

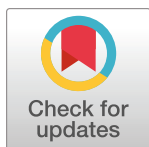
RESEARCH ARTICLE

# Multiscale effective connectivity analysis of brain activity using neural ordinary differential equations

Yin-Jui Chang<sup>1</sup>, Yuan-I Chen<sup>1</sup>, Hannah M. Stealey<sup>1</sup>, Yi Zhao<sup>1</sup>, Hung-Yun Lu<sup>1</sup>, Enrique Contreras-Hernandez<sup>1</sup>, Megan N. Baker<sup>1</sup>, Edward Castillo<sup>1</sup>, Hsin-Chih Yeh<sup>1,2</sup>, Samantha R. Santacruz<sup>1,3,4\*</sup>

**1** Biomedical Engineering, University of Texas at Austin, Austin, TX, United States of America, **2** Texas Materials Institute, University of Texas at Austin, Austin, TX, United States of America, **3** Electrical & Computer Engineering, University of Texas at Austin, Austin, TX, United States of America, **4** Institute for Neuroscience, University of Texas at Austin, Austin, TX, United States of America

\* [s.santacruz@austin.utexas.edu](mailto:s.santacruz@austin.utexas.edu)



## OPEN ACCESS

**Citation:** Chang Y-J, Chen Y-I, Stealey HM, Zhao Y, Lu H-Y, Contreras-Hernandez E, et al. (2024) Multiscale effective connectivity analysis of brain activity using neural ordinary differential equations. PLoS ONE 19(12): e0314268. <https://doi.org/10.1371/journal.pone.0314268>

**Editor:** Marko Čanadija, Faculty of Engineering, University of Rijeka, CROATIA

**Received:** April 17, 2024

**Accepted:** November 7, 2024

**Published:** December 4, 2024

**Peer Review History:** PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pone.0314268>

**Copyright:** © 2024 Chang et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All analyses were implemented using custom Python code. Code and data to replicate the main results is available at <https://github.com/santacruzlab/msDyNODE>.

## Abstract

Neural mechanisms and underlying directionality of signaling among brain regions depend on neural dynamics spanning multiple spatiotemporal scales of population activity. Despite recent advances in multimodal measurements of brain activity, there is no broadly accepted multiscale dynamical models for the collective activity represented in neural signals. Here we introduce a neurobiological-driven deep learning model, termed multiscale neural dynamics neural ordinary differential equation (msDyNODE), to describe multiscale brain communications governing cognition and behavior. We demonstrate that msDyNODE successfully captures multiscale activity using both simulations and electrophysiological experiments. The msDyNODE-derived causal interactions between recording locations and scales not only aligned well with the abstraction of the hierarchical neuroanatomy of the mammalian central nervous system but also exhibited behavioral dependences. This work offers a new approach for mechanistic multiscale studies of neural processes.

## Introduction

The brain is a complex system exhibiting computational structure involving multiple spatial scales (from molecules to whole brain) and temporal scales (from submilliseconds to the entire lifespan) [1]. Effective connectivity (EC) is a type of brain connectivity that characterizes relationships between brain regions [2]. Unlike structural connectivity for anatomical links and functional connectivity for statistical dependencies, EC refers to a pattern of causal interactions between distinct areas. Multiscale effective connectivity (msEC) among brain regions provides essential information about human cognition [3] and behaviors such as motor preparation [4], motor adaptation [5], motor timing [6], decision making [7], and working memory [8]. To date, much research has primarily focused on extracting EC from a single modality of neural measurements (e.g., electrophysiology, functional magnetic resonance imaging, and <sup>18</sup>F-

**Funding:** This work was supported by the National Science Foundation (Award No. 2145412, SRS), the Cockrell School of Engineering at the University of Texas at Austin (Start-up funds, SRS), the National Institutes of Health (Award No. DA060543, HCY), and the National Science Foundation (Award No. 2404334, HCY). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

**Competing interests:** The authors have declared that no competing interests exist.

fluorodeoxyglucose positron emission tomography [3]) and typically makes simplifying assumptions in which neural dynamics are linear [9] or log-linear [10]. However, the lack of the integration between multiple modalities and the reality of nonlinear neural dynamics prevents us from uncovering a deeper and more comprehensive understanding of system-level mechanisms of motor behavior [11, 12].

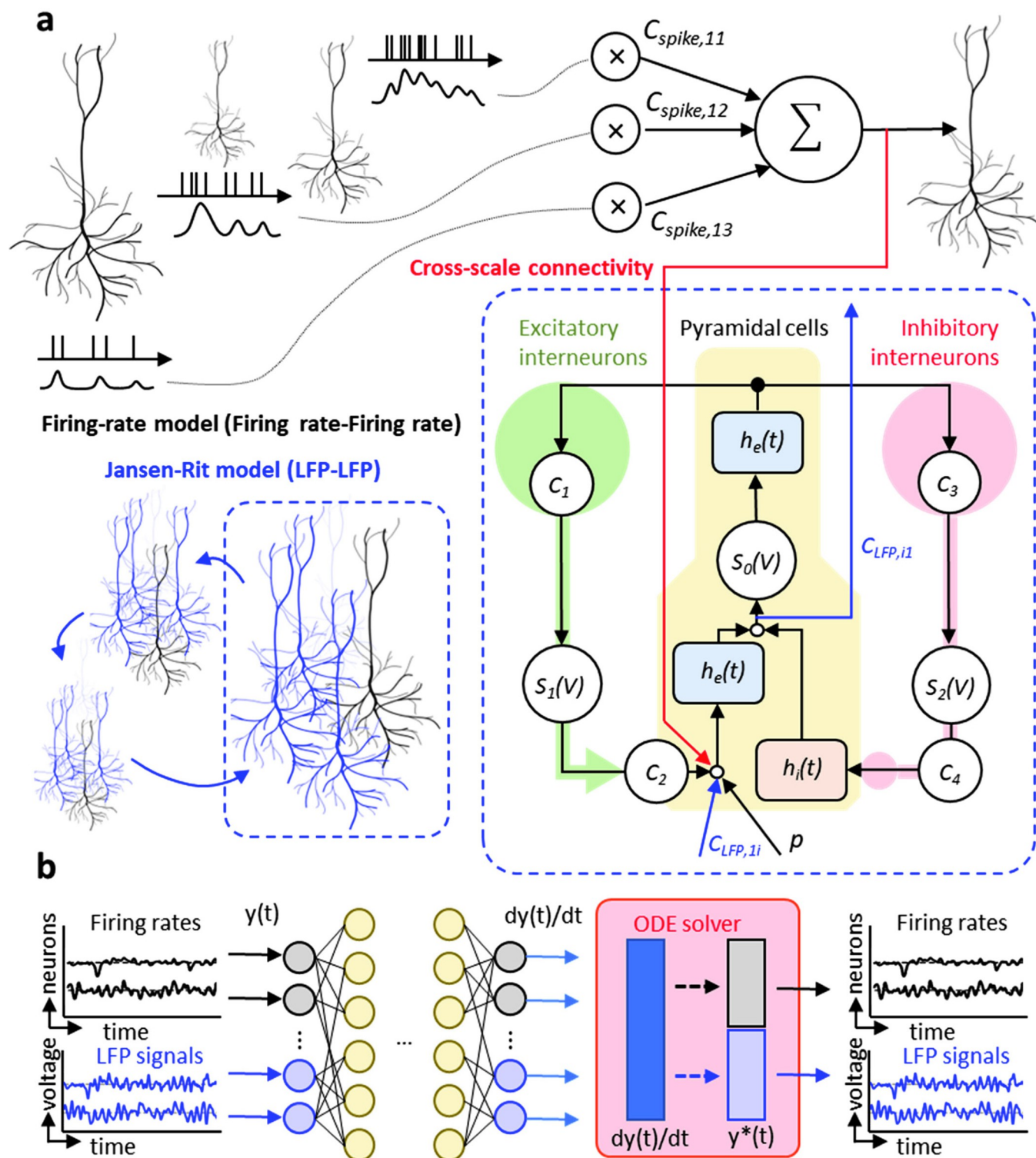
msEC can be divided into within-scale and cross-scale EC, where the former indicates the causal interactions between neural elements at the same spatial and temporal scales and the latter specifies the causal interactions between neural elements at different spatial or temporal scales. Previous work has largely focused on inferring within-scale EC *via* multivariate autoregressive models [13], vector autoregressive models [14], psycho-physiological interactions [15], structural equation modeling [16–19], or dynamic causal modeling [20]. Despite emergence of the cross-scale analyses such as source localization [21] and cross-level coupling (CLC) [22], the fidelity of experimental implementation of source localization is limited and only the statistical dependencies are quantified by CLC. To reveal the directed interactions across spatiotemporal scales of brain activity, recent work has developed the generalized linear model-based multi-scale method [23]. However, experimental data indicate that local brain dynamics rely on nonlinear phenomena [24]. Nonlinear models may be required to generate the rich temporal behavior matching that of the measured data [25]. Taking the nature of nonlinearity in brain computations, we have previously proposed the NBGNet, a sparsely-connected recurrent neural network (RNN) where the sparsity is based on the electrophysiological relationships between modalities, to capture cross-scale EC [26]. Despite the success of capturing complex dynamics using a nonlinear model, we still lack an integrative method that can infer nonlinear msEC.

To analyze multiscale neural activity in an integrative manner, we introduce a multiscale modeling framework termed msDyNODE (multiscale neural dynamics neural ordinary differential equation). Neural ordinary differential equation (NODE) is a new family of deep neural networks that naturally models the continuously-defined dynamics [27]. In our method, within-scale dynamics are determined based on neurobiological models at each scale, and cross-scale dynamics are added as the connections between latent states at disparate scales (Fig 1). Using both simulation and an experimental dataset, we demonstrate that msDyNODE not only reconstructs well the multi-scale data, even for the perturbation tasks, but also uncovers multi-scale causal interactions driving cognitive behavior.

## Results

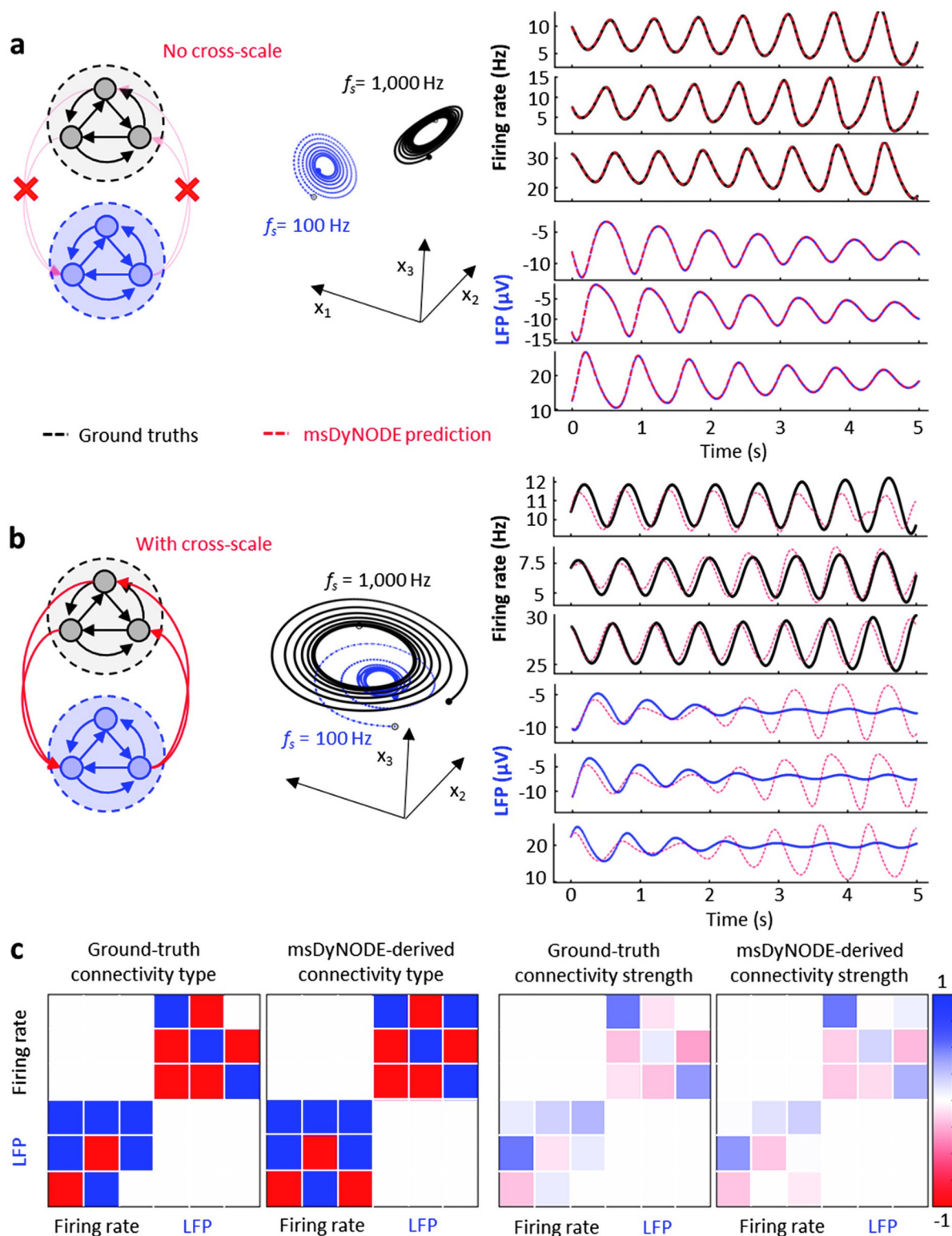
### Validation of msDyNODE framework using simulated Lorenz attractor

Since the Lorenz attractor model is a standard nonlinear dynamical system in the field with its simplicity and straightforward state space visualization [28, 29], we first demonstrate the msDyNODE framework using the simulated Lorenz attractor dataset. A Python program is employed to generate synthetic stochastic neuronal firing rates and local field potentials from deterministic nonlinear system. Two sets of Lorenz attractor systems are implemented to simulate activity at two scales: one to simulate firing rates at the single-neuron scale and another to stimulate local field potentials (LFPs) at the neuronal population scale. Without causal interactions between scales, the msDyNODE reconstructs well the Lorenz attractor parameters, simulated firing rates and LFPs (mean absolute error = 0.64 Hz for firing rate; = 0.18  $\mu$ V for LFPs; Fig 2A). To evaluate the performance of the msDyNODE in the multiscale system, we mimic cross-scale interactions by adding causal connections between latent states of the two systems (Fig 2B). Although the fitting accuracy is relatively poorer than the systems without causal interactions (mean absolute error = 1.43 Hz for firing rate; = 2.58  $\mu$ V for LFPs), the



**Fig 1. The architecture of msDyNODE applied to multiscale LFP and firing rate.** (a) Firing rate-Firing rate model follows the firing-rate model. LFP-LFP model follows the Jansen-Rit model. Cross-scale connectivity between firing rates and LFPs is added between latent variables of two systems. (b) The schematics of msDyNODE for multiscale firing rate-LFP model.

<https://doi.org/10.1371/journal.pone.0314268.g001>



**Fig 2. msDyNODE applied to Lorenz attractor.** (a) The evolution of the Lorenz system in its 3-dimensional state space for firing rates (black) and LFPs (blue; left). The synthetic firing rates (black) and LFPs (blue), as well as the msDyNODE predictions (red dashed line), were plotted as a function of time (right). (b) The same as **a** but with cross-scale causal interactions. (c) Ground-truth and identified cross-scale communication types (left) and causal interactions (right) between synthetic firing rates and LFPs.

<https://doi.org/10.1371/journal.pone.0314268.g002>



**Table 1. msDyNODE captures the Lorenz attractor parameters.** The predictions are summarized from 10 repeats of model training individually.

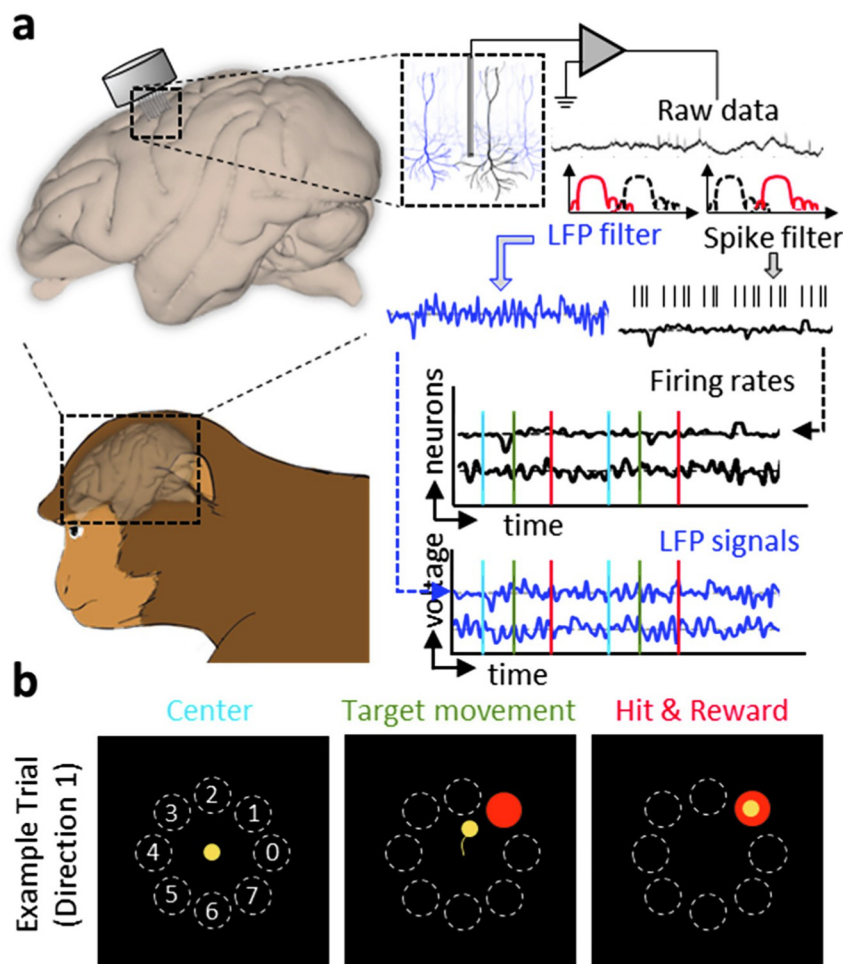
Model parameters	Ground truth	Predictions (n = 10)
$\sigma_1$	10	$10.09 \pm 0.04$
$\rho_1$	28	$28.02 \pm 0.06$
$\beta_1$	2.67	$2.69 \pm 0.03$
$\sigma_2$	8	$7.87 \pm 0.08$
$\rho_2$	20	$19.82 \pm 0.06$
$\beta_2$	3.33	$3.45 \pm 0.03$

<https://doi.org/10.1371/journal.pone.0314268.t001>

msDyNODE still captures the signals and the Lorenz attractor parameters (Table 1). Notably, with the cross-scale interactions between systems, the msDyNODE can reconstruct the ground truth accurately for 2.5 seconds. Furthermore, we assess if the msDyNODE can identify the types (excitatory or inhibitory) and the strength of causal interactions (Fig 2C). Positive and negative causal strengths correspond to excitatory and inhibitory effects, respectively. The positive causality identified by the msDyNODE is true positive when the ground truth is also positive. It became a false positive if the ground truth is negative. The identification accuracy is  $77 \pm 6\%$  (Fig 2C left). We also find that msDyNODE successfully captures the cross-scale causal interactions (mean absolute difference between the ground-truth and estimated causality = 0.07; Fig 2C right). These simulations verify that msDyNODE is a reliable framework for modeling multiscale systems.

### msDyNODE outputs reconstruct well experimentally-acquired firing rate and field potential signals

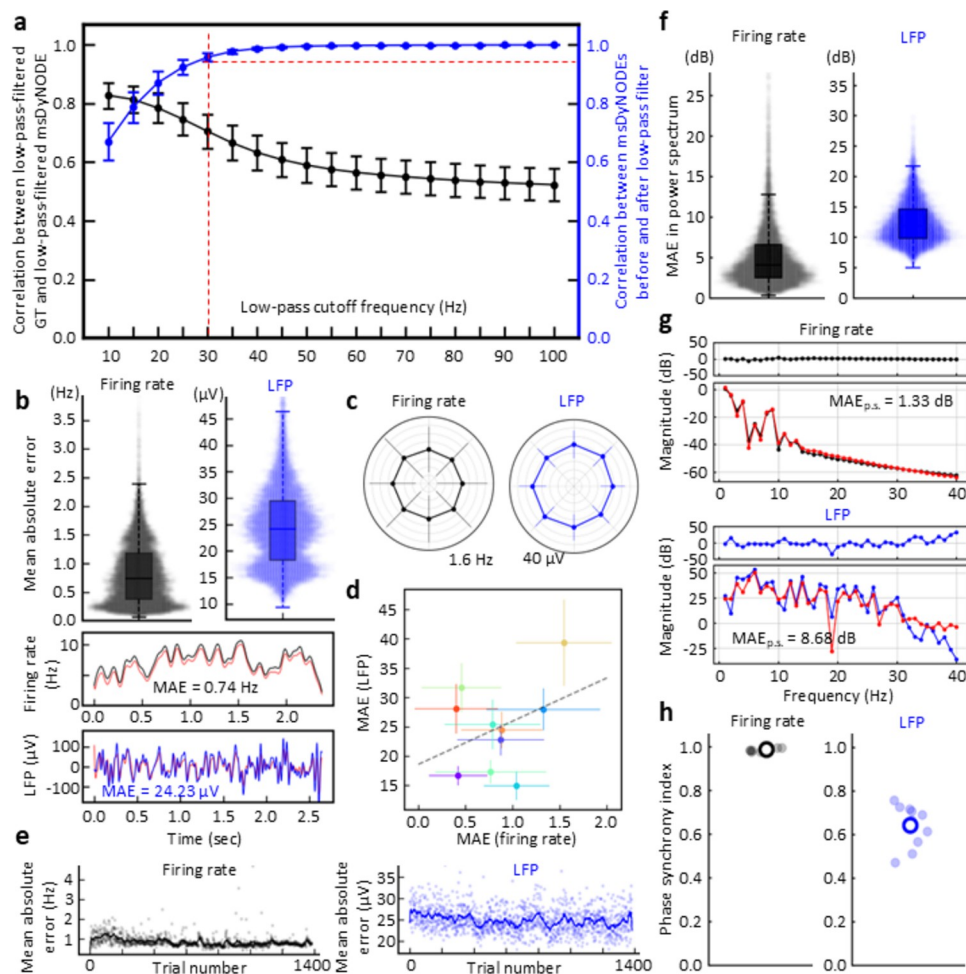
Firing rate and LFP activity are simultaneously recorded in the left dorsal premotor (PMd) and primary motor cortex (M1) of rhesus macaques (N = 2) while performing a center-out brain-machine interface (BMI) task [30–34] (Fig 3; see Materials and Methods). Multi-scale firing rate and LFP are acquired with the same set of electrodes but undergoing different pre-processing procedures (Fig 3A). During the center-out BMI task, the subjects volitionally modulate brain activity to move the cursor from the center to one of the eight peripheral targets. When BMI perturbation task is implemented, the subjects need to reassociate the existing neural patterns with the new direction [32, 35]. The increasing deviation shown in our simulation (Fig 2) is not the problem in our case with the average trial less than 2.5 seconds. The msDyNODE for the firing rate-LFP modeling is developed based on rate model [36–38] and Jansen-Rit model [39] (Fig 1; see Materials and Methods). By fitting the msDyNODE to the experimental datasets, we demonstrate the goodness-of-fit of the proposed multiscale framework in modeling multiscale brain activity using correlation and mean absolute error metrics (Fig 4). Correlation between ground truth data and the msDyNODE-predicted data defines a linear relationship between the real and predicted signal, with a strong correlation ( $> 0.7$ ) indicating consistent temporal co-variation between the two data up to a constant amplitude scaling. Mean absolute error (MAE), on the other hand, measures error in signal amplitude timepoint by timepoint but without describing the overall relationship between the data. Together, high correlation and low MAE indicate that the data co-vary together and any scaling difference between the real and predicted data is small. We find that indeed there is high correlation between ground truth data and msDyNODE-predictions, with msDyNODE primarily capturing the LFP activity below 30 Hz (Fig 4A). This observation is consistent with the fact that LFP neural dynamics are dominated by lower frequencies. Therefore, for the rest of



**Fig 3. Data acquisition and experimental task design for multiscale neural signals.** (a) Simultaneous recording of firing rates and LFP signals. (b) The visual feedback task contains eight different cursor movements, each corresponding to one of the eight outer targets. The color-coded tasks are also indicated in a.

<https://doi.org/10.1371/journal.pone.0314268.g003>

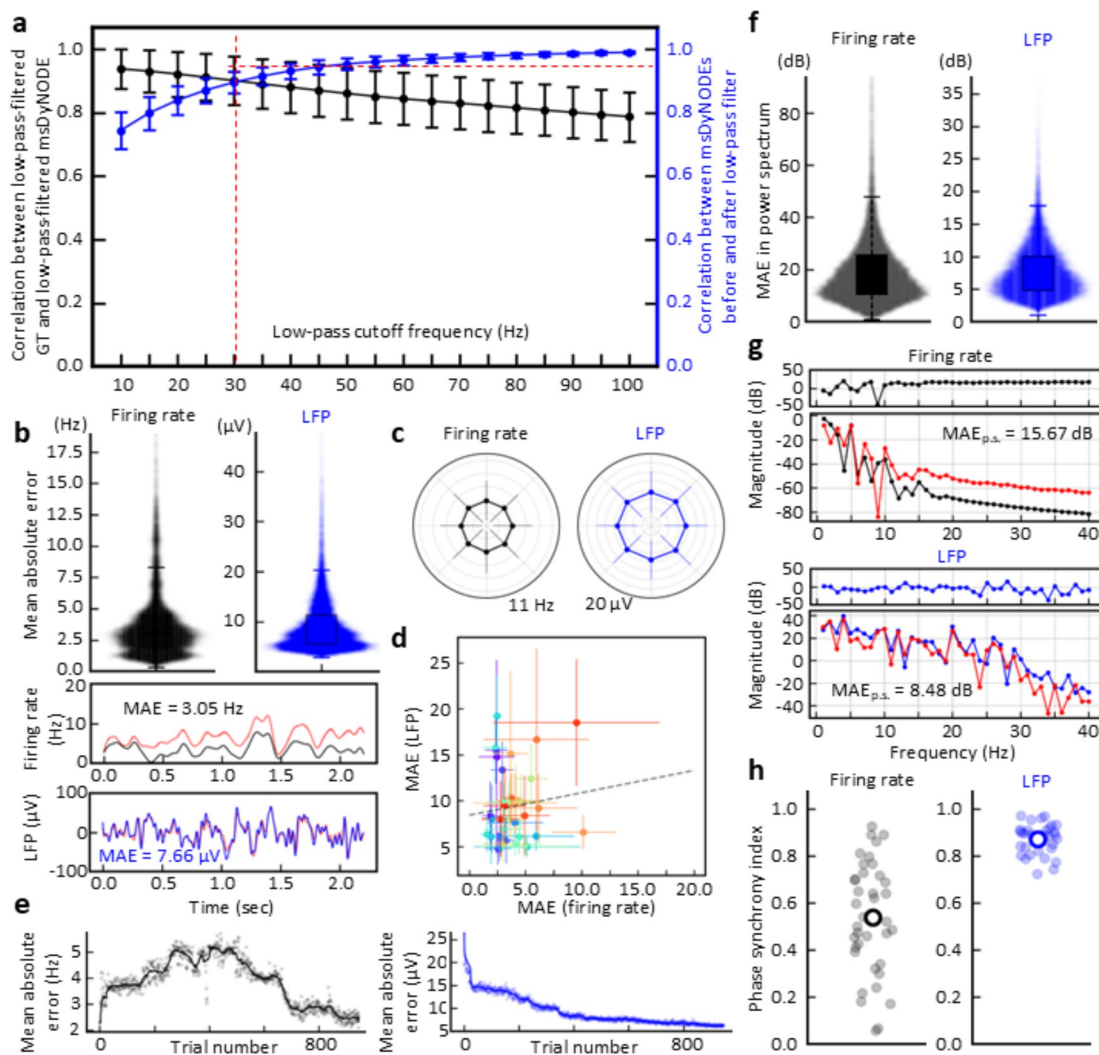
the evaluations, we focus on the performance in the frequency range of 0 and 30 Hz. Overall, the msDyNODE well reconstructed the firing rates (median of MAE = 0.74 Hz) and LFPs (median MAE = 24.23  $\mu$ V; Fig 4B). In addition, we find that the performance of the msDyNODE is target direction-independent, with a similar MAE over eight target directions for both firing rates and LFPs (Fig 4C). Interestingly, the reconstruction performances for firing rates and LFPs are not independent (Fig 4D). Good performance on certain channels indicated similarly good performance for different signal types, and vice versa. Surprisingly, the modeling performance for firing rates remains high over hundreds of trials even when a perturbation is introduced to increase the task difficulty (Fig 4E). However, the modeling performance for LFP gradually improves over trials, which may indicate that LFP dynamics become more predictable. Furthermore, the performance holds when applying the msDyNODE to a different monkey dataset (i.e., that it is not trained on), indicating that the msDyNODE is generalizable across different sessions and subjects (Fig 5). With a larger number of spiking units and LFPs recorded in this subject, it is expected that the msDyNODE can reconstruct LFP more accurately. The only difference in the reconstruction performance is that the firing rate predictions were worse during the first half of the experimental sessions, followed by



**Fig 4. msDyNODE captures and reconstructs the latent dynamics in the center-out BMI task for Monkey A.** (a) Correlation coefficient between ground truth (GT) signals and msDyNODE predictions (black) as a function of low-pass cutoff frequency (error bars, s.t.d.). In addition, we show correlation between msDyNODE before and after low-pass filter (blue). (b) Boxplots and swarmplots of the mean absolute errors in firing rates and LFPs (top). The representative GT and msDyNODE with the MAE equaling to the median values of all the MAEs (bottom). (c) Error bars of the MAE over eight different target directions presented in polar coordinates (error bars, s.t.d.). (d) Scatter plots of the MAE over recording channels (error bars, s.t.d.). (e) MAE values of firing rates and LFPs over trials. Dim points represent average MAE ( $n = 10$ ) at each trial. (f) Boxplots and swarmplots of the mean absolute errors in power spectrum for firing rates and LFPs. (g) The representative power spectrum from GT and msDyNODE with the selected example in Fig 4B. (h) Scatter plots of PSI values for firing rates and LFPs. Empty circles indicate overall average PSI values. Dim points represent average PSI over trials for each recording channel.

<https://doi.org/10.1371/journal.pone.0314268.g004>

increasing accuracy for the second half of the recording sessions (Fig 5E). This may indicate the neural dynamics were less stable during the first half of the sessions and thus more challenging to be captured. Beyond MAE in the time domain, we also assess MAE in the frequency domain and phase synchronization in the phase domain (Figs 4F–4H, 5F–5H; see Materials and Methods). Overall, the msDyNODE captures the signal's power for both Monkey A (Fig 4F and 4G) and Monkey B (Fig 5F and 5G). Notably, phase synchronization is recognized as a fundamental neural mechanism that supports neural communication and plasticity [40]. Therefore, the model performance in the phase domain is crucial. We demonstrated that msDyNODE-predictions are in sync with ground truth by showing most of the predictions have a phase synchrony index greater than 0.5 (Figs 4H and 5H). These experimental results



**Fig 5. msDyNODE captures and reconstructs the latent dynamics in the center-out BMI task for Monkey B.** (a) Correlation coefficient between ground truths (GT) and msDyNODE (black) and between msDyNODE before and after low-pass filter (blue) as a function of low-pass cutoff frequency (error bars, s.d.). (b) Boxplots and swarmplots of the mean absolute errors in firing rates and LFPs (top). The representative GT and msDyNODE with the MAE equaling to the median values of all the MAEs (bottom). (c) Error bars of the MAE over eight different target directions presented in polar coordinates (error bars, s.d.). (d) Scatter plots of the MAE over recording channels (error bars, s.d.). (e) MAE values of firing rates and LFPs over trials. Dim points represent average MAE ( $n = 38$ ) at each trial. (f) Boxplots and swarmplots of the mean absolute errors in power spectrum for firing rates and LFPs. (g) The representative power spectrum from GT and msDyNODE with the selected example in Fig 5B. (h) Scatter plots of PSI values for firing rates and LFPs. Empty circles indicate overall average PSI values. Dim points represent average PSI over trials for each recording channel.

<https://doi.org/10.1371/journal.pone.0314268.g005>

validated that msDyNODE can capture the dynamics hidden in the multiscale brain systems, and msDyNODE can be generalized to different sessions and different subjects.

### msDyNODE decodes underlying behavior *via* multiscale effective connectivity

In msDyNODE, the msEC can be derived from the parameters that indicate the causal influence that the latent states of a neural system exert over those of another system. The average connectivity for each target direction is calculated by subtracting the grand-averaged

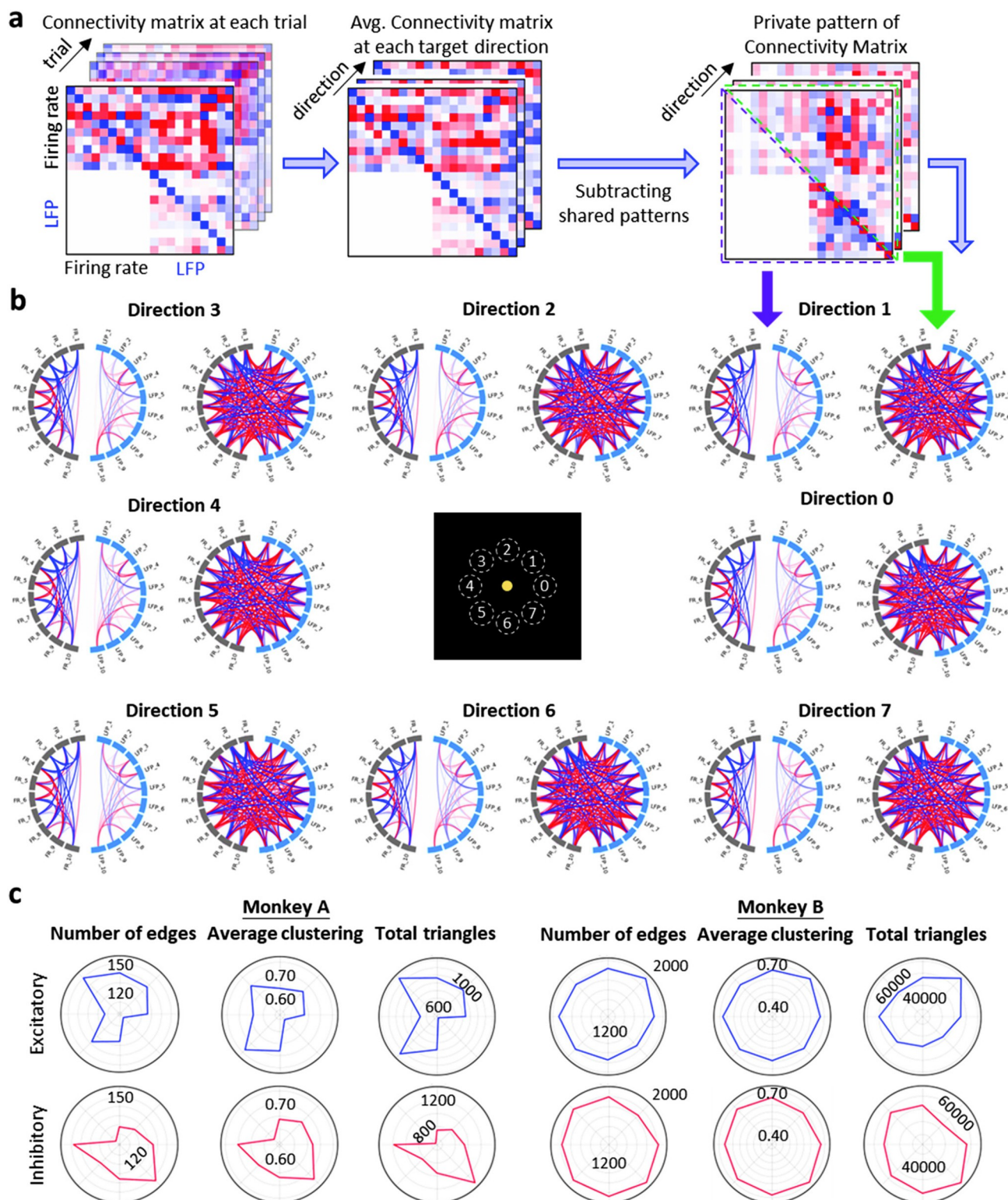


connectivity from the average connectivity within each target (Fig 6A). For each direction, the bi-directional msEC is divided into two parts (upper and lower triangular connectivity matrix) and visualized respectively (Fig 6B). Most of the msEC remained similar across target directions, indicating the common patterns of voluntary movement. To investigate if there existed unique patterns of excitatory and inhibitory subnetworks across directions, we quantified the individual subnetworks using common graph properties such as number of edges, average clustering, and total triangles (Fig 6C). Interestingly, these graph properties are different across the eight target directions, revealing the excitatory and inhibitory neural dynamics exhibited unique connectivity patterns relating to target direction. Thus, msDyNODE is demonstrated to be capable of capturing the multiscale effective connectivity patterns underlying behaviors.

## Discussion

Large populations of individual neurons coordinate their activity together to achieve a specific cognitive task, highlighting the importance of studying the coordination of neural activity. Over the past decades, we have learned much about the human cognitive behaviors and viewed an explosive growth in the understanding of single neurons and synapses [41, 42]. However, we still lack a fundamental understanding of multiscale interactions. For decades a critical barrier to multiscale study was the recording technologies available, forcing scientists to choose either the microscale or macroscale, with few researchers addressing on the interactions between scales. Neurophysiologists, for example, often focused on single-neuronal activity to investigate the sensory consequences of motor commands with the bottom-up approach [43], without the consideration of brain rhythm. Instead, cognitive neuroscientists pay attention to the neural oscillations at a larger scale (e.g., electroencephalography) with the top-down approach to establish the links between brain rhythm and cognitive behaviors [44], disregarding the spiking activity of single neurons. With the advancement of multi-modal measurements, there is an unmet need for an integrative framework to analyze multiscale systems. In the present study, we propose msDyNODE to model the multiscale signals of firing rates and field potentials, and then infer multiscale causal interactions that exhibit distinct patterns for different motor behaviors.

To the best of our knowledge, this is the first demonstration of a NODE applied to model multiscale neural activity. Assuming brain computation as a nonlinear operator [45–51], we employ a deep learning technique to approximate the nonlinear mapping of the state variables in dynamic systems. Different deep learning architectures are tailored for specific tasks. Common examples include convolution neural networks for image recognition [52], recurrent neural networks (RNN) for sequential data [53], transformers for natural language processing tasks [54], and generative adversarial networks for generating authentic new data [55] and denoising [56]. While RNNs are a powerful approach to solve the dynamic equations [57, 58], it may fail to capture faster dynamics information or introduce artifacts by matching the sampling rates between signals. In contrast to the RNN which describes the complicated state transformation at discretized steps for time-series inference, the proposed msDyNODE models continuous dynamics by learning the derivative of the state variables [27], indicating that both slow and fast dynamics can be captured. Such a capability is crucial for multiscale modeling since the system dynamics vary at different scales. Additionally, NODE allows us to define the multiscale system by customizing the differential equations in the network, in which we can investigate the physiological interpretation of the modeled systems. It is worth noting that the nonconstant sampling can be addressed by preprocessing the NODE output with the observation mask [59]. Therefore, unmatched sampling rates between modalities can be resolved by feeding individual observation masks, respectively. Furthermore, in the real world,



**Fig 6. msDyNODE captures msEC patterns underlying behaviors.** (a) Workflow to obtain the private pattern of connectivity matrix for each target direction from msDyNODE-inferred msEC. (b) Circular connectivity graphs of lower (left) and upper (right) triangular msEC matrix for each target direction. (c) Graph properties (number of edges, average clustering, number of total triangles) over eight different target directions presented in polar coordinates for Monkey A and B, and excitatory and inhibitory subnetworks, respectively.

<https://doi.org/10.1371/journal.pone.0314268.g006>

not all the signals can be measured at fixed time intervals. The missing data issue can thus introduce artefacts using a conventional approach which assumes the signals are sampled regularly. While there exists several methods, such as dropping variables, last observation carried forward and next observation carried backward, linear interpolation, linear regression, or imputation dealing with missing data [60], none of them serves as good way to deal with this issue because they add no new information but only increase the sample size and lead to an underestimate of the errors. The proposed framework also holds great potential to be an alternative approach dealing with missing data commonly seen in the real world.

Comparing with existing biophysical models of brain functioning, including NetPyNE [61], modified spectral graph theory model (M-SGM [62]), and SGM integrated with simulation-based inference for Bayesian inference (SBI-SGM [63]), we demonstrate that msDyNODE is superior these approaches. msDyNODE showed smaller MAEs in both time and frequency domains, and greater phase synchronization with the ground truth signals (S1 Fig). The potential reason for relatively poor performance in NetPyNE may be due to inaccurate modeling. Indeed, NetPyNE is a powerful tool to define the model at molecular, cellular, and circuit scales when the model parameters such as populations, cell properties, and connectivity are accurate. Although NetPyNE also provides evolutionary algorithms beyond the grid parameter search to perform parameter optimization and exploration, the improper selection of parameters and their ranges to be optimized can degrade the performance. Furthermore, msDyNODE exhibits better performance than both versions of SGMs (S2 Fig). The rationale for why msDyNODE models the real multiscale brain signals better than SGMs may be due to the consideration of nonlinear brain dynamics and spatially varying parameters. Another advantage of msDyNODE over NetPyNE, M-SGM, and SBI-SGM is the adaptability of the new model. For msDyNODE, the user can easily modify the differential equation sections in the script. However, NetPyNE requires the development of an external module in NEURON [64, 65]. For M-SGM and SBI-SGM, the new transfer function is required to derive from the new or customized models.

While msDyNODE provides accurate analysis for multiscale systems, its cost lies in the optimal selections of neural models. At the scale of firing rate, integrate-and-fire model and its variants (leaky integrate-and-fire [66, 67] and quadratic integrate-and-fire [68, 69]) are all plausible options. At the scale of field potential activity, the candidate model includes Jansen-Rit model [39] that characterizes three populations (pyramidal cells, excitatory interneurons, and inhibitory interneurons) and Wilson-Cowan model [70] that refers to two coupled populations (excitatory and inhibitory). Suboptimal selections of neural models may result in misleading conclusions. To avoid suboptimal model selection, probabilistic statistical measures such as Akaike Information Criterion [71, 72], Bayesian Information Criterion [73], and minimum description length [74, 75] can be implemented to ensure the correct selection of the neural models. Furthermore, hours of network training time are another issue for quick implementation. In future work, transfer learning [76] from the previously trained network may be a possible strategy to improve computation time by potentially speeding up the convergence of the learning process.

Recent evidence suggests that signal changes on multiple timescales at multiple levels in the motor system allow arbitration between exploration and exploitation to achieve a goal [77–80]. Still, the role of cross-scale, as well as within-scale, causal interactions in motor learning remains incompletely understood [78, 79, 81]. In this work, we utilize the msDyNODE to study the essential brain function that modulates the motor commands to achieve desired actions, showing distinct dynamic patterns underlying different behaviors. Although existing estimators of causal brain connectivity (e.g., Granger causality [82] and directed transfer function (DTF [83])) provide disparate graph properties (S3 and S4 Figs), Granger causality

supports our observation that both excitatory and inhibitory msEC exhibit unique patterns relating to target directions. In contrast, DTF failed to demonstrate unique patterns, which may be due to its inability to be divided into excitatory and inhibitory subnetworks. While both existing estimators are powerful tools for characterizing functional coupling between brain regions, they primarily reflect the patterns of statistical dependence. To better reveal the causal interactions that align with the actual mechanisms of brain function, it is suggested to assess effective connectivity using a mechanistic model, such as msDyNODE. Taken together, our work represents an important step forward towards multiscale modeling of brain networks for mechanistic understanding of neural signaling. The underlying multiscale dynamics embedded in msDyNODE illustrate how the individual neurons and populations of neurons communicate across scales, which is a key factor in uncovering the mechanisms of brain computations and the mediation of the behaviors.

## Materials and methods

### Ethics statement

All the experiments were performed in compliance with the regulation of the Institutional Animal Care and Use Committee at the University of Texas at Austin.

### Experimental protocol

Two male rhesus macaques are used in the behavioral and electrophysiological experiments. Before the experimental session, we run a calibration session. During calibration, the subject passively observes a cursor moving from the center target toward a randomly generated peripheral target (in one of eight possible positions), followed by the cursor movement back to the center. In addition to providing continuous visual feedback, we also reinforce the behavior and neural activity by delivering a small juice reward directly into the subject's mouth. The neural data is recorded for approximately three and a half minutes (or reaching ~6 trials per target direction). A Kalman filter (KF) is employed as the decoder to map the spike count from each unit to a two-dimensional cursor control output signal [84, 85]. While the KF decodes both the intended position and velocity, only the velocity is used to estimate the position at the next time point based on kinematic equations of motion. To increase the initial performance and reduce directional bias, we conduct daily, 10-minute closed-loop decoder adaptation (CLDA) [85–89] sessions. Both the decoder and neural activity adapt to complete center-out tasks with consistent trial times and straight path lengths to each target. After the calibration session, the main task is manually initiated. The subject then completes a BMI task called “center-out” [90–92]. During the task, spiking activity is recorded online to produce cursor control commands in real-time. Spikes for each unit are summed over a window of 100 millisecond and serve as the input to the decoder. The neural activity is then transformed into a “neural command” by applying the dot product of the spike count vector to the Kalman gain matrix. Cursor position is updated iteratively by adding the cursor position to the product of velocity, which is determined by the neural command, times update time (100 ms). In each trial, the subjects control the velocity of a computer cursor to move from the center target toward one of eight outer targets. Only one peripheral target is presented on a given trial. The order of the appearance of the target is pseudorandomly selected; for every eight consecutive trials, each target is shown once in a random order. The 8 targets were radially distributed from 0° to 360° (0°, 45°, 90°, 135°, 180°, 225°, 270°, 315°) at equal distances from the center (10 centimeters). Upon successful completion of moving and holding the cursor at the peripheral target for 0.2 seconds, the target turns green (cue for success), and a small juice reward is dispensed directly into the subject's mouth. The cursor then automatically appears at the



center of the screen to initiate another new trial. Subjects can fail the task in two ways: (1) failure in holding the cursor at the center target or the peripheral target for 0.2 seconds or (2) failure in reaching the peripheral target within specified time (10 seconds). The subject has 10 chances to complete a successful trial before the task automatically moves onto the next target. During the BMI tasks, we also implement perturbation task by perturbing the decoder using a visuomotor rotation in which the cursor movements are rotated by an angle. The subjects then need to reassociate the existing neural patterns with new directions [32, 35].

### Spike trains and LFP data

The extracellular single and multi-unit activity in the left primary motor cortex (M1) and dorsal premotor cortex (PMd) are recorded using a 64- or 128-channel chronic array (Innovative Neurophysiology, Inc., Durham, NC; Fig 3A) from both subjects. The spike trains are acquired at 30 kHz sampling frequency, and the LFPs are acquired at 1 kHz sampling frequency. After excluding the recording channels that fail to capture activity (average firing rate < 1 Hz), 10 (Monkey A) and 38 (Monkey B) channels are considered for analysis. Cursor movements are tracked using the custom-built Python-based software suite. Neuronal signals are recorded using Trellis (Ripple Neuro, UT, USA) interfacing with Python (v3.6.5) via the Xipppy library (v1.2.1), amplified, digitized, and filtered with the Ripple Grapevine System (Ripple Neuro, UT, USA).

### Multiscale dynamics modeling with neurobiological constraints

We define a multi-scale dynamics network as a collection of neural recordings from different modalities (e.g., spike trains, LFPs, EEGs, fast-scan cyclic voltammetry, calcium imaging, functional magnetic resonance imaging, and functional near-infrared spectroscopy). A generic multi-scale dynamics system, where the evolution of latent variables and the output was described by the nonlinear functions of latent states and corresponding inputs, for  $M$  modalities is as follows,

$$\dot{\mathbf{x}}_i = \sum_{j=1}^M f_{ij}(\mathbf{x}_j, \theta_j, t), \quad i \in [1, 2, \dots, M]$$

$$\mathbf{y}_i = g_{ii}(\mathbf{x}_i, \theta_i, t), \quad i \in [1, 2, \dots, M]$$

where  $\mathbf{x}_i$ ,  $\mathbf{y}_i$  represent the latent state variables and the observations for  $i^{\text{th}}$  modality, respectively,  $f_{ij}$  denotes within-scale ( $i = j$ ) and cross-scale ( $i \neq j$ ) dynamics parameterized by  $\theta_j$ , and  $g_{ii}$  is the observation model in each modality. In this work, we focus on firing rates and LFPs (Figs 1 and 3), referred to as multi-scale signals. In addition, to enable the interpretability of the deep learning model, we introduce neurobiological constraints in our proposed network. Constraints including integration of modeling across different scales, the nature of the neuron model, regulation and control through interplay between excitatory and inhibitory neurons, and both local within- and global between-area connectivity have been reported to make neural network models more biologically plausible [93]. How are these neurobiological constraints implemented in the proposed approach are described in the following sections.

The multi-scale dynamics modeling for firing rate activity and LFP are based on well-established neurobiological models can be divided into three parts: (1) firing rate-firing rate within-scale model, (2) LFP-LFP within-scale model, and (3) firing rate-LFP cross-scale model. The rate model is employed as the firing rate-firing rate inference model with  $N_{\text{tot}}$  coupled neurons

[36–38]:

$$\frac{dx_{FR,i}}{dt} = \frac{x_{FR,i} + \text{sigm}\left(\sum_{j=1}^{N_{tol}} \left(C_{hidden\ FR,ij} \frac{dx_{FR,j}}{dt}(t) + C_{FR,ij} x_{FR,j}(t)\right)\right)}{-\tau_m},$$

where  $x_{FR,i}$  represents the membrane voltage of neuron  $i$ ,  $\tau_m$  denotes the membrane time constant,  $C_{FR,ij}$  and  $C_{hidden\ FR,ij}$  represents two types of causal interactions between presynaptic neuron  $j$  and postsynaptic neuron  $i$ . For the LFP-LFP within-scale model, we implement the Jansen-Rit model to describe the local cortical circuit by second-order ODEs [39]:

$$\ddot{x}_{LFP,i0} = Aa \text{sigm}_0(x_{LFP,i1}(t) - x_{LFP,i2}(t)) - 2a\dot{x}_{LFP,i0} - a^2 x_{LFP,i0}(t),$$

$$\ddot{x}_{LFP,i1} = Aa[p_i(t) + C_2 \text{sigm}_1(C_1 x_{LFP,i0}(t)) + p_{mu}] - 2a\dot{x}_{LFP,i1} - a^2 x_{LFP,i1}(t),$$

$$\ddot{x}_{LFP,i2} = BbC_4 \text{sigm}_2(C_3 x_{LFP,i0}(t)) - 2b\dot{x}_{LFP,i2} - b^2 x_{LFP,i2}(t),$$

$$p_i(t) = \sum_j^{N_{tol}} C_{LFP,ij} \text{sigm}_0(x_{LFP,j1}(t) - x_{LFP,j2}(t)),$$

where  $\text{sigm}()$  is a sigmoid function,  $A$  and  $B$  represent the maximum amplitude of the excitatory and inhibitory postsynaptic potentials (PSPs),  $a$  and  $b$  denote the reciprocal of the time constants of excitatory and inhibitory PSPs,  $p_{mu}(t)$  represents the excitatory input noise of the neuron  $i$ , and  $p(t)$  represents the excitatory input of the neuron  $i$  from other neurons.

For the cross-scale model that identifies and quantifies cross-scale communications, we consider the causal interactions between the hidden states (membrane voltage of single neuron for spike; membrane potential of pyramidal, inhibitory, and excitatory neurons) as the effective connectivity:

$$h_{LFP} = Ch_{FR} + \varepsilon,$$

where  $C$  represents the cross-scale causal interactions, and  $\varepsilon$  denotes the error, which includes inputs from other units which are not explicitly considered. Note here that the cross-scale interactions are defined to be unidirectional and linear due to fact that the LFP are defined as the summed and synchronous electrical activity of the individual neurons. After implementing the cross-scale causal interactions as the excitatory input of the neurons, the second ordinary differential equation in the Jansen-Rit model becomes as follows,

$$\ddot{x}_{LFP,i1} = Aa[p_i(t) + C_2 \text{sigm}_1(C_1 x_{LFP,i0}(t)) + p_{mu} + C_{FR-LFP,ij} x_{FR,j}] - 2a\dot{x}_{LFP,i1} - a^2 x_{LFP,i1}(t),$$

Taken together, combining the above equations, our multiscale dynamics model for spike and field potential can be written as follows, where  $F_{FR-FR}$  and  $F_{LFP-LFP}$  represent the within-scale dynamics equations while  $F_{FR-LFP}$  denotes the cross-scale dynamics equations:

$$\frac{d\mathbf{x}}{dt} = \begin{bmatrix} \frac{dx_{FR}}{dt} \\ \frac{dx_{LFP}}{dt} \end{bmatrix} = \begin{bmatrix} F_{FR-FR} & \mathbf{0} \\ \mathbf{0} & F_{LFP-LFP} \end{bmatrix} \begin{bmatrix} \mathbf{x}_{FR} \\ \mathbf{x}_{LFP} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ F_{FR-LFP} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \frac{dx_{FR}}{dt} \\ \frac{dx_{LFP}}{dt} \end{bmatrix} + \begin{bmatrix} \mathbf{b}_{FR} \\ \mathbf{0} \end{bmatrix}.$$

**Multiscale neural dynamics neural ordinary differential equation (msDyNODE).** Popular models such as recurrent neural networks and residual networks learn a complicated

transformation by applying a sequence of transformations to the hidden states [27]:  $\mathbf{h}_{t+1} = \mathbf{h}_t + f(\mathbf{h}_t, \theta_t)$ . Such iterative updates can be regarded as the discretization of a continuous transformation. In the case of infinitesimal update steps, the continuous dynamics of the hidden states can be parameterized with an ordinary differential equation (ODE):

$$\frac{d\mathbf{h}(t)}{dt} = f(\mathbf{h}(t), \theta, t).$$

A new family of deep neural networks, termed the neural ODE (NODE), was thus introduced to parameterize the  $f$  using a neural network [27]. The output of the NODE was then computed using any differential equation solver (e.g., Euler, Runge-Kutta methods). In this work, we utilize Runge-Kutta method with a fixed time step of 1 ms. The resulting msDy-NODE model consists of 7 layers with 1,480 and 18,392 trainable parameters for Monkey A and B, respectively. NODE exhibits several benefits, including memory efficiency, adaptive computation, and the capability of incorporating data arriving at arbitrary times. Recent work proposed a NODE-based approach with a Bayesian update network to model the *sporadically* observed (i.e., irregular sampling) multi-dimensional time series dataset [59]. Therefore, NODE serves as powerful tool for multi-scale data analysis.

**Synthetic Lorenz attractor.** The Lorenz attractor is a simple but standard model of a non-linear, chaotic dynamical system in the field [28, 94]. It consists of nonlinear equations for three dynamic variables. The state evolutions are derived as follows,

$$\dot{x}_1 = \sigma(x_2 - x_1)$$

$$\dot{x}_2 = x_1(\rho - x_3) - x_2$$

$$\dot{x}_3 = x_1x_2 - \beta x_3$$

The standard parameters are  $\sigma = 10$ ,  $\rho = 28$ , and  $\beta = 8/3$ . The Euler integration is used with  $\Delta t = 0.001$  (i.e. 1 ms). We first simulate two sets of Lorenz attractor systems with different parameter sets ( $\sigma_1 = 10$ ,  $\rho_1 = 28$ ,  $\beta_1 = 8/3$ ,  $\sigma_2 = 8$ ,  $\rho_2 = 20$ , and  $\beta_2 = 10/3$ ) but without cross-scale interactions:

$$\dot{x}_1 = \sigma_1(x_2 - x_1),$$

$$\dot{x}_2 = x_1(\rho_1 - x_3) - x_2,$$

$$\dot{x}_3 = x_1x_2 - \beta_1x_3,$$

$$\dot{x}_4 = \sigma_2(x_5 - x_4),$$

$$\dot{x}_5 = x_4(\rho_2 - x_6) - x_5,$$

$$\dot{x}_6 = x_4x_5 - \beta_2x_6,$$

with one system for a population of neurons with firing rates given by the Lorenz variables and another system for LFPs given by the Lorenz variables (Fig 2). We start the Lorenz system with a random initial state vector and run it for 6 seconds. We hypothesize that the neural activity consists of multiple marginally stable modes [95, 96]. The last five seconds were selected to ensure marginal stability in the simulation. Three different firing rates and LFPs were then generated with different sampling rates (1,000 Hz for spikes and 100 Hz for LFPs).

Models are trained with ten batches of 1-second data with randomly selected starting points for 1,000 iterations.

To evaluate the fitting performance of the msDyNODE with the Lorenz systems with cross-scale interactions, we then simulate two sets of Lorenz attractor systems with different parameter sets ( $\sigma_1 = 8, \rho_1 = 28, \beta_1 = 8/3, \sigma_2 = 10, \rho_2 = 20$ , and  $\beta_2 = 10/3$ ) and cross-scale interactions:

$$\dot{x}_1 = \sigma_1(x_2 - x_1) + 0.1x_4 + 0.2x_5 + 0.3x_6,$$

$$\dot{x}_2 = x_1(\rho_1 - x_3) + 0.5x_4 - 0.1x_5 + 0.1x_6,$$

$$\dot{x}_3 = x_1x_2 - \beta_1x_3 - 0.2x_4 + 0.1x_5,$$

$$\dot{x}_4 = \sigma_2(x_5 - x_4) + 0.5x_1 - 0.1x_2,$$

$$\dot{x}_5 = x_4(\rho_2 - x_6) - x_5 - 0.2x_1 + 0.1x_2 - 0.3x_3,$$

$$\dot{x}_6 = x_4x_5 - \beta_2x_6 - 0.1x_1 - 0.2x_2 + 0.4x_3,$$

All the other simulation settings remain the same as above.

**Phase synchrony assessment.** We apply the Hilbert transform,  $\mathbf{HT}[\cdot]$ , on a pair of signals,  $\mathbf{s}_1(t)$  and  $\mathbf{s}_2(t)$ , in order to obtain the analytical signals,  $\mathbf{z}_1(t)$  and  $\mathbf{z}_2(t)$ .

$$z_i(t) = s_i(t) + j\mathbf{HT}[s_i(t)] = \mathbf{A}_i(t)e^{j\Phi_i(t)}$$

$$\mathbf{HT}[s_i(t_k)] = s_i(t_k) * \frac{1}{2\pi} \left[ \int_{-\pi}^0 j \cdot e^{jwk} dw - \int_0^{\pi} j \cdot e^{jwk} dw \right]$$

where  $k = 1$  to  $T$ ,  $\mathbf{A}_i(t)$  represents the instantaneous amplitude, and  $\Phi_i(t)$  represents the instantaneous phase. The instantaneous phase synchronous (IPS [97]), which measured the phase similarity at each timepoint, can be calculated by the following,

$$IPS(t) = 1 - \sin\left(\frac{|\Phi_1(t) - \Phi_2(t)|}{2}\right)$$

where the phase is in the unit of degree. IPS spans the range of 0–1, where a larger value indicates a stronger synchrony. We then define a quarter of the whole range of phase difference ( $180^\circ$ ),  $45^\circ$ , as the threshold. When the phase difference is less than  $45^\circ$ , IPS was greater than 0.62, thus revealing a better performance. We finally calculated the PSI by the ratio of the time with the IPS greater than 0.62,

$$PSI = \frac{t_{IPS > 0.62}}{T}$$

## Supporting information

**S1 Fig. Benchmark with NetPyNE.** Scatter plots of MAE in the time domain, MAE in the frequency domain and PSI in the phase domain. Empty circles indicate overall average MAEs and PSI values for msDyNODE (black: firing rate, blue: LFP) and NetPyNE (red). Dim points represent average MAEs and PSI over trials for each recording channel. \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$  using two-sided Wilcoxon's rank-sum test. (TIF)



**S2 Fig. Benchmark with M-SGM and SBI-SGM.** Periodograms of MAEs in frequency responses spanning from 0 to 40 Hz.

(TIF)

**S3 Fig.** Granger causality-based graph properties over eight different target directions for Monkey A and B. Number of edges, average clustering, and number of total triangles derived from Granger causality-based excitatory (blue) and inhibitory (red) subnetworks are presented in polar coordinated for Monkey A (top) and B (bottom), respectively.

(TIF)

**S4 Fig. DTF-based graph properties over eight different target directions for Monkey A and B.** Number of edges, average clustering, and number of total triangles derived from Granger causality-based network are presented in polar coordinated for Monkey A (top) and B (bottom), respectively.

(TIF)

## Acknowledgments

We thank José del R. Millán from Clinical Neuroprosthetics and Brain Interaction lab at University of Texas at Austin for extensive discussion and suggestions.

## Author Contributions

**Conceptualization:** Yin-Jui Chang, Samantha R. Santacruz.

**Data curation:** Yin-Jui Chang, Hannah M. Stealey, Yi Zhao, Hung-Yun Lu, Enrique Contreras-Hernandez, Megan N. Baker, Samantha R. Santacruz.

**Formal analysis:** Yin-Jui Chang, Yuan-I Chen, Samantha R. Santacruz.

**Funding acquisition:** Hsin-Chih Yeh, Samantha R. Santacruz.

**Investigation:** Yin-Jui Chang, Hannah M. Stealey, Yi Zhao, Hung-Yun Lu, Enrique Contreras-Hernandez, Megan N. Baker.

**Methodology:** Yin-Jui Chang, Yuan-I Chen, Edward Castillo, Hsin-Chih Yeh, Samantha R. Santacruz.

**Project administration:** Hsin-Chih Yeh, Samantha R. Santacruz.

**Resources:** Samantha R. Santacruz.

**Software:** Yin-Jui Chang.

**Supervision:** Edward Castillo, Hsin-Chih Yeh, Samantha R. Santacruz.

**Validation:** Yin-Jui Chang.

**Visualization:** Yin-Jui Chang.

**Writing – original draft:** Yin-Jui Chang, Samantha R. Santacruz.

**Writing – review & editing:** Yin-Jui Chang, Yuan-I Chen, Hannah M. Stealey, Yi Zhao, Hung-Yun Lu, Enrique Contreras-Hernandez, Megan N. Baker, Edward Castillo, Hsin-Chih Yeh, Samantha R. Santacruz.

## References

1. Presigny C, De Vico Fallani F. Colloquium: Multiscale modeling of brain network organization. *Rev Mod Phys*. 2022 Aug 2; 94(3):031002.
2. Friston KJ. Functional and Effective Connectivity: A Review. *Brain Connectivity*. 2011 Jan 1; 1(1):13–36. <https://doi.org/10.1089/brain.2011.0008> PMID: 22432952
3. Riedl V, Utz L, Castrillón G, Grimmer T, Rauschecker JP, Ploner M, et al. Metabolic connectivity mapping reveals effective connectivity in the resting human brain. *Proceedings of the National Academy of Sciences*. 2016 Jan 12; 113(2):428–33.
4. Churchland MM, Cunningham JP, Kaufman MT, Ryu SI, Shenoy KV. Cortical Preparatory Activity: Representation of Movement or First Cog in a Dynamical Machine? *Neuron*. 2010 Nov 4; 68(3):387–400. <https://doi.org/10.1016/j.neuron.2010.09.015> PMID: 21040842
5. Vyas S, O'Shea DJ, Ryu SI, Shenoy KV. Causal Role of Motor Preparation during Error-Driven Learning. *Neuron*. 2020 Apr 22; 106(2):329–339.e4. <https://doi.org/10.1016/j.neuron.2020.01.019> PMID: 32053768
6. Mauk MD, Buonomano DV. The Neural Basis of Temporal Processing. *Annual Review of Neuroscience*. 2004; 27(1):307–40. <https://doi.org/10.1146/annurev.neuro.27.070203.144247> PMID: 15217335
7. Chaisangmongkon W, Swaminathan SK, Freedman DJ, Wang XJ. Computing by Robust Transience: How the Fronto-Parietal Network Performs Sequential, Category-Based Decisions. *Neuron*. 2017 Mar 22; 93(6):1504–1517.e4. <https://doi.org/10.1016/j.neuron.2017.03.002> PMID: 28334612
8. Chaudhuri R, Fiete I. Computational principles of memory. *Nature Neuroscience*. 2016 Mar; 19(3):394–403. <https://doi.org/10.1038/nn.4237> PMID: 26906506
9. Macke JH, Buesing L, Sahani M. Estimating state and parameters in state space models of spike trains. In: Chen Z, editor. *Advanced State Space Methods for Neural and Clinical Data* [Internet]. Cambridge: Cambridge University Press; 2015 [cited 2020 Sep 9]. p. 137–59. Available from: [https://www.cambridge.org/core/product/identifier/CBO9781139941433A054/type/book\\_part](https://www.cambridge.org/core/product/identifier/CBO9781139941433A054/type/book_part)
10. Byron MY, Cunningham JP, Santhanam G, Ryu SI, Shenoy KV, Sahani M. Gaussian-Process Factor Analysis for Low-Dimensional Single-Trial Analysis of Neural Population Activity. *Journal of Neurophysiology*. 2009 Jul 1; 102(1):614–35. <https://doi.org/10.1152/jn.90941.2008> PMID: 19357332
11. Buschman TJ, Kastner S. From behavior to neural dynamics: An integrated theory of attention. *Neuron*. 2015 Oct 7; 88(1):127–44. <https://doi.org/10.1016/j.neuron.2015.09.017> PMID: 26447577
12. Harbecke J. The methodological role of mechanistic-computational models in cognitive science. *Synthese*. 2020 Feb 17;1–23.
13. Harrison L, Penny WD, Friston K. Multivariate autoregressive modeling of fMRI time series. *NeuroImage*. 2003 Aug 1; 19(4):1477–91. [https://doi.org/10.1016/s1053-8119\(03\)00160-5](https://doi.org/10.1016/s1053-8119(03)00160-5) PMID: 12948704
14. Goebel R, Roebroeck A, Kim DS, Formisano E. Investigating directed cortical interactions in time-resolved fMRI data using vector autoregressive modeling and Granger causality mapping. *Magnetic Resonance Imaging*. 2003 Dec 1; 21(10):1251–61. <https://doi.org/10.1016/j.mri.2003.08.026> PMID: 14725933
15. Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ. Psychophysiological and Modulatory Interactions in Neuroimaging. *NeuroImage*. 1997 Oct 1; 6(3):218–29. <https://doi.org/10.1006/nimg.1997.0291> PMID: 9344826
16. McIntosh AR, Gonzalez-Lima F. Structural modeling of functional neural pathways mapped with 2-deoxyglucose: effects of acoustic startle habituation on the auditory system. *Brain Research*. 1991 May 3; 547(2):295–302. [https://doi.org/10.1016/0006-8993\(91\)90974-z](https://doi.org/10.1016/0006-8993(91)90974-z) PMID: 1884204
17. Büchel C, Friston KJ. Modulation of connectivity in visual pathways by attention: cortical interactions evaluated with structural equation modelling and fMRI. *Cerebral Cortex (New York, NY: 1991)*. 1997; 7(8):768–78. <https://doi.org/10.1093/cercor/7.8.768> PMID: 9408041
18. Bullmore E, Horwitz B, Honey G, Brammer M, Williams S, Sharma T. How Good Is Good Enough in Path Analysis of fMRI Data? *NeuroImage*. 2000 Apr 1; 11(4):289–301. <https://doi.org/10.1006/nimg.2000.0544> PMID: 10725185
19. McIntosh AR, Gonzalez-Lima F. Structural equation modeling and its application to network analysis in functional brain imaging. *Human Brain Mapping*. 1994; 2(1–2):2–22.
20. Penny WD, Stephan KE, Mechelli A, Friston KJ. Modelling functional integration: a comparison of structural equation and dynamic causal models. *Neuroimage*. 2004 Jan 1; 23 Suppl 1:S264–74. <https://doi.org/10.1016/j.neuroimage.2004.07.041> PMID: 15501096
21. Michel CM, Brunet D. EEG Source Imaging: A Practical Review of the Analysis Steps. *Front Neurol*. 2019; 10:325. <https://doi.org/10.3389/fneur.2019.00325> PMID: 31019487

22. Canolty RT, Ganguly K, Carmena JM. Task-Dependent Changes in Cross-Level Coupling between Single Neurons and Oscillatory Activity in Multiscale Networks. *PLOS Computational Biology*. 2012 Dec 20; 8(12):e1002809. <https://doi.org/10.1371/journal.pcbi.1002809> PMID: 23284276
23. Wang C, Pesaran B, Shanechi MM. Modeling multiscale causal interactions between spiking and field potential signals during behavior. *J Neural Eng*. 2022 Mar; 19(2):026001. <https://doi.org/10.1088/1741-2552/ac4e1c> PMID: 35073530
24. Characterization of regional differences in resting-state fMRI with a data-driven network model of brain dynamics | *Science Advances* [Internet]. [cited 2023 Aug 1]. Available from: <https://www.science.org/doi/10.1126/sciadv.abq7547>
25. Heitmann S, Breakspear M. Putting the “dynamic” back into dynamic functional connectivity. *Network Neuroscience*. 2018 Jun 1; 02(02):150–74. [https://doi.org/10.1162/netn\\_a\\_00041](https://doi.org/10.1162/netn_a_00041) PMID: 30215031
26. Chang YJ, Chen YI, Yeh HC, Santacruz SR. Neurobiologically realistic neural network enables cross-scale modeling of neural dynamics. *Sci Rep*. 2024 Mar 1; 14(1):5145. <https://doi.org/10.1038/s41598-024-54593-w> PMID: 38429297
27. Chen RTQ, Rubanova Y, Bettencourt J, Duvenaud D. Neural Ordinary Differential Equations. arXiv:180607366 [cs, stat] [Internet]. 2019 Dec 13 [cited 2022 Apr 28]; Available from: <http://arxiv.org/abs/1806.07366>
28. Zhao Y, Park IM. Variational Latent Gaussian Process for Recovering Single-Trial Dynamics from Population Spike Trains. *Neural Computation*. 2017 May; 29(5):1293–316. [https://doi.org/10.1162/NECO\\_a\\_00953](https://doi.org/10.1162/NECO_a_00953) PMID: 28333587
29. Pandarinath C, O'Shea DJ, Collins J, Jozefowicz R, Stavisky SD, Kao JC, et al. Inferring single-trial neural population dynamics using sequential auto-encoders. *Nature Methods*. 2018 Oct; 15(10):805–15. <https://doi.org/10.1038/s41592-018-0109-9> PMID: 30224673
30. Ganguly K, Carmena JM. Emergence of a Stable Cortical Map for Neuroprosthetic Control. *PLOS Biology*. 2009 Jul 21; 7(7):e1000153. <https://doi.org/10.1371/journal.pbio.1000153> PMID: 19621062
31. Athalye VR, Ganguly K, Costa RM, Carmena JM. Emergence of Coordinated Neural Dynamics Underlies Neuroprosthetic Learning and Skillful Control. *Neuron*. 2017 Feb 22; 93(4):955–970.e5. <https://doi.org/10.1016/j.neuron.2017.01.016> PMID: 28190641
32. Golub MD, Sadtler PT, Oby ER, Quick KM, Ryu SI, Tyler-Kabara EC, et al. Learning by neural reassociation. *Nat Neurosci*. 2018 Apr; 21(4):607–16. <https://doi.org/10.1038/s41593-018-0095-3> PMID: 29531364
33. Oby ER, Golub MD, Hennig JA, Degenhart AD, Tyler-Kabara EC, Yu BM, et al. New neural activity patterns emerge with long-term learning. *Proceedings of the National Academy of Sciences*. 2019 Jul 23; 116(30):15210–5. <https://doi.org/10.1073/pnas.1820296116> PMID: 31182595
34. Zippi EL, You AK, Ganguly K, Carmena JM. Selective modulation of cortical population dynamics during neuroprosthetic skill learning. *Sci Rep*. 2022 Sep 24; 12(1):15948. <https://doi.org/10.1038/s41598-022-20218-3> PMID: 36153356
35. Sadtler PT, Quick KM, Golub MD, Chase SM, Ryu SI, Tyler-Kabara EC, et al. Neural constraints on learning. *Nature*. 2014 Aug; 512(7515):423–6. <https://doi.org/10.1038/nature13665> PMID: 25164754
36. Dayan P, Abbott LF. *Theoretical neuroscience*, vol. 806. 2001;
37. Nordbø Ø, Wyller J, Einevoll GT. Neural network firing-rate models on integral form. *Biol Cybern*. 2007 Sep 1; 97(3):195–209.
38. Nordlie E, Tetzlaff T, Einevoll G. Rate Dynamics of Leaky Integrate-and-Fire Neurons with Strong Synapses. *Frontiers in Computational Neuroscience* [Internet]. 2010 [cited 2023 Apr 18]; 4. Available from: <https://www.frontiersin.org/articles/10.3389/fncom.2010.00149> PMID: 21212832
39. Jansen BH, Rit VG. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol Cybern*. 1995 Sep 1; 73(4):357–66. <https://doi.org/10.1007/BF00199471> PMID: 7578475
40. Fell J, Axmacher N. The role of phase synchronization in memory processes. *Nat Rev Neurosci*. 2011 Feb; 12(2):105–18. <https://doi.org/10.1038/nrn2979> PMID: 21248789
41. Nichols MJ, Newsome WT. The neurobiology of cognition. *Nature*. 1999 Dec; 402(6761):C35–8. <https://doi.org/10.1038/35011531> PMID: 10591223
42. Lisman J. The challenge of understanding the brain: where we stand in 2015. *Neuron*. 2015 May 20; 86(4):864–82. <https://doi.org/10.1016/j.neuron.2015.03.032> PMID: 25996132
43. Abbott LF. *Theoretical Neuroscience Rising*. *Neuron*. 2008 Nov 6; 60(3):489–95. <https://doi.org/10.1016/j.neuron.2008.10.019> PMID: 18995824
44. Kriegeskorte N, Douglas PK. Cognitive computational neuroscience. *Nature Neuroscience*. 2018 Sep; 21(9):1148–60. <https://doi.org/10.1038/s41593-018-0210-5> PMID: 30127428

45. McKenna TM, McMullen TA, Shlesinger MF. The brain as a dynamic physical system. *Neuroscience*. 1994 Jun 1; 60(3):587–605. [https://doi.org/10.1016/0306-4522\(94\)90489-8](https://doi.org/10.1016/0306-4522(94)90489-8) PMID: 7936189
46. Freeman WJ. Mesoscopic neurodynamics: From neuron to brain. *Journal of Physiology-Paris*. 2000 Dec 1; 94(5):303–22. [https://doi.org/10.1016/s0928-4257\(00\)01090-1](https://doi.org/10.1016/s0928-4257(00)01090-1) PMID: 11165902
47. Stam CJ. Nonlinear dynamical analysis of EEG and MEG: Review of an emerging field. *Clinical Neurophysiology*. 2005 Oct 1; 116(10):2266–301. <https://doi.org/10.1016/j.clinph.2005.06.011> PMID: 16115797
48. Breakspear M. Dynamic models of large-scale brain activity. *Nature Neuroscience*. 2017 Mar; 20(3):340–52. <https://doi.org/10.1038/nn.4497> PMID: 28230845
49. Roberts JA, Gollo LL, Abeysuriya RG, Roberts G, Mitchell PB, Woolrich MW, et al. Metastable brain waves. *Nat Commun*. 2019 Mar 5; 10(1):1056. <https://doi.org/10.1038/s41467-019-08999-0> PMID: 30837462
50. Bansal K, Garcia JO, Thompson SH, Verstynen T, Vettel JM, Muldoon SF. Cognitive chimera states in human brain networks. *Science Advances*. 2019 Apr 3; 5(4):eaau8535. <https://doi.org/10.1126/sciadv.aau8535> PMID: 30949576
51. Lynn CW, Bassett DS. The physics of brain network structure, function and control. *Nat Rev Phys*. 2019 May; 1(5):318–32.
52. Chauhan R, Ghanshala KK, Joshi RC. Convolutional Neural Network (CNN) for Image Detection and Recognition. In: 2018 First International Conference on Secure Cyber Computing and Communication (ICSCCC) [Internet]. 2018 [cited 2024 Oct 17]. p. 278–82. Available from: <https://ieeexplore.ieee.org/abstract/document/8703316>
53. Pan Y, Wang J. Model Predictive Control of Unknown Nonlinear Dynamical Systems Based on Recurrent Neural Networks. *IEEE Transactions on Industrial Electronics*. 2012 Aug; 59(8):3089–101.
54. Wolf T, Debut L, Sanh V, Chaumond J, Delangue C, Moi A, et al. Transformers: State-of-the-Art Natural Language Processing. In: Liu Q, Schlangen D, editors. *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations* [Internet]. Online: Association for Computational Linguistics; 2020 [cited 2024 Oct 17]. p. 38–45. Available from: <https://aclanthology.org/2020.emnlp-demos.6>
55. Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, et al. Generative Adversarial Nets. In: *Advances in Neural Information Processing Systems* [Internet]. Curran Associates, Inc.; 2014 [cited 2023 Apr 26]. Available from: [https://papers.nips.cc/paper\\_files/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html](https://papers.nips.cc/paper_files/paper/2014/hash/5ca3e9b122f61f8f06494c97b1afccf3-Abstract.html)
56. Chen YI, Chang YJ, Liao SC, Nguyen TD, Yang J, Kuo YA, et al. Generative adversarial network enables rapid and robust fluorescence lifetime image analysis in live cells. *Commun Biol*. 2022 Jan 11; 5(1):1–11.
57. Voelker A, Kajić I, Eliasmith C. Legendre Memory Units: Continuous-Time Representation in Recurrent Neural Networks. In: *Advances in Neural Information Processing Systems* [Internet]. Curran Associates, Inc.; 2019 [cited 2023 Aug 1]. Available from: <https://proceedings.neurips.cc/paper/2019/hash/952285b9b7e7a1be5aa7849f32ffff05-Abstract.html>
58. Chang B, Chen M, Haber E, Chi EH. AntisymmetricRNN: A Dynamical System View on Recurrent Neural Networks [Internet]. arXiv; 2019 [cited 2023 Aug 1]. Available from: <http://arxiv.org/abs/1902.09689>
59. De Brouwer E, Simm J, Arany A, Moreau Y. GRU-ODE-Bayes: Continuous modeling of sporadically-observed time series. arXiv:190512374 [cs, stat] [Internet]. 2019 Nov 28 [cited 2022 Apr 28]; Available from: <http://arxiv.org/abs/1905.12374>
60. Kang H. The prevention and handling of the missing data. *Korean J Anesthesiol*. 2013 May 24; 64(5):402–6. <https://doi.org/10.4097/kjae.2013.64.5.402> PMID: 23741561
61. Dura-Bernal S, Suter BA, Gleeson P, Cantarelli M, Quintana A, Rodriguez F, et al. NetPyNE, a tool for data-driven multiscale modeling of brain circuits. *Bhalla US, Calabrese RL, Sterratt D, Wójcik DK, editors. eLife*. 2019 Apr 26; 8:e44494. <https://doi.org/10.7554/eLife.44494> PMID: 31025934
62. Verma P, Nagarajan S, Raj A. Spectral graph theory of brain oscillations—Revisited and improved. *NeuroImage*. 2022 Apr 1; 249:118919. <https://doi.org/10.1016/j.neuroimage.2022.118919> PMID: 35051584
63. Jin H, Verma P, Jiang F, Nagarajan SS, Raj A. Bayesian inference of a spectral graph model for brain oscillations. *NeuroImage*. 2023 Oct 1; 279:120278. <https://doi.org/10.1016/j.neuroimage.2023.120278> PMID: 37516373
64. Carnevale NT, Hines ML. *The NEURON book*. Cambridge University Press; 2006.
65. Tikidji-Hamburyan RA, Narayana V, Bozkus Z, El-Ghazawi TA. Software for Brain Network Simulations: A Comparative Study. *Front Neuroinform* [Internet]. 2017 Jul 20 [cited 2024 Oct 17]; 11. Available from:



<https://www.frontiersin.org/journals/neuroinformatics/articles/10.3389/fninf.2017.00046/full> PMID: 28775687

66. Stein RB, Hodgkin AL. The frequency of nerve action potentials generated by applied currents. *Proceedings of the Royal Society of London Series B Biological Sciences*. 1997 Jan; 167(1006):64–86.
67. Potjans TC, Diesmann M. The Cell-Type Specific Cortical Microcircuit: Relating Structure and Activity in a Full-Scale Spiking Network Model. *Cerebral Cortex*. 2014 Mar 1; 24(3):785–806. <https://doi.org/10.1093/cercor/bhs358> PMID: 23203991
68. Montbrió E, Pazó D, Roxin A. Macroscopic Description for Networks of Spiking Neurons. *Phys Rev X*. 2015 Jun 19; 5(2):021028.
69. Schmidt H, Avitabile D, Montbrió E, Roxin A. Network mechanisms underlying the role of oscillations in cognitive tasks. *PLOS Computational Biology*. 2018 Sep 6; 14(9):e1006430. <https://doi.org/10.1371/journal.pcbi.1006430> PMID: 30188889
70. Wilson HR, Cowan JD. Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys J*. 1972 Jan; 12(1):1–24. [https://doi.org/10.1016/S0006-3495\(72\)86068-5](https://doi.org/10.1016/S0006-3495(72)86068-5) PMID: 4332108
71. Arnold TW. Uninformative Parameters and Model Selection Using Akaike's Information Criterion. *The Journal of Wildlife Management*. 2010; 74(6):1175–8.
72. Vrieze SI. Model selection and psychological theory: A discussion of the differences between the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). *Psychological Methods*. 2012; 17(2):228–43. <https://doi.org/10.1037/a0027127> PMID: 22309957
73. Chen J, Chen Z. Extended Bayesian information criteria for model selection with large model spaces. *Biometrika*. 2008 Sep 1; 95(3):759–71.
74. Grünwald P. Model Selection Based on Minimum Description Length. *Journal of Mathematical Psychology*. 2000 Mar; 44(1):133–52. <https://doi.org/10.1006/jmps.1999.1280> PMID: 10733861
75. Hansen MH, Yu B. Model selection and the principle of minimum description length. *Journal of the American Statistical Association*. 2001; 96(454):746–74.
76. Pan SJ, Yang Q. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*. 2010; 22(10):1345–59.
77. Sternad D. It's not (only) the mean that matters: variability, noise and exploration in skill learning. *Current Opinion in Behavioral Sciences*. 2018 Apr 1; 20:183–95. <https://doi.org/10.1016/j.cobeha.2018.01.004> PMID: 30035207
78. Waschke L, Kloosterman NA, Obleser J, Garrett DD. Behavior needs neural variability. *Neuron*. 2021 Mar 3; 109(5):751–66. <https://doi.org/10.1016/j.neuron.2021.01.023> PMID: 33596406
79. Sun X, O'Shea DJ, Golub MD, Trautmann EM, Vyas S, Ryu SI, et al. Cortical preparatory activity indexes learned motor memories. *Nature*. 2022 Jan 26; 1–6. <https://doi.org/10.1038/s41586-021-04329-x> PMID: 35082444
80. Churchland MM, Yu BM, Cunningham JP, Sugrue LP, Cohen MR, Corrado GS, et al. Stimulus onset quenches neural variability: a widespread cortical phenomenon. *Nat Neurosci*. 2010 Mar; 13(3):369–78. <https://doi.org/10.1038/nn.2501> PMID: 20173745
81. Dhawale AK, Smith MA, Ölveczky BP. The Role of Variability in Motor Learning. *Annual Review of Neuroscience*. 2017; 40(1):479–98. <https://doi.org/10.1146/annurev-neuro-072116-031548> PMID: 28489490
82. Granger CWJ. Investigating Causal Relations by Econometric Models and Cross-spectral Methods. *Econometrica*. 1969; 37(3):424–38.
83. Kaminski MJ, Blinowska KJ. A new method of the description of the information flow in the brain structures. *Biol Cybern*. 1991 Jul 1; 65(3):203–10. <https://doi.org/10.1007/BF00198091> PMID: 1912013
84. Wu W, Gao Y, Bienenstock E, Donoghue JP, Black MJ. Bayesian Population Decoding of Motor Cortical Activity Using a Kalman Filter. *Neural Computation*. 2006 Jan 1; 18(1):80–118. <https://doi.org/10.1162/089976606774841585> PMID: 16354382
85. Dangi S, Gowda S, Moorman HG, Orsborn AL, So K, Shanechi M, et al. Continuous closed-loop decoder adaptation with a recursive maximum likelihood algorithm allows for rapid performance acquisition in brain-machine interfaces. *Neural Comput*. 2014 Sep; 26(9):1811–39. [https://doi.org/10.1162/NECO\\_a\\_00632](https://doi.org/10.1162/NECO_a_00632) PMID: 24922501
86. Orsborn AL, Moorman HG, Overduin SA, Shanechi MM, Dimitrov DF, Carmena JM. Closed-Loop Decoder Adaptation Shapes Neural Plasticity for Skillful Neuroprosthetic Control. *Neuron*. 2014 Jun 18; 82(6):1380–93. <https://doi.org/10.1016/j.neuron.2014.04.048> PMID: 24945777
87. Shanechi MM. Brain-machine interfaces from motor to mood. *Nature Neuroscience*. 2019 Oct; 22(10):1554–64. <https://doi.org/10.1038/s41593-019-0488-y> PMID: 31551595

88. Shenoy KV, Carmena JM. Combining Decoder Design and Neural Adaptation in Brain-Machine Interfaces. *Neuron*. 2014 Nov 19; 84(4):665–80. <https://doi.org/10.1016/j.neuron.2014.08.038> PMID: 25459407
89. Orsborn AL, Pesaran B. Parsing learning in networks using brain-machine interfaces. *Curr Opin Neurobiol*. 2017 Oct; 46:76–83. <https://doi.org/10.1016/j.conb.2017.08.002> PMID: 28843838
90. Schwarz DA, Lebedev MA, Hanson TL, Dimitrov DF, Leheve G, Meloy J, et al. Chronic, wireless recordings of large-scale brain activity in freely moving rhesus monkeys. *Nat Methods*. 2014 Jun; 11(6):670–6. <https://doi.org/10.1038/nmeth.2936> PMID: 24776634
91. Ganguly K, Dimitrov DF, Wallis JD, Carmena JM. Reversible large-scale modification of cortical networks during neuroprosthetic control. *Nat Neurosci*. 2011 May; 14(5):662–7. <https://doi.org/10.1038/nn.2797> PMID: 21499255
92. Inoue Y, Mao H, Suway SB, Orellana J, Schwartz AB. Decoding arm speed during reaching. *Nat Commun*. 2018 Dec 7; 9(1):5243. <https://doi.org/10.1038/s41467-018-07647-3> PMID: 30531921
93. Pulvermüller F, Tomasello R, Henningsen-Schomers MR, Wennekers T. Biological constraints on neural network models of cognitive function. *Nat Rev Neurosci*. 2021 Aug; 22(8):488–502. <https://doi.org/10.1038/s41583-021-00473-5> PMID: 34183826
94. Linderman S, Johnson M, Miller A, Adams R, Blei D, Paninski L. Bayesian learning and inference in recurrent switching linear dynamical systems. In *PMLR*; 2017. p. 914–22.
95. Gray R, Robinson P. Stability constraints on large-scale structural brain networks. *Frontiers in Computational Neuroscience* [Internet]. 2013 [cited 2023 Aug 1]; 7. Available from: <https://www.frontiersin.org/articles/10.3389/fncom.2013.00031> PMID: 23630490
96. Xu T, Barak O. Dynamical Timescale Explains Marginal Stability in Excitability Dynamics. *J Neurosci*. 2017 Apr 26; 37(17):4508–24. <https://doi.org/10.1523/JNEUROSCI.2340-16.2017> PMID: 28348138
97. Pedersen M, Omidvarnia A, Walz JM, Zalesky A, Jackson GD. Spontaneous brain network activity: Analysis of its temporal complexity. *Network Neuroscience*. 2017 Jun 1; 1(2):100–15. [https://doi.org/10.1162/NETN\\_a\\_00006](https://doi.org/10.1162/NETN_a_00006) PMID: 29911666