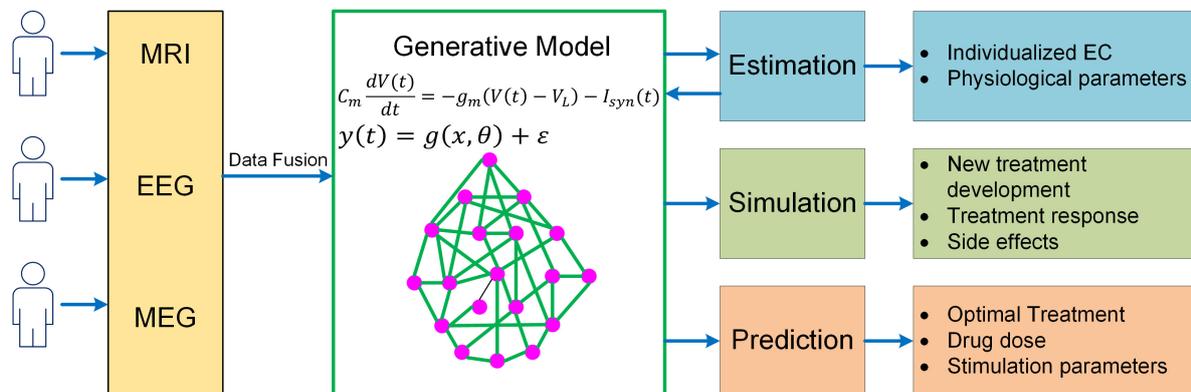


FIGURE 5

A unified mechanistic pipeline for generative model-based neuroimaging analysis to treat brain disorders. In this analysis pipeline, individual subjects first undergo multiple neuroimaging scans/recordings such as MRI, EEG and MEG. The multi-modal neuroimaging data are then combined with data fusion and fed into the generative model to estimate individualized EC and other relevant physiological parameters such as neuromodulatory levels. The estimated model parameters are then fed back into the generative model to simulate existing treatment response and new treatment development as well as their side effects. Based on the simulation outcome, the generative model will predict optimal treatment strategy for the patient along with drug dose or stimulation parameters.

hand, though BNM incorporates more physiologically grounded neuronal models for fMRI generation, its identifiability is limited to one or a small subset of parameters often at the

group-average level. Effort to combine the advantages of DCM with BNM has led to the development of a different type of modeling framework that can be categorized as Dynamic

Neural Model (DNM) with direct parameterization. DNM seeks to model population-level neuronal dynamics accurately with biophysically plausible NMMs and estimate physiologically meaningful parameters for individual connections at single subject level. It should be noted that the boundary among these three types of generative models is diminishing with the latest developments of DCM and BNM which utilize more biophysically informed models and are equipped with the capability for efficient EC estimation of large-scale networks at both individual connection and individual subject levels. One should expect the convergence of DCM, BNM and DNM continues in the future.

While much progress has been made, more needs to be done to meet the challenges in neuroscience. It should be recognized that even the most sophisticated BNM is only a highly simplified representation of the human brain, yet more complex models would make parameter inference much more difficult, raising the question of how to determine the right level of complexity in generative modeling. One important rule of thumb is that models should only be considered that are in the right ballpark of complexity to address the question at hand. That is, they need to have parameters relating to the quantities of interest (interpretability), while not being more complex than the data can accommodate (given the limited resolution of fMRI). This principle has been well implemented in DCM *via* Bayesian model selection, a process where different candidates of models are iteratively generated and compared to reach the models that have the optimal level of complexity (Stephan et al., 2009; Rosa et al., 2012). Specifically, the optimal model optimizes the trade-off between accuracy and complexity, which is quantified by the log model evidence (i.e.,  $\log p(y|m)$ ; Zeidman et al., 2019a,b). The topic of complexity deserves more consideration in future generative modeling studies given the need to explore more physiologically based neuroscience questions (e.g., neuromodulatory effects on cognitive functions). Also, to have a thorough understanding of the neural mechanisms of cognition, a truly multiscale model is wanted which has the capability to link cellular, circuit, network and system dynamics with behavioral response. Moreover, to enable more accurate estimation of model parameters, generative models need to integrate multimodal neuroimaging data (e.g., fMRI, MEG, and EEG) into a unified framework. Lastly, in order

to apply to clinical interventions, generative models should be able to explore new treatment paradigms such as non-invasive brain stimulation for brain disorders, predict optimal personalized treatment and simulate the treatment outcome (Figure 5). Addressing such grand challenges will lead to a new class of generative models for neuroimaging data that not only revolutionizes the field of human connectomics but also significantly advances our understanding of the human brain and neuropsychiatric disorders.

## Author contributions

GL and P-TY contributed to the conception of the work, revised the manuscript, and approved the final version of the manuscript. GL drafted the manuscript. Both authors contributed to the article and approved the submitted version.

## Funding

This work was supported in part by the United States National Institutes of Health (NIH) (grant nos. EB008374, MH125479, and EB006733).

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## References

- Ashby, F. G., and Spiering, B. J. (2004). The neurobiology of category learning. *Behav. Cogn. Neurosci. Rev.* 3, 101–113.
- Baliyan, V., Das, C. J., Sharma, R., and Gupta, A. K. (2016). Diffusion weighted imaging: Technique and applications. *World J. Radiol.* 8, 785–798.
- Barttfeld, P., Uhrig, L., Sitt, J. D., Sigman, M., Jarraya, B., and Dehaene, S. (2015). Signature of consciousness in the dynamics of resting-state brain activity. *Proc. Natl. Acad. Sci. U.S.A.* 112, 887–892.
- Becker, C. O., Bassett, D. S., and Preciado, V. M. (2018). Large-scale dynamic modeling of task-fMRI signals via subspace system identification. *J. Neural Eng.* 15:066016. doi: 10.1088/1741-2552/aad8c7
- Bolton, T. A. W., Tarun, A., Sterpenich, V., Schwartz, S., and Van De Ville, D. (2018). Interactions between large-scale functional brain networks are captured by sparse coupled HMMs. *IEEE Trans. Med. Imag.* 37, 230–240. doi: 10.1109/TMI.2017.2755369

- Braun, U., Schaefer, A., Betzel, R. F., Tost, H., Meyer-Lindenberg, A., and Bassett, D. S. (2018). From maps to multi-dimensional network mechanisms of mental disorders. *Neuron* 97, 14–31. doi: 10.1016/j.neuron.2017.11.007
- Brzdil, M., Mikl, M., Marecek, R., Krupa, P., and Rektor, I. (2007). Effective connectivity in target stimulus processing: A dynamic causal modeling study of visual oddball task. *Neuroimage* 35, 827–835. doi: 10.1016/j.neuroimage.2006.12.020
- Breakspear, M. (2017). Dynamic models of large-scale brain activity. *Nat. Neurosci.* 20, 340–352.
- Cabral, J., Luckhoo, H., Woolrich, M., Joensson, M., Mohseni, H., Baker, A., et al. (2014). Exploring mechanisms of spontaneous functional connectivity in MEG: How delayed network interactions lead to structured amplitude envelopes of band-pass filtered oscillations. *Neuroimage* 90, 423–435. doi: 10.1016/j.neuroimage.2013.11.047
- Carnevale, N. T., and Hines, M. L. (2006). *The NEURON Book*. Cambridge, UK: Cambridge University Press.
- Chow, M. S., Wu, S. L., Webb, S. E., Gluskin, K., and Yew, D. T. (2017). Functional magnetic resonance imaging and the brain: A brief review. *World J. Radiol.* 9, 5–9. doi: 10.4329/wjr.v9.i1.5
- Craddock, R. C., Jbabdi, S., Yan, C. G., Vogelstein, J. T., Castellanos, F. X., Di Martino, A., et al. (2013). Imaging human connectomes at the macroscale. *Nat. Methods* 10, 524–539.
- Dauwizeau, J., David, O., and Stephan, K. E. (2011). Dynamic causal modelling: A critical review of the biophysical and statistical foundations. *Neuroimage* 58, 312–322. doi: 10.1016/j.neuroimage.2009.11.062
- Davidson, R. J., Pizzagalli, D., Nitschke, J. B., and Putnam, K. (2002). Depression: Perspectives from affective neuroscience. *Annu. Rev. Psychol.* 53, 545–574.
- Deco, G., Cruzat, J., Cabral, J., Tagliazucchi, E., Laufs, H., Logothetis, N. K., et al. (2019). Awakening: Predicting external stimulation to force transitions between different brain states. *Proc. Natl. Acad. Sci. U.S.A.* 116, 18088–18097. doi: 10.1073/pnas.1905534116
- Deco, G., and Jirsa, V. K. (2012). Ongoing cortical activity at rest: Criticality, multistability, and ghost attractors. *J. Neurosci.* 32, 3366–3375. doi: 10.1523/JNEUROSCI.2523-11.2012
- Deco, G., Jirsa, V. K., and McIntosh, A. R. (2013a). Resting brains never rest: Computational insights into potential cognitive architectures. *Trends Neurosci.* 36, 268–274.
- Deco, G., Ponce-Alvarez, A., Mantini, D., Romani, G. L., Hagmann, P., and Corbetta, M. (2013b). Resting-state functional connectivity emerges from structurally and dynamically shaped slow linear fluctuations. *J. Neurosci.* 33, 11239–11252. doi: 10.1523/JNEUROSCI.1091-13.2013
- Deco, G., and Kringelbach, M. L. (2014). Great expectations: Using whole-brain computational connectomics for understanding neuropsychiatric disorders. *Neuron* 84, 892–905. doi: 10.1016/j.neuron.2014.08.034
- Deco, G., McIntosh, A. R., Shen, K., Hutchison, R. M., Menon, R. S., Everling, S., et al. (2014a). Identification of optimal structural connectivity using functional connectivity and neural modeling. *J. Neurosci.* 34, 7910–7916.
- Deco, G., Ponce-Alvarez, A., Hagmann, P., Romani, G. L., Mantini, D., and Corbetta, M. (2014b). How local excitation-inhibition ratio impacts the whole brain dynamics. *J. Neurosci.* 34, 7886–7898.
- Deco, G., Tononi, G., Boly, M., and Kringelbach, M. L. (2015). Rethinking segregation and integration: Contributions of whole-brain modelling. *Nat. Rev. Neurosci.* 16, 430–439. doi: 10.1038/nrn3963
- Demirtaş, M., Burt, J. B., Helmer, M., Ji, J. L., Adkinson, B. D., Glasser, M. F., et al. (2019). Hierarchical heterogeneity across human cortex shapes large-scale neural dynamics. *Neuron* 101, 1181–1194. doi: 10.1016/j.neuron.2019.01.017
- Deneve, S., and Machens, C. K. (2016). Efficient codes and balanced networks. *Nat. Neurosci.* 19, 375–382.
- Di, X., and Biswal, B. B. (2014). Identifying the default mode network structure using dynamic causal modeling on resting-state functional magnetic resonance imaging. *Neuroimage* 86, 53–59.
- Disner, S. G., Beevers, C. G., Haigh, E. A., and Beck, A. T. (2011). Neural mechanisms of the cognitive model of depression. *Nat. Rev. Neurosci.* 12, 467–477.
- Dozat, T. (2016). “Incorporating nesterov momentum into adam,” in *Proceedings of 4<sup>th</sup> International Conference on Learning Representations*, (San Juan, Puerto Rico: Workshop Track).
- Drevets, W. C. (2001). Neuroimaging and neuropathological studies of depression: Implications for the cognitive-emotional features of mood disorders. *Curr. Opin. Neurobiol.* 11, 240–249. doi: 10.1016/s0959-4388(00)00203-8
- Drysdale, A. T., Grosenick, L., Downar, J., Dunlop, K., Mansouri, F., Meng, Y., et al. (2017). Resting-state connectivity biomarkers define neurophysiological subtypes of depression. *Nat. Med.* 23, 28–38.
- Durstewitz, D., Seamans, J. K., and Sejnowski, T. J. (2000). Dopamine-mediated stabilization of delay-period activity in a network model of prefrontal cortex. *J. Neurophysiol.* 83, 1733–1750. doi: 10.1152/jn.2000.83.3.1733
- Dutta, A., McKie, S., and Deakin, J. F. (2014). Resting state networks in major depressive disorder. *Psychiatry Res. Neuroimaging* 224, 139–151.
- Elam, J. S., Glasser, M. F., Harms, M. P., Sotiropoulos, S. N., Andersson, J. L. R., Burgess, G. C., et al. (2021). The Human Connectome Project: A retrospective. *Neuroimage* 244:118543.
- Escrachs, A., Sanz Perl, Y., Martínez-Molina, N., Biarnes, C., Garre-Olmo, J., Fernández-Real, J. M., et al. (2022). The effect of external stimulation on functional networks in the aging healthy human brain. *Cereb. Cortex* 21:bhac064.
- Esmeñio, S., Soares, J. M., Oliveira-Silva, P., Zeidman, P., Razi, A., Gonçalves Óf, Friston, K., et al. (2019). Using resting-state DMN effective connectivity to characterize the neurofunctional architecture of empathy. *Sci. Rep.* 9:2603. doi: 10.1038/s41598-019-38801-6
- Fales, C. L., Barch, D. M., Rundle, M. M., Mintun, M. A., Snyder, A. Z., Cohen, J. D., et al. (2008). Altered emotional interference processing in affective and cognitive-control brain circuitry in major depression. *Biol. Psychiatry* 63, 377–384. doi: 10.1016/j.biopsych.2007.06.012
- Felleman, D. J., and Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex* 1, 1–47.
- Frässle, S., Harrison, S. J., Heinze, J., Clementz, B. A., Tamminga, C. A., Sweeney, J. A., et al. (2021a). Regression dynamic causal modeling for resting-state fMRI. *Hum. Brain Mapp.* 42, 2159–2180.
- Frässle, S., Manjaly, Z. M., Do, C. T., Kasper, L., Pruessmann, K. P., and Stephan, K. E. (2021b). Whole-brain estimates of directed connectivity for human connectomics. *NeuroImage* 225:117491.
- Frässle, S., Lomakina, E. I., Kasper, L., Manjaly, Z. M., Leff, A., Pruessmann, K. P., et al. (2018). A generative model of whole-brain effective connectivity. *Neuroimage* 179, 505–529.
- Frässle, S., Lomakina, E. I., Razi, A., Friston, K. J., Buhmann, J. M., and Stephan, K. E. (2017). Regression DCM for fMRI. *Neuroimage* 155, 406–421.
- Frässle, S., Paulus, F. M., Krach, S., Schweinberger, S. R., Stephan, K. E., and Jansen, A. (2016). Mechanisms of hemispheric lateralization: Asymmetric interhemispheric recruitment in the face perception network. *NeuroImage* 124, 977–988. doi: 10.1016/j.neuroimage.2015.09.055
- Friston, K. J. (2011). Functional and effective connectivity: A review. *Brain Connect.* 1, 13–36.
- Friston, K. J., Bastos, A., Litvak, V., Stephan, E. K., Fries, P., and Moran, R. J. (2012). DCM for complex-valued data: Cross-spectra, coherence and phase-delays. *Neuroimage* 59, 439–455. doi: 10.1016/j.neuroimage.2011.07.048
- Friston, K. J., Harrison, L., and Penny, W. (2003). Dynamic causal modeling. *Neuroimage* 19, 1273–1302.
- Friston, K. J., Kahan, J., Biswal, B., and Razi, A. A. (2014). A DCM for resting state fMRI. *Neuroimage* 94, 396–407.
- Friston, K. J., Litvak, V., Oswal, A., Razi, A., Stephan, K. E., van Wijk, B. C. M., et al. (2016). Bayesian model reduction and empirical Bayes for group (DCM) studies. *Neuroimage* 128, 413–431. doi: 10.1016/j.neuroimage.2015.11.015
- Friston, K. J., Mattout, J., Trujillo-Barreto, N., Ashburner, J., and Penny, W. (2007). Variational free energy and the Laplace approximation. *Neuroimage* 34, 220–234.
- Friston, K. J., Preller, K. H., Mathys, C., Cagnan, H., Heinze, J., Razi, A., et al. (2019). Dynamic causal modelling revisited. *Neuroimage* 199, 730–744.
- Froemke, R. C. (2015). Plasticity of cortical excitatory-inhibitory balance. *Annu. Rev. Neurosci.* 38, 195–219.
- Fukushima, M., Yamashita, O., Knosche, T. R., and Sato, M. A. (2015). MEG source reconstruction based on identification of directed source interactions on whole-brain anatomical networks. *Neuroimage* 105, 408–427. doi: 10.1016/j.neuroimage.2014.09.066
- Gewaltig, M., and Diesmann, M. (2007). NEST (NEural Simulation Tool). *Scholarpedia* 2:1430.
- Gilson, M., Deco, G., Friston, K. J., Hagmann, P., Mantini, D., Betti, V., et al. (2018). Effective connectivity inferred from fMRI transition dynamics during movie viewing points to a balanced reconfiguration of cortical interactions. *Neuroimage* 180, 534–546. doi: 10.1016/j.neuroimage.2017.09.061
- Gilson, M., Moreno-Bote, R., Ponce-Alvarez, A., Ritter, P., and Deco, G. (2016). Estimation of directed effective connectivity from fMRI functional connectivity

- hints at asymmetries of cortical connectome. *PLoS Comput. Biol.* 12:e1004762. doi: 10.1371/journal.pcbi.1004762
- Gilson, M., Zamora-López, G., Pallarés, V., Adhikari, M. H., Senden, M., Campo, A. T., et al. (2020). Model-based whole-brain effective connectivity to study distributed cognition in health and disease. *Netw. Neurosci.* 4, 338–373. doi: 10.1162/netn\_a\_00117
- Glover, G. H. (2011). Overview of functional magnetic resonance imaging. *Neurosurg. Clin. N. Am.* 22, 133–139.
- Goodman, D. F., and Brette, R. (2009). The Brian simulator. *Front. Neurosci.* 3, 192–197. doi: 10.3389/neuro.01.026.2009
- Gore, J. C. (2003). Principles and practice of functional MRI of the human brain. *J. Clin. Investig.* 112, 4–9.
- Hahn, G., Skeide, M. A., Mantini, D., Ganzetti, M., Destexhe, A., Friederici, A. D., et al. (2019). A new computational approach to estimate whole-brain effective connectivity from functional and structural MRI, applied to language development. *Sci. Rep.* 9:8479. doi: 10.1038/s41598-019-44909-6
- Hamilton, J. P., Etkin, A., Furman, D. J., Lemus, M. G., Johnson, R. F., and Gotlib, I. H. (2012). Functional neuroimaging of major depressive disorder: A meta-analysis and new integration of baseline activation and neural response data. *Am. J. Psychiatry* 169, 693–703. doi: 10.1176/appi.ajp.2012.11071105
- Hass, J., Hertäg, L., and Durstewitz, D. (2016). A detailed data-driven network model of prefrontal cortex reproduces key features of in vivo activity. *PLoS Comput. Biol.* 12:e1004930. doi: 10.1371/journal.pcbi.1004930
- Hasselmo, M. E., Schnell, E., and Barkai, E. (1995). Dynamics of learning and recall at excitatory recurrent synapses and cholinergic modulation in rat hippocampal region CA3. *J. Neurosci.* 15, 5249–5262. doi: 10.1523/JNEUROSCI.15-07-05249.1995
- Havlicek, M., Friston, K. J., Jan, J., Brazdil, M., and Calhoun, V. D. (2011). Dynamic modeling of neuronal responses in fMRI using cubature Kalman filtering. *Neuroimage* 56, 2109–2128.
- Hodgkin, A. L., and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiology* 117, 500–544.
- Honey, C. J., Kötter, R., Breakspear, M., and Sporns, O. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *Proc. Natl. Acad. Sci. U.S.A.* 104, 10240–10245.
- Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J. P., Meuli, R., et al. (2009). Predicting human resting-state functional connectivity from structural connectivity. *Proc. Natl. Acad. Sci. U.S.A.* 106, 2035–2040.
- Jia, K., Xue, X., Lee, J. H., Fang, F., Zhang, J., and Li, S. (2018). Visual perceptual learning modulates decision network in the human brain: The evidence from psychophysics, modeling, and functional magnetic resonance imaging. *J. Vis.* 18:9. doi: 10.1167/18.12.9
- Jiang, X., Bradley, E., Rini, R. A., Zeffiro, T., VanMeter, J., and Riesenhuber, M. (2007). Categorization training results in shape- and category-selective human neural plasticity. *Neuron* 53, 891–903. doi: 10.1016/j.neuron.2007.02.015
- Jiang, X., Chevillet, M. A., Rauschecker, J. P., and Riesenhuber, M. (2018). Training humans to categorize monkey calls: Auditory feature- and category-selective neural tuning changes. *Neuron* 98, 405–416. doi: 10.1016/j.neuron.2018.03.014
- Jirsa, V. K., Proix, T., Perdikis, D., Woodman, M. M., Wang, H., Gonzalez-Martinez, J., et al. (2017). The virtual epileptic patient: Individualized whole-brain models of epilepsy spread. *Neuroimage* 145, 377–388.
- Jirsa, V. K., Sporns, O., Breakspear, M., Deco, G., and McIntosh, A. R. (2010). Towards the virtual brain: Network modeling of the intact and the damaged brain. *Arch. Ital. Biol.* 148, 189–205.
- Jung, K., Friston, K. J., Pae, C., Choi, H. H., Tak, S., Choi, Y. K., et al. (2018). Effective connectivity during working memory and resting states: A DCM study. *Neuroimage* 169, 485–495. doi: 10.1016/j.neuroimage.2017.12.067
- Just, M. A., and Varma, S. (2007). The organization of thinking: What functional brain imaging reveals about the neuroarchitecture of complex cognition. *Cogn. Affect. Behav. Neurosci.* 7, 153–191. doi: 10.3758/cabn.7.3.153
- Karahanoglu, F. I., and Van De Ville, D. (2017). Dynamics of large-scale fMRI networks: Deconstruct brain activity to build better models of brain function. *Curr. Opin. Biomed. Eng.* 3, 28–36.
- Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., and Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of DSM-IV disorders in the National Comorbidity Survey Replication. *Arch. Gen. Psychiatry* 62, 593–602.
- Kringelbach, M. L., and Deco, G. (2020). Brain states and transitions: Insights from computational neuroscience. *Cell Rep.* 32:108128.
- Kringelbach, M. L., McIntosh, A. R., Ritter, P., Jirsa, V. K., and Deco, G. (2015). The rediscovery of slowness: Exploring the timing of cognition. *Trends Cogn. Sci.* 19, 616–628. doi: 10.1016/j.tics.2015.07.011
- Li, B., Daunizeau, J., Stephan, K. E., Penny, W., Hu, D., and Friston, K. J. (2011). Generalized filtering and stochastic DCM for fMRI. *Neuroimage* 58, 442–457.
- Li, G., and Cleland, T. A. (2018). Generative Biophysical modeling of dynamical networks in the olfactory system. *Methods Mol. Biol.* 1820, 265–288.
- Li, G., Liu, Y., Zheng, Y., Wu, Y., Li, D., Liang, X., et al. (2021). Multiscale neural modeling of resting-state fMRI reveals executive-limbic malfunction as a core mechanism in major depressive disorder. *Neuroimage Clin.* 31:102758. doi: 10.1016/j.nicl.2021.102758
- Li, G., Nair, S. S., and Quirk, G. J. (2009). A biologically realistic network model of acquisition and extinction of conditioned fear associations in lateral amygdala neurons. *J. Neurophysiol.* 101, 1629–1646. doi: 10.1152/jn.90765.2008
- Lopatina, O. L., Malinovskaya, N. A., Komleva, Y. K., Gorina, Y. V., Shuvaev, A. N., Olovyanikova, R. Y., et al. (2019). Excitation/inhibition imbalance and impaired neurogenesis in neurodevelopmental and neurodegenerative disorders. *Rev. Neurosci.* 30, 807–820. doi: 10.1515/revneuro-2019-0014
- Lumaca, M., Dietz, M. J., Hansen, N. C., Quiroga-Martinez, D. R., and Vuust, P. (2021). Perceptual learning of tone patterns changes the effective connectivity between Heschl's gyrus and planum temporale. *Hum. Brain Mapp.* 42, 941–952. doi: 10.1002/hbm.25269
- Malone, P. S., Eberhardt, S. P., Wimmer, K., Sprouse, C., Klein, R., Glomb, K., et al. (2019). Neural mechanisms of vibrotactile categorization. *Hum. Brain Mapp.* 40, 3078–3090.
- Markram, H. (2006). The Blue Brain Project. *Nat. Rev. Neurosci.* 7, 153–160.
- Markram, H., Muller, E., Ramaswamy, S., Reimann, M. W., Abdellah, M., Sanchez, C. A., et al. (2015). Reconstruction and simulation of neocortical microcircuitry. *Cell* 163, 456–492.
- Marreiros, A. C., Kiebel, S. J., and Friston, K. J. (2008). Dynamic causal modelling for fMRI: A two-state model. *Neuroimage* 39, 269–278.
- Mayberg, H. S. (1997). Limbic-cortical dysregulation: A proposed model of depression. *J. Neuropsychiatry Clin. Neurosci.* 9, 471–481.
- Mayberg, H. S. (2002). "Mapping mood: An evolving emphasis on frontal-limbic interactions," in *Principles of Frontal Lobe Function*, eds D. T. Stuss and R. T. Knight (Oxford: Oxford University Press), 376–391.
- Mayberg, H. S. (2003). Modulating dysfunctional limbic-cortical circuits in depression: Towards development of brain-based algorithms for diagnosis and optimised treatment. *Br. Med. Bull.* 65, 193–207. doi: 10.1093/bmb/65.1.193
- Meisel, C., Klaus, A., Kuehn, C., and Pleniz, D. (2015). Critical slowing down governs the transition to neuron spiking. *PLoS Comput. Biol.* 11:e1004097. doi: 10.1371/journal.pcbi.1004097
- Menon, V. (2011). Large-scale brain networks and psychopathology: A unifying triple network model. *Trends Cogn. Sci.* 15, 483–506. doi: 10.1016/j.tics.2011.08.003
- Moran, R. J., Symmonds, M., Stephan, K. E., Friston, K. J., and Dolan, R. J. (2011). An in vivo assay of synaptic function mediating human cognition. *Curr. Biol.* 21, 1320–1325. doi: 10.1016/j.cub.2011.06.053
- Mulders, P. C., van Eijndhoven, P. F., Scheme, A. H., Beckmann, C. F., and Tendolcar, I. (2015). Resting-state functional connectivity in major depressive disorder: A review. *Neurosci. Biobehav. Rev.* 56, 330–344.
- Olier, I., Trujillo-Barreto, N. J., and El-Deredey, W. (2013). A switching multiscale dynamical network model of EEG/MEG. *Neuroimage* 83, 262–287. doi: 10.1016/j.neuroimage.2013.04.046
- Park, H.-J., Eo, J., Pae, C., Son, J., Park, S. M., and Kang, J. (2021). State-dependent effective connectivity in resting-state fMRI. *Front. Neural Circuits* 15:719364. doi: 10.3389/fncir.2021.719364
- Power, J. D., Schlaggar, B. L., and Petersen, S. E. (2014). Studying brain organization via spontaneous fMRI signal. *Neuron* 84, 681–696.
- Prando, G., Zorzi, M., Bertoldo, A., Corbetta, M., Zorzi, M., and Chiuso, A. (2019). Sparse DCM for whole-brain effective connectivity from resting-state fMRI data. *Neuroimage* 208:116367. doi: 10.1016/j.neuroimage.2019.116367
- Razi, A., Seghier, M. L., Zhou, Y., McColgan, P., Zeidman, P., Park, H. J., et al. (2017). Large-scale DCMs for resting-state fMRI. *Netw. Neurosci.* 1, 222–241. doi: 10.1162/NETN\_a\_00015
- Riesenhuber, M., and Poggio, T. (2000). Models of object recognition. *Nat. Neurosci.* 3, 1199–1204.
- Rosa, M. J., Friston, K., and Penny, W. (2012). Post-hoc selection of dynamic causal models. *J. Neurosci. Methods* 208, 66–78.

- Sadeghi, S., Mier, D., Gerchen, M. F., Schmidt, S. N. L., and Hass, J. (2020). Dynamic Causal Modeling for fMRI with Wilson-Cowan-based neuronal equations. *Front. Neurosci.* 14:593867. doi: 10.3389/fnins.2020.593867
- Sanz-Leon, P., Knock, S. A., Spiegler, A., and Jirsa, V. K. (2015). Mathematical framework for large-scale brain network modeling in the virtual brain. *Neuroimage* 111, 385–430.
- Sanz-Leon, P., Knock, S. A., Woodman, M. M., Domide, L., Mersmann, J., McIntosh, A. R., et al. (2013). The Virtual Brain: A simulator of primate brain network dynamics. *Front. Neuroinf.* 7:10. doi: 10.3389/fninf.2013.00010
- Siegle, G. J., Steinhauer, S. R., Thase, M. E., Stenger, V. A., and Carter, C. S. (2002). Can't shake that feeling: Event-related fMRI assessment of sustained amygdala activity in response to emotional information in depressed individuals. *Biol. Psychiatry* 51, 693–707. doi: 10.1016/s0006-3223(02)01314-8
- Singh, M. F., Braver, T. S., Cole, M. W., and Ching, S. (2020). Estimation and validation of individualized dynamic brain models with resting state fMRI. *Neuroimage* 221:117046.
- Soares, J. M., Magalhães, R., Moreira, P. S., Sousa, A., Ganz, E., Sampaio, A., et al. (2016). A Hitchhiker's guide to functional Magnetic Resonance Imaging. *Front. Neurosci.* 10:515. doi: 10.3389/fnins.2016.00515
- Sokolov, A. A., Zeidman, P., Razi, A., Erb, M., Ryvlin, P., Pavlova, M. A., et al. (2020). Asymmetric high-order anatomical brain connectivity sculpts effective connectivity. *Netw. Neurosci.* 4, 871–890. doi: 10.1162/netn\_a\_00150
- Sporns, O. (2013). Network attributes for segregation and integration in the human brain. *Curr. Opin. Neurobiol.* 23, 162–171.
- Sporns, O., Tononi, G., and Kotter, R. (2005). The human connectome: A structural description of the human brain. *PLoS Comput. Biol.* 1:e42. doi: 10.1371/journal.pcbi.0010042
- Stephan, K. E., Iglesias, S., Heinzle, J., and Diaconescu, A. O. (2015). Translational perspectives for computational neuroimaging. *Neuron* 87, 716–732.
- Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J., and Friston, K. J. (2009). Bayesian model selection for group studies. *Neuroimage* 46, 1004–1017.
- Stimberg, M., Brette, R., and Goodman, D. F. (2019). Brian 2, an intuitive and efficient neural simulator. *elife* 8:e47314.
- Tagliazucchi, E., Crossley, N., Bullmore, E. T., and Laufs, H. (2016). Deep sleep divides the cortex into opposite modes of anatomical-functional coupling. *Brain Struct. Funct.* 221, 4221–4234. doi: 10.1007/s00429-015-1162-0
- Tognoli, E., and Kelso, J. A. (2014). The metastable brain. *Neuron* 81, 35–48.
- Tuckwell, H. C. (1988). *Introduction to Theoretical Neurobiology: Volume 1, Linear Cable Theory and Dendritic Structure*. Cambridge, UK: Cambridge University Press.
- Unal, G., Pare, J. F., Smith, Y., and Pare, D. (2013). Differential connectivity of short- vs. long-range extrinsic and intrinsic cortical inputs to perirhinal neurons. *J. Comp. Neurol.* 521, 2538–2550. doi: 10.1002/cne.23297
- Van Essen, D. C., and Glasser, M. F. (2016). The Human Connectome Project: Progress and Prospects. *Cerebrum* 2016, cer–10–16.
- Vidaurre, D., Abeysuriya, R., Becker, R., Quinn, A. J., Alfaro-Almagro, F., Smith, S. M., et al. (2017). Discovering dynamic brain networks from big data in rest and task. *Neuroimage* 180, 646–656. doi: 10.1016/j.neuroimage.2017.06.077
- Wang, X., Wu, Q., Egan, L., Gu, X., Liu, P., Gu, H., et al. (2019b). Anterior insular cortex plays a critical role in interoceptive attention. *elife* 8:e42265. doi: 10.7554/eLife.42265
- Wang, P., Kong, R., Kong, X., Liégeois, R., Orban, C., Deco, G., et al. (2019a). Inversion of a large-scale circuit model reveals a cortical hierarchy in the dynamic resting human brain. *Sci. Adv.* 5:eaat7854. doi: 10.1126/sciadv.aat7854
- Wang, R., Liu, M., Cheng, X., Wu, Y., Hildebrandt, A., and Zhou, C. (2021). Segregation, integration, and balance of large-scale resting brain networks configure different cognitive abilities. *Proc. Natl. Acad. Sci. U.S.A.* 118:e2022288118. doi: 10.1073/pnas.2022288118
- Wilson, H. R., and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* 12, 1–24.
- Zeidman, P., Jafarian, A., Corbin, N., Seghier, M. L., Razi, A., Price, C. J., et al. (2019a). A guide to group effective connectivity analysis, part 1: First level analysis with DCM for fMRI. *Neuroimage* 200, 174–190.
- Zeidman, P., Jafarian, A., Seghier, M. L., Litvak, V., Cagnan, H., Price, C. J., et al. (2019b). A guide to group effective connectivity analysis, part 2: Second level analysis with PEB. *Neuroimage* 200, 12–25. doi: 10.1016/j.neuroimage.2019.06.032