ADVANCES IN COMPUTATIONAL NEUROSCIENCE

EDITED BY: Thomas Nowotny, Sacha Jennifer van Albada, Tatyana Sharpee, Jean-Marc Fellous, Renaud Blaise Jolivet, Julie S. Haas and Christoph Metzner

PUBLISHED IN: Frontiers in Computational Neuroscience

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88974-501-2
DOI 10.3389/978-2-88974-501-2

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

ADVANCES IN COMPUTATIONAL NEUROSCIENCE

Topic Editors:
Thomas Nowotny, University of Sussex, United Kingdom
Sacha Jennifer van Albada, Julich Research Center, Helmholtz Association of German Research Centres (HZ), Germany
Tatyana Sharpee, Salk Institute for Biological Studies, United States
Jean-Marc Fellous, University of Arizona, United States
Renaud Blaise Jolivet, Université de Genève, Switzerland
Julie S. Haas, Lehigh University, United States
Christoph Metzner, Technical University of Berlin, Germany

Citation: Nowotny, T., van Albada, S. J., Sharpee, T., Fellous, J.-M., Jolivet, R. B., Haas, J. S., Metzner, C., eds. (2022). Advances in Computational Neuroscience. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88974-501-2

Table of Contents

frontiers in Computational Neuroscience

EDITORIAL
published: 24 January 2022
doi: 10.3389/fncom.2021.824899

Editorial: Advances in Computational Neuroscience

Thomas Nowotny[1], Sacha J. van Albada[2,3], Jean-Marc Fellous[4], Julie S. Haas[5], Renaud B. Jolivet[6], Christoph Metzner[7,8,9] and Tatyana Sharpee[10,11]*

[1] School of Engineering and Informatics, University of Sussex, Brighton, United Kingdom, [2] Jülich Research Centre, Institute of Neuroscience and Medicine (INM-6) and Institute for Advanced Simulation (IAS-6) and JARA-Institute Brain Structure-Function Relationships (INM-10), Jülich, Germany, [3] Department of Biology, Faculty of Mathematics and Natural Sciences, Institute of Zoology, University of Cologne, Cologne, Germany, [4] Departments of Psychology and Biomedical Engineering, University of Arizona, Tucson, AZ, United States, [5] Department of Biological Sciences, Lehigh University, Bethlehem, PA, United States, [6] Maastricht Centre for Systems Biology, Maastricht University, Maastricht, Netherlands, [7] Department of Electrical Engineering and Computer Science, Institute of Software Engineering and Theoretical Computer Science, Technische Universit4t Berlin, Berlin, Germany, [8] Department of Child and Adolescent Psychiatry, Charité Universitätsmedizin, Berlin, Germany, [9] School of Physics, Engineering and Computer Science, University of Hertfordshire, Hatfield, United Kingdom, [10] Computational Neurobiology Laboratory, Salk Institute for Biological Studies, La Jolla, CA, United States, [11] Department of Physics, University of California, San Diego, La Jolla, CA, United States

Keywords: computational neuroscience, computational modeling, data analysis, brain circuits, neuronal networks, neural dynamics

Editorial on the Research Topic

Advances in Computational Neuroscience

The 28th Annual Computational Neuroscience Meeting CNS*2019 took place from 13 to 17 July 2019 in the city of Barcelona. The conference encompassed a wide diversity of Research Topics and welcomed participants from around the world, with keynotes on "Brain networks, adolescence, and schizophrenia" by Professor Ed Bullmore, "Neural circuits for mental simulation" by Professor Kenji Doya, "One network, many states: varying the excitability of the cerebral cortex" by Professor Maria Sanchez-Vives, and "Neural circuits for flexible memory and navigation" by Professor Ila Fiete. The present Research Topic, "*Advances in Computational Neuroscience*," contains some of the leading-edge Computational Neuroscience research presented and discussed at the conference.

Like CNS*2019, the articles in this Research Topic reflect the diversity and richness of computational neuroscience research, expanding from subcellular scales to networks, from biological details to *in silico* technology, and from computational methods to brain theory.

At the sub-neuronal level, in "ROOTS: An algorithm to generate biologically realistic cortical axons and an application to electroceutical modeling," Bingham et al. develop computational methods for building more accurate computational models, extending the capability of generative methods for producing neuronal morphology of highly branched cortical axon terminal arbors. In a similar domain, in "Serotonergic Axons as Fractional Brownian Motion Paths: Insights into the Self-organization of Regional Densities," Janušonis et al. describe how a computational model based on reflected Fractional Brownian Motion can generate steady-state distributions that approximate the experimentally observed serotonergic fiber distributions in physical brain sections. Gontier and Pfister, in "Identifiability of a binomial synapse," expand model principles by introducing a definition of when a statistical model is practically identifiable and apply this concept to models of synapses. Felton et al. in "Assessing the Impact of I_h Conductance on Cross-Frequency Coupling in Model Pyramidal Neurons," analyze the role of the hyperpolarization-activated mixed cation current (I_h) in the dynamical phenomenon of cross-frequency coupling. In a similar vein, Mergenthal et al. in "A Computational Model of the Cholinergic Modulation of CA1 Pyramidal Cell Activity," present a computational model of pyramidal cells that includes unprecedented detail

Edited and reviewed by:
Si Wu,
Peking University, China

****Correspondence:***
Thomas Nowotny
t.nowotny@sussex.ac.uk

Received: *29 November 2021*
Accepted: *06 December 2021*
Published: *24 January 2022*

Citation:
Nowotny T, van Albada SJ, Fellous J-M, Haas JS, Jolivet RB, Metzner C and Sharpee T (2022) Editorial: Advances in Computational Neuroscience. Front. Comput. Neurosci. 15:824899. doi: 10.3389/fncom.2021.824899

on receptor activations in response to changes in extracellular acetylcholine concentration, and describe their effects on cellular excitability and downstream intracellular calcium dynamics.

Other authors work at the network level. In "From topological analyses to functional modeling: the case of hippocampus," Dabaghian presents an approach to network modeling in which a combination of topological analyses provides insights into information processing in mammalian hippocampus. Similarly, Hasanzadeh et al. in "Necessary Conditions for Reliable Propagation of Time-Varying Firing Rate," investigate which network properties in feed-forward networks allow stable propagation of asynchronous spikes, while Zirkle and Rubchinsky, in "Spike-Timing Dependent Plasticity Effect on the Temporal Patterning of Neural Synchronization," investigate in a small network model how plasticity of synapses can alter synchronization dynamics and induce intermittent synchronization akin to experimental observations. Tian and Zhou, on the other hand, focus more on simulation algorithms in "Exponential Time Differencing Algorithm For Pulse-coupled Hodgkin-Huxley Neural Networks."

A third group of works relates computational modeling with higher-level experimental observations. Endo et al. in "Evaluation of resting spatio-temporal dynamics of a neural mass model using resting fMRI connectivity and EEG microstates," use the Larter-Breakspear neural mass model to relate to both, fast EEG/MEG microstates and slow fluctuations observed with fMRI. Relatedly, Don et al. in "Topological View of Flows inside the BOLD Spontaneous Activity of the Human Brain," use computational topology of data and discover previously unknown vortex structures in activated brain regions. Finally, within this category, Li et al. in "Effects of cholinergic neuromodulation on thalamocortical rhythms during NREM sleep: a model study," study a new model of the thalamo-cortical network with cholinergic modulation that exhibits the hallmarks of NREM sleep and allows formulating hypotheses for the role of this important neuromodulator in generating these hallmarks.

Other works are inspired by principles from statistical physics and information theory. For instance, in "A computational framework for controlling the self-restorative brain based on the free energy and degeneracy principles," Park and Kang speculate about a method based on the free energy principle to identify ways of using the brain's self-restoration capabilities to induce desired changes in brain activity. Sorooshyari et al. in "Object Recognition at Higher Regions of the Ventral Visual Stream via Dynamic Inference," hypothesize about the computations performed by the ventral visual stream during object recognition based on dynamic inference, and present simulations of object identification by inferior temporal cortex. On the other hand, Yamakou et al. in "Optimal self-induced stochastic resonance in multiplex neural networks: electrical vs. chemical synapses," are inspired by stochastic resonance, and investigate in a computational model the differential roles of electrical and chemical synapses and their properties for supporting this mechanism. Finally, Ludl and Soriano, in "Impact of physical obstacles on the structural and effective connectivity of *in silico* neuronal circuits," take us all the way to cultured brain circuits and find that in their simulations physical obstacles placed into the path of neurons growing on silicon substrates lead to the formation of local effective microcircuits.

While the breadth and diversity of the submitted work could hardly be larger, there are also striking commonalities. Many of the works draw on a rich set of advances from several disciplines. This may reflect the typically multi-disciplinary background of the author teams, but we believe is also a sign of the future of the maturing field of computational neuroscience, which has long since moved beyond building models of isolated aspects of brains, computation, or behavior. Our future lies in the cross-disciplinary amalgamation of methods, theories, models, and experimental data. In this spirit, we will continue to be an open and welcoming community as the Computational Neuroscience Meeting enters its 4th decade.

AUTHOR CONTRIBUTIONS

TN wrote the initial draft. SA, J-MF, JH, RJ, CM, and TS edited and approved the manuscript.

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.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 17 January 2020
doi: 10.3389/fncom.2019.00091

Evaluation of Resting Spatio-Temporal Dynamics of a Neural Mass Model Using Resting fMRI Connectivity and EEG Microstates

Hidenori Endo[1,2], *Nobuo Hiroe*[2] *and Okito Yamashita*[2,3]*

[1] *Graduate School of Information Science, Nara Institute of Science and Technology, Nara, Japan,* [2] *ATR Neural Information Analysis Laboratories, Kyoto, Japan,* [3] *Center for Advanced Intelligence Project, RIKEN, Tokyo, Japan*

Resting-state brain activities have been extensively investigated to understand the macro-scale network architecture of the human brain using non-invasive imaging methods such as fMRI, EEG, and MEG. Previous studies revealed a mechanistic origin of resting-state networks (RSNs) using the connectome dynamics modeling approach, where the neural mass dynamics model constrained by the structural connectivity is simulated to replicate the resting-state networks measured with fMRI and/or fast synchronization transitions with EEG/MEG. However, there is still little understanding of the relationship between the slow fluctuations measured with fMRI and the fast synchronization transitions with EEG/MEG. In this study, as a first step toward evaluating experimental evidence of resting state activity at two different time scales but in a unified way, we investigate connectome dynamics models that simultaneously explain resting-state functional connectivity (rsFC) and EEG microstates. Here, we introduce empirical rsFC and microstates as evaluation criteria of simulated neuronal dynamics obtained by the Larter-Breakspear model in one cortical region connected with those in other cortical regions based on structural connectivity. We optimized the global coupling strength and the local gain parameter (variance of the excitatory and inhibitory threshold) of the simulated neuronal dynamics by fitting both rsFC and microstate spatial patterns to those of experimental ones. As a result, we found that simulated neuronal dynamics in a narrow optimal parameter range simultaneously reproduced empirical rsFC and microstates. Two parameter groups had different inter-regional interdependence. One type of dynamics was synchronized across the whole brain region, and the other type was synchronized between brain regions with strong structural connectivity. In other words, both fast synchronization transitions and slow BOLD fluctuation changed based on structural connectivity in the two parameter groups. Empirical microstates were similar to simulated microstates in the two parameter groups. Thus, fast synchronization transitions correlated with slow BOLD fluctuation based on structural connectivity yielded

OPEN ACCESS

Edited by:
Sacha Jennifer van Albada, Julich Research Centre, Germany

Reviewed by:
Stavros I. Dimitriadis, Cardiff University, United Kingdom
Jamie Sleigh, The University of Auckland, New Zealand

***Correspondence:**
Okito Yamashita
oyamashi@atr.jp

Received: *01 October 2019*
Accepted: *20 December 2019*
Published: *17 January 2020*

Citation:
Endo H, Hiroe N and Yamashita O (2020) Evaluation of Resting Spatio-Temporal Dynamics of a Neural Mass Model Using Resting fMRI Connectivity and EEG Microstates. Front. Comput. Neurosci. 13:91. doi: 10.3389/fncom.2019.00091

characteristics of microstates. Our results demonstrate that a bottom-up approach, which extends the single neuronal dynamics model based on empirical observations into a neural mass dynamics model and integrates structural connectivity, effectively reveals both macroscopic fast, and slow resting-state network dynamics.

Keywords: resting-state networks, resting-state functional connectivity, microstates, neural mass model, cortico-cortical dynamics

INTRODUCTION

Research on resting-state networks is attracting much attention in human neuroimaging. Resting-state functional connectivity (rsFC), i.e., co-activation patterns of slowly fluctuating BOLD signals measured with fMRI (on the order of seconds), has shown interesting empirical evidence on functional subnetworks and their relevance to individual differences (Smith et al., 2013). On the other hand, the microstates, i.e., fast-transient spatial patterns of human scalp potential measured with EEG (on the order of 10–100 ms), have been regarded as the building blocks of human information processing, and four canonical microstates appear in resting-state consistently across subjects and studies (Pascual-Marqui et al., 1995; Koenig et al., 2002; Michel and Koenig, 2018a). In addition, simultaneous fMRI and EEG measurements have been used to reveal the relationship between the slow fluctuation related to rsFC and the fast synchronization transition related to microstates in terms of the spatiotemporal dynamics of the human brain's information processing (Britz et al., 2010; Van de Ville et al., 2010; Yuan et al., 2012; Schwab et al., 2015; Bréchet et al., 2019). However, few mechanistic explanations of these two phenomena have been presented.

Recently, the connectome dynamics models, based on models of neural dynamics constrained by the whole brain's structural connectivity (called connectome), have been investigated to clarify the generative mechanism of functional brain activities and networks. Several computational studies have used simulated neuronal dynamics to understand the mechanistic origins of rsFC patterns (Breakspear et al., 2007; Honey et al., 2009; Deco and Jirsa, 2012; Deco et al., 2013), dynamic rsFC patterns (Hansen et al., 2015; Fukushima and Sporns, 2018), and static FC related to fast synchronization measured by MEG (Nakagawa et al., 2014; Deco et al., 2017; Abeysuriya et al., 2018). Furthermore, recent studies have tried to uncover the relationships between fast synchronization transition and slow fluctuation by combining experimental fMRI with EEG or/and MEG data. Schirner et al. proposed a connectome dynamics model that has EEG source currents in the alpha band as input and demonstrated that the model replicated multiple experimental observations measured with fMRI (Schirner et al., 2018). Demirtaş et al. proposed a locally heterogeneous connectome dynamics model that improved the replication performance of rsFC and MEG power spectrum spatial distribution (Demirtaş et al., 2019). Roberts et al. showed that the Larter-Breakspear model (Sanz-Leon et al., 2015) constrained by the connectome generated rich repertoires of rapidly changing spatiotemporal patterns that are in agreement with the temporal statistics of experimental data such as electrical waves in cortical tissue, sequential spatiotemporal patterns in the resting state MEG data, and large-scale waves in human electrocorticography as well as static rsFC (Roberts et al., 2019). However, similarities between experimental and simulated fast-transient spatial patterns have not yet been investigated.

In this study, to evaluate experimental evidence of resting-state activity on two different time scales but in a unified way, we investigated a connectome dynamics model that explains both experimental rsFC and microstates. We used the Larter-Breakspear model, in which the inhibitory and excitatory neurons in one region are connected with those in other regions based on a connectome measured with diffusion MRI. We optimized the global coupling strength and the local gain parameter (variance of the excitatory and inhibitory threshold) of the simulated neuronal dynamics by fitting both rsFC and microstate spatial patterns to those of the experimental ones. As a result, we found that fast synchronization transitions correlated with slow BOLD fluctuation based on structural connectivity yielded characteristics of empirical microstates. In detail, we found that the parameter sets with high fitting performance to rsFC overlapped with those with high fitting performance to microstates and that the optimal parameter range was greatly reduced by adding microstates as evaluation criteria compared with not adding them as in a previous work (Honey et al., 2009). We found two parameter regions where both rsFC and microstate spatial patterns were reproduced with moderately high accuracy: One had a high local gain (high variance of the excitatory and inhibitory threshold) and weak global coupling strength, while the other had a low local gain (low variance of the excitatory and inhibitory threshold) and strong global coupling strength. In investigating the neural mass dynamics generated from these two parameter sets, the former showed highly periodic and synchronized activation; the latter showed fewer synchronized and periodic activations. The temporal transition of the simulated microstates for the former parameters persisted for about 200 ms, and that for the latter parameters persisted for 150 ms. Both resulted in longer durations than the experimental data. Our results demonstrate that a bottom-up approach, which extends microscopic models of single-neuron dynamics based on empirical studies (Hodgkin and Huxley, 1952; Morris and Lecar, 1981) into a mesoscopic neural mass dynamics model (Larter et al., 1999) and integrates macroscopic structural connectivity, can effectively reveal both macroscopic fast and slow resting-state network dynamics that are observed in human neuroimaging measurements.

MATERIALS AND METHODS

RSNs are characterized by an rsFC based on slow fluctuations observed by fMRI and microstates based on fast synchronization transitions observed by EEG. Simulated rsFCs and microstates, which were obtained by the Larter-Breakspear model that integrated the empirical structural connectivity, were compared with the empirical rsFCs and the microstates. Regarding slow fluctuation, the mean excitatory membrane potentials were converted into blood-oxygen-level dependent (BOLD) signals by the Balloon-Windkessel model. Next, after global fluctuations were regressed out of the BOLD signals, simulated rsFCs were obtained by calculating the cross-correlation coefficients among the BOLD signals. The empirical and simulated rsFCs were evaluated for their spatial pattern similarity. Regarding the fast synchronization transitions, a simulated EEG was obtained by multiplying the lead field and transformed into microstates by applying modified k-means clustering. We evaluated the empirical and simulated microstates for their spatial pattern similarity and non-stationary switching of microstates (**Figure 1**).

Data Acquisition

Structural and Diffusion MRI Data

To obtain a structural connectivity matrix using a fiber-tracking algorithm, we measured the T1-weighted structural (TR: 2,300 ms, TE: 2.98 ms, Flip angle: 9°, TI: 900 ms, thickness: 1 mm, FOV: 256, matrix: 256 × 256, iso-voxel) and diffusion MRI data (gradient directions: 64, *b*-value: 1,000, thickness: 2 mm, iso-voxel) that were acquired on a 3T Trio (Siemens, Erlangen, Germany) from 13 participants (11 males and 2 females, aged 28.7 ± 8.47 years). All of the 13 participants gave informed written consent. All of the experiments in this study were conducted according to the Declaration of Helsinki and were approved by the Ethics Committee of the Advanced Telecommunications Research Institute International, Japan.

Resting-State Functional MRI Data

We recorded the resting-state brain activities for 10 min. The same 13 participants who took part in the dMRI experiment fixated on a cross, let their mind wander, and avoided focusing on any one thing. The resting-state functional imaging data (TR: 2,500 ms, TE: 30 ms, FOV: 212 mm, flip angle: 80°, matrix: 64 × 64, thickness: 3.2 mm, gap: 0.8 mm, 40 slices × 244 volumes) were acquired on a 3T Trio (Siemens, Erlangen, Germany).

Resting-State EEG Data

The resting-state EEGs were recorded for 5 min. The four participants who took part in the fMRI experiment fixated on the cross, let their mind wander, and avoided focusing on any one thing. All of the four participants gave informed written consent. Their EEGs were recorded with a whole-head 63-channel system (BrainAmp; Brain Products GmbH, Germany). The sampling frequency was 1 kHz. Electrooculogram (EOG) signals were simultaneously recorded and then stored in the EEG.

Data Analysis

Empirical Structural Connectivity

We computed the experimental structural connectivity matrix in accordance with a previous work (Fukushima et al., 2015). Briefly, the seed and target ROIs used for fiber-tracking were obtained by FreeSurfer. The participants' motions were corrected by the FMRIB Software Library (FSL). Fractional anisotropy images were then calculated from the corrected images and used for registering the diffusion-space to the T1-space by a non-linear registration tool (FNIRT) in FSL. The local model of the fiber orientations was the fiber orientation distribution (FOD), reconstructed at each voxel by constrained spherical deconvolution (Tournier et al., 2007) with six-dimensional spherical harmonics for the response function. Based on the reconstructed FOD, fibers were probabilistically tracked by MRtrix. The fiber tracks were generated 105 times from each ROI. We calculated the structural connectivity strength as the number of fibers within each ROI pair *ft* divided by the total number of fibers generated from seed ROI *fs* with voxel size normalization: $(ft/vt)/(fs/vs)$, where *vt* and *vs* are the number of voxels in the target and the seed ROI. Since the direction of the structural connectivity strength was not determined by a measurement principle, the structural connectivity matrix was symmetrized by assigning the higher strength to both directions. All parameters were determined as done in a previous work (Fukushima et al., 2015). The representative structural connectivity matrix was obtained by averaging the structural connectivity matrices of all participants with respect to the participants' common cortical ROI (**Table S1**).

Empirical Resting-State Functional Connectivity

The resting-state functional imaging was preprocessed with SPM8 software (Wellcome Trust Center for Neuroimaging, University College London, UK) in MATLAB (R2013a, Mathworks, USA) as follows. First, the raw functional images were corrected for slice-timing and realigned to the mean image of that sequence to compensate for the head motion. Second, the structural images were co-registered to the mean functional image and segmented into three tissue classes in the Montreal Neurological Institute (MNI) space. The functional images were then normalized and resampled in a 2 × 2 × 2 mm grid and smoothed by a Gaussian of 8 *mm* full-width at half-maximum.

We computed the functional connectivity matrix using parcellation defined by anatomical automatic labeling (AAL) for each participant. We extracted a representative time course in each region by averaging the time courses of the voxels therein. A band-pass filter (transmission range, 0.008–0.1 Hz) was applied to these sets of time courses prior to the following regression procedure. The filtered time courses were linearly regressed by the temporal fluctuations of the white matter, the cerebrospinal fluid, and the entire brain. Here, the fluctuation in each tissue class was determined from the average time course of the voxels within a mask created by the segmentation procedure of the T1 image. These extracted time courses were bandpass-filtered (transmission range, 0.008–0.1 Hz) before the linear regression, as was done for the regional time courses. All parameters were determined as done in a previous work (Yahata et al., 2016).

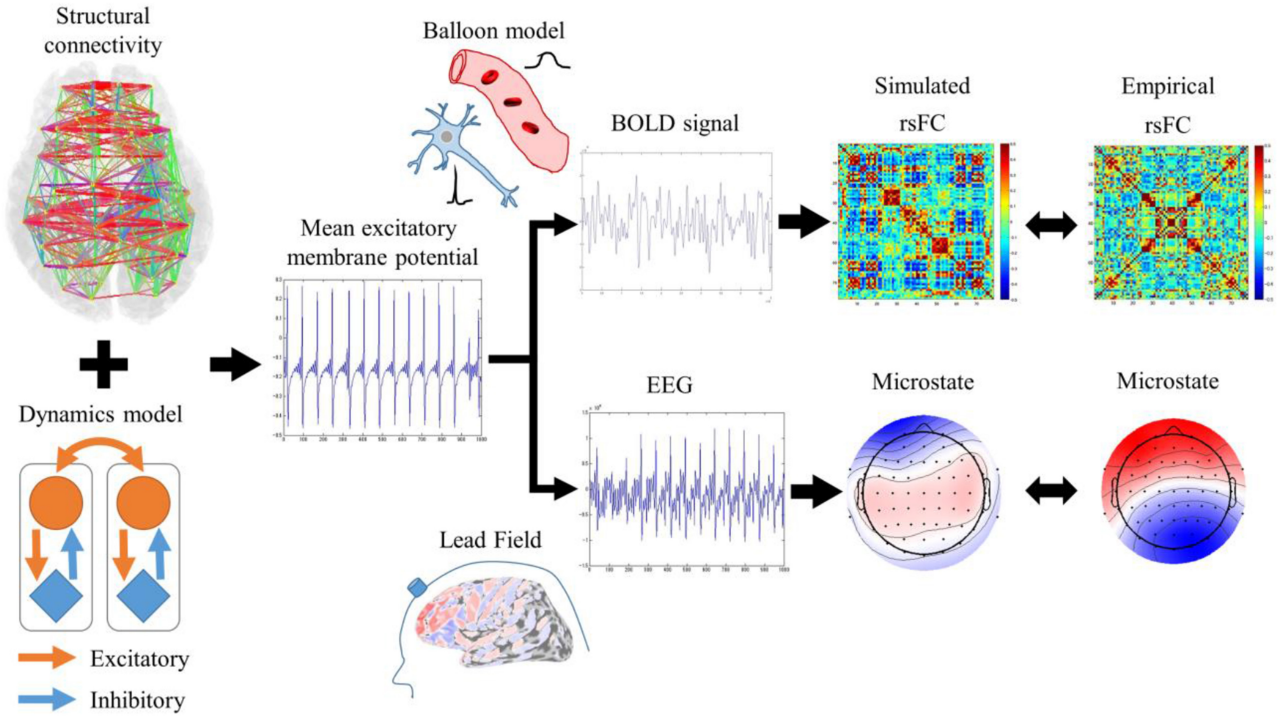

FIGURE 1 | Evaluation procedures for rsFC and microstates. Mean excitatory membrane potentials obtained by Larter-Breakspear model integrating structural connectivity were converted into simulated rsFC and microstates by Balloon-Windkessel model and lead field and compared with empirical rsFC and microstates.

Then a representative rsFC was obtained by calculating the cross-correlation coefficients among the BOLD signals for each participant and averaging the rsFC of each participant with respect to the common cortical ROI in all participants.

Microstates

The EEG data were preprocessed with a low-pass FIR filter with a cutoff frequency of 50 Hz, downsampled at 100 Hz, and passed through a high-pass FIR filter with a cutoff frequency of 0.5 Hz. After a common average reference, the EOG artifacts were removed by generating a multiple linear regression model to predict the eye-movement-related components in the EEG data using the EOG data. Cardiac artifacts and sensor noise were removed by ICA. All of the EEG data were converted into empirical microstates.

Since microstates are seen as building blocks of human information processing as noted in a previous work (Koenig et al., 2002), they were introduced as the criteria to compare the fast dynamics between empirical and simulated neuronal dynamics. First, the standard deviation of all EEG signals, called the global field power (GFP), was calculated as the criterion of the signal-to-noise ratio (SNR). Second, four microstates were identified by the basic N-microstate algorithm applied to the normalized EEG at the local maxima in the GFP curve. The optimal number of microstates was set to four based on a large-scale study on microstates (Koenig et al., 2002), and probabilistically discrete initial values for clustering were chosen based on the k-means++ algorithm (Arthur and Vassilvitskii, 2007). If the spatial correlation value between one microstate and another microstate was 0.9 or more, these two microstates were merged. Third, the transition of the microstates was calculated by a segmentation-smoothing algorithm. All parameters were determined as done in a previous work (Pascual-Marqui et al., 1995). That is, convergence criterion parameter $\epsilon = 10^{-6}$, window size parameter $b = 3$, and non-smoothness penalty parameter $\lambda = 5$. The occupation ratio is defined as the time allocated to a microstate divided by the total time.

Computational Modeling

Larter-Breakspear Model

The Larter-Breakspear model is a phenomenological scheme that describes the electrophysiological neuronal dynamics in each region based on structural connectivity. This model consists of the mean membrane potential of the excitatory neurons (V) and the inhibitory neurons (Z), and the average number of open potassium ion channels (W). The mean firing rate for excitatory and inhibitory populations are described by Q_V and Q_Z. The voltage-dependent fractions of open ion channels are described by m_{ion}. These sigmoidal functions describe averaging over a population of ion channels and cell firing rates under Gaussian distribution. Excitatory interactions between region i and j are described by $\langle Q_V \rangle^i$. Simulations were performed using ode23, which automatically chooses the step size, maintains a specified accuracy, and solves ordinary differential equations in MATLAB. We repeated simulations for each parameter set 10 times with

TABLE 1 | Parameter values for the Larter-Breakspear model.

Parameter	Description	Value
T_{Ca}	Threshold value for Ca channels	−0.01
δ_{Ca}	Variance of Ca channel threshold	0.15
g_{Ca}	Conductance of population of Ca channels	1
V_{Ca}	Ca Nernst potential	1
T_K	Threshold value for K channels	0.0
δ_K	Variance of K channel threshold	0.30
g_K	Conductance of population of K channels	2.0
V_K	K Nernst potential	−0.7
T_{Na}	Threshold value for Na channels	0.3
δ_{Na}	Variance of Na channel threshold	0.15
g_{Na}	Conductance of population of Na channels	6.7
V_{Na}	Na Nernst potential	0.53
V_L	Nernst potential leak channels	−0.5
g_L	Conductance of population of leak channels	0.5
V_T	Threshold potential for excitatory neurons	0.0
Z_T	Threshold potential for inhibitory neurons	0.0
δ_Z	Variance of inhibitory threshold	Same value as δ_V
$Q_{V_{max}}$	Maximal firing rate for excitatory populations	1.0
Q_{Zmin}	Maximal firing rate for inhibitory populations	1.0
I	Subcortical input strength	0.30
a_{ee}	Excitatory-to-excitatory synaptic strength	0.36
a_{ei}	Excitatory-to-inhibitory synaptic strength	2
a_{ie}	Inhibitory-to-excitatory synaptic strength	2
a_{ne}	Non-specific-to-excitatory synaptic strength	1
a_{ni}	Non-specific-to-inhibitory synaptic strength	0.4
b	Time constant scaling factor	0.1
φ	Temperature scaling factor	0.7
τ_K	Time constant for K relaxation time	1
r_{NMDA}	Ratio of NMDA to AMPA receptors	0.25
δ	Random modulation of subcortical input	0

different initial values to reduce the influence of initial values. The simulation length was set to 10 min. We discarded the first 2 min to eliminate the influence of the initial values. All simulation parameters in **Table 1** were determined based on a previous work (Roberts et al., 2019). In the parameter search, the balance of intra- and inter-regional excitatory synaptic connection strength was changed by the global coupling strength C, and the oscillation of the excitatory and inhibitory neural populations was changed by the variance of the excitatory and inhibitory threshold. It took about 2 days to complete 10 min of simulation for a particular parameter set and an initial value using our high performance computer server.

$$m_{ion} = 0.5\left(1 + tanh\left(\frac{V^i - T_{ion}}{\delta_{ion}}\right)\right), \tag{1}$$

$$Q_V = 0.5Q_{V_{max}}\left(1 + tanh\left(\frac{V^i - V_T}{\delta_V}\right)\right), \tag{2}$$

$$Q_Z = 0.5Q_{Z_{max}}\left(1 + tanh\left(\frac{Z^i - Z_T}{\delta_Z}\right)\right), \tag{3}$$

$$\begin{aligned}\frac{dV^i}{dt} = &-\left(g_{Ca} + (1-C)\, r_{NMDA} a_{ee} Q_V^i\right.\\ &\left.+ Cr_{NMDA} a_{ee} \langle Q_V\rangle^i\right) m_{Ca}\left(V^i - V_{Ca}\right)\\ &- g_K W\left(V^i - V_K\right) - g_L\left(V^i - V_L\right)\\ &-\left(g_{Na} m_{Na} + (1-C)\, a_{ee} Q_V^i + Ca_{ee}\langle Q_V\rangle^i\right)\left(V^i - V_{Na}\right)\\ &- a_{ie} Z Q_Z^i + a_{ne} I, \end{aligned} \tag{4}$$

$$\frac{dZ^i}{dt} = b\left(a_{ni} I + a_{ei} V^i Q_V^i\right), \tag{5}$$

$$\frac{dW^i}{dt} = \phi\frac{m_K - W^i}{\tau_K} \tag{6}$$

$$\langle Q_V\rangle^i = \frac{\sum_j u_{ij} Q_V^j}{\sum_j u_{ij}}. \tag{7}$$

Here, u_{ij} is the structural connectivity strength between region i and region j.

Simulated Bold

To calculate the simulated rsFC, the mean membrane potential of the excitatory neurons was converted into simulated BOLD signals by the Balloon-Windkessel hemodynamic model. In this paper, neuronal activity was given by the absolute value of the time derivative of the mean excitatory membrane potential within each brain region. For the ith region, neuronal activity z_i increased vasodilatory signal s_i, which is subject to autoregulatory feedback. Inflow f_i responds in proportion to this signal with concomitant changes in blood volume v_i and deoxyhemoglobin content q_i. The following are the related equations:

$$\frac{ds_i}{dt} = z_i - \kappa_i s_i - \gamma_i\left(f_i - 1\right), \tag{8}$$

$$\frac{df_i}{dt} = s_i, \tag{9}$$

$$\tau_i\frac{dv_i}{dt} = f_i - v_i^{1/\alpha}, \tag{10}$$

$$\tau_i\frac{dq_i}{dt} = \frac{f_i(1-(1-\rho_i^{1/f_i}))}{\rho_i} - \frac{v_i^{1/\alpha} q_i}{v_i}, \tag{11}$$

where $\kappa_i = 0.65$ is the rate of signal decay, $\alpha = 0.32$ is Grubb's exponent, $\tau_i = 0.98$ is the hemodynamic transit time, and $\rho = 0.34$ is the resting oxygen extraction fraction. The BOLD signal is a static non-linear function of volume and deoxyhemoglobin comprised of a volume-weighted sum of the extra- and intravascular signals:

$$y_i = V_0\left(7\rho_i\left(1 - q_i\right) + 2\left(1 - \frac{q_i}{v_i}\right) + (2\rho_i - 0.2)(1 - v_i)\right), \tag{12}$$

where $V_0 = 0.02$ is the resting-blood volume fraction. All of the simulation parameters were determined as done in a previous work (Friston et al., 2003).

Lead Field

The mean membrane potentials of the excitatory neurons were converted into simulated EEGs by the lead field that

represents the linear relation between the sources on the cortical surface and the measurements on each EEG channel as a gain matrix. Polygon models of the cortical surfaces (20,004 vertex points) were constructed from T1 structural images of the same four participants who took part in the EEG experiment using FreeSurfer software (Dale et al., 1999). A single current dipole was assumed at each vertex point to be perpendicular to the cortical surface. The brain structures were approximated as a three-layer model by identifying three boundaries, i.e., for the cerebrospinal fluid (CSF), the skull, and the scalp, assuming that the conductivities were 0.33, 0.0042, and 0.33, respectively (Waberski et al., 1998). These surfaces obtained by FreeSurfer were slightly modified using gray/white/CSF segmentation by SPM8 (Welcome Department of Cognitive Neurology, UK) and morphological operations. All of the parameters were determined as done in a previous work (Aihara et al., 2012).

Singular value decomposition was applied to all points belonging to one ROI out of 20,004 points. The singular values were arranged in descending order and the lead field at 80% accumulation rate of singular values was averaged in one ROI. (That is, it has almost the same effect as averaging the lead field within an ROI). We calculated the low-dimensional lead field (63 channels $\times$ 78 ROIs) of each participant by averaging the components that contribute 80% for each ROI since the Later-Breakspear model has the same number of neurons in each ROI. This way, we removed the influence of heterogeneous ROI sizes in the empirical data. Then the low-dimensional lead field matrices of all subjects were averaged.

Simulated Microstates

To compare the empirical and simulated microstates, we converted the mean membrane potential of the excitatory neurons into simulated EEG signals by the lead field. First, the simulated EEG signals were calculated by multiplying the mean membrane potential of the excitatory neurons by the lead field. Second, a common average reference was applied for simulated EEG signals and the simulated EEG signals were bandpass-filtered between 10 and 15 Hz. Third, the bandpass-filtered signals were downsampled from 1,000 to 100 Hz. Finally, the simulated microstates were acquired by applying the basic N-microstate algorithm and the segmentation smoothing algorithm to the obtained signals.

Quantitative Evaluation of Empirical and Simulated Results

The spatial correlation between the empirical and simulated rsFCs was obtained by averaging the cross-correlation coefficients of the lower triangular components between the empirical and simulated rsFCs for each parameter set.

We evaluated the empirical and simulated microstates in terms of spatial similarity, occupation ratio, mean transition time, and global explained variance (GEV). GEV represents the goodness of fit between the microstates and the normalized EEG weighted by GFP (Khanna et al., 2014). We obtained the spatial correlation between the empirical and simulated microstates by the following procedures. For the 24 combinations between the empirical and simulated microstates, we calculated the average cross-correlation coefficients of each combination and chose the one that maximized it. In the simulation, two microstates that showed similar spatial patterns (similarity over 0.9) were merged. This happened for several parameter combinations. If the number of simulated microstates was >4, the cross-correlation coefficients were set to 0 for the missing simulated microstates. The occupation ratio was obtained by dividing the time allocated to each microstate by the total simulation time. The mean transition time was obtained by averaging required time to transition from one microstate to another.

Phase-Locking Matrix

The phase-locking matrix quantitatively visualizes the inter-regional interdependence of the non-linear simulated neuronal dynamics. First, the 10–15 Hz bandpass filter was applied to the mean membrane potentials. Second, these signals were Hilbert-transformed, extracting the phase θ. Finally, we calculated the phase-locking values (PLVs) between regions p and q by the following formula:

$$\mathrm{PLVs} = \frac{\left| \sum_{t=1}^{N_T} e^{i(\theta_q^t - \theta_p^t)} \right|}{N_T}, \tag{13}$$

where N_T represents the sampling number.

RESULTS

Spatial Pattern Similarity of rsFC and Microstates

In the two parameter groups of strong global coupling strength and small variance of threshold or weak global coupling strength and large variance of threshold, the simulated rsFC and microstates indicated a high spatial similarity to the empirical rsFCs and microstates (**Figure 2**). We selected the optimal parameter groups by considering both the spatial similarities of rsFC, which were moderately high over a broad range, and the spatial similarities of the microstates, which were high in a narrow range. The spatial similarities in the optimal parameter groups obtained with structural connectivity based on diffusion MRI were significantly higher than those obtained with shuffled structural connectivity (**Figure S1**).

The spatial pattern similarity between empirical and simulated microstates in the $(\mathrm{C},\ \delta_{\mathrm{V,\,Z}}) = (0.35,\ 0.61)$ condition is high but the occupation ratio and the mean transition time are greatly different from empirical microstates (**Figure 2B**, **Figure S2**). The occupation ratio and the mean transition time in the $(\mathrm{C},\ \delta_{\mathrm{V,\,Z}}) = (0.30,\ 0.61)$ and $(0.25,\ 0.61)$ are almost the same as in the $(\mathrm{C},\ \delta_{\mathrm{V,\,Z}}) = (0.35,\ 0.61)$ condition.

Comparison of Spatiotemporal Patterns Between Empirical and Simulated Microstates

Empirical and simulated microstates were obtained by applying the basic N-microstate algorithm and the segmentation smoothing algorithm for EEG time-series. Roughly, one microstate probabilistically transitions to another microstate when GFP is at a local minimum (**Figure S3**).

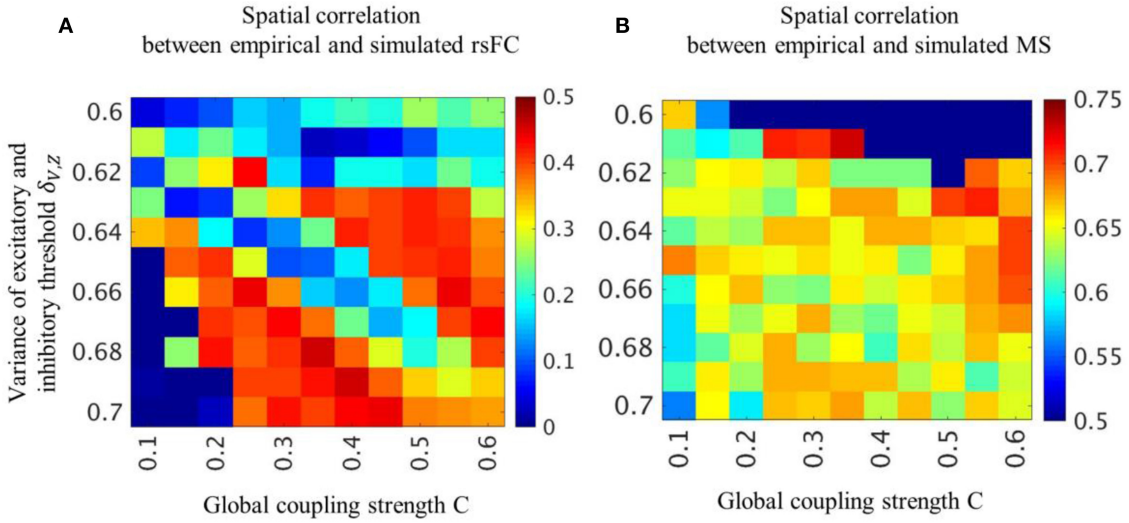

FIGURE 2 | Spatial pattern similarity of rsFC and microstates for each parameter combination. Averaged cross-correlation coefficients between empirical and simulated rsFC **(A)** and microstates **(B)** are obtained by varying global coupling strength *C* and variance of threshold $\delta_{V,Z}$. Color bars indicate strength of cross-correlation coefficients.

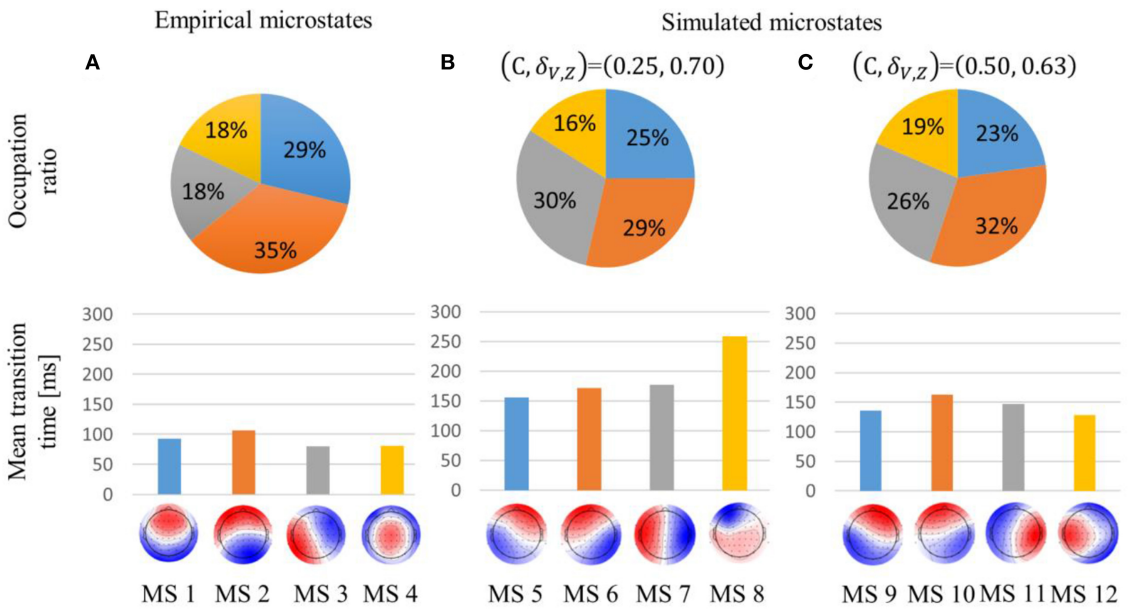

FIGURE 3 | Occupation ratio and mean transition time of empirical **(A)** and simulated **(B,C)** microstates. global explained variance (GEV) was 59, 66, and 69% for the empirical, simulated weak global coupling strength, and simulated strong global coupling strength microstates, respectively. Each simulated microstate is indicated with the same color as the empirical microstate with which it is most highly correlated. Interchanging red and blue would correspond to the identical pattern.

Concerning the empirical microstates, MS2 accounted for the highest proportion, and their mean transition times were about 100 ms, as in the previous research (**Figure 3A**).

In the $(C, \delta_{V,Z}) = (0.25, 0.70)$ condition, the simulated microstates were sustained about twice as long compared with the empirical microstates. In contrast to the empirical microstates, MS6 and MS7 accounted for a higher proportion than MS5. Furthermore, MS8's temporal transition was also intermediate between MS2 and MS4 because MS8 accounted for the lowest proportion, even though it was sustained for the longest time (**Figure 3B**).

In the $(C, \delta_{V,Z}) = (0.50, 0.63)$ condition, the simulated microstates were sustained about 1.5 times longer than the empirical microstates. Unlike the $(C, \delta_{V,Z}) = (0.25, 0.70)$ condition, the occupation ratio of MS10 was about 5% higher than that of MS11.

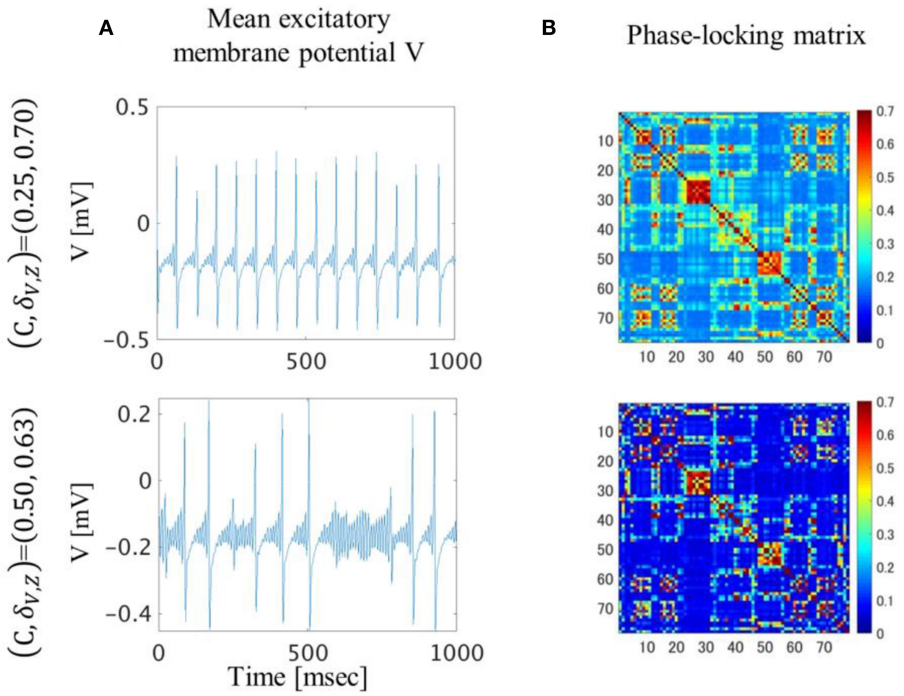

FIGURE 4 | Excitatory mean membrane potentials for one second **(A)** and averaged phase-locking matrix **(B)**. In the $(C, \delta_{V,Z}) = (0.25, 0.70)$ condition, excitatory mean membrane potentials fluctuated in a regular manner and were synchronized across the whole brain region. In the $(C, \delta_{V,Z}) = (0.50, 0.63)$ condition, excitatory mean membrane potentials fluctuated irregularly and were synchronized between brain regions with strong structural connectivity.

Moreover, the spatial pattern of MS12 was biased to the left hemisphere (**Figure 3C**).

Comparison of Neuronal Dynamics and Phase-Locking

In the $(C, \delta_{V,Z}) = (0.25, 0.70)$ condition, the excitatory mean membrane potentials were synchronized across the whole brain region with weak global coupling strength because the self-recurrent excitation was higher than the low variance of the threshold (**Figures 4A**,**B** top). In the $(C, \delta_{V,Z}) = (0.50, 0.63)$ condition, the excitatory mean membrane potentials were synchronized between the brain regions with strong structural connectivity and were not synchronized between the brain regions with weak or no structural connectivity (**Figure 4B** bottom, **Figure S1A**). Strong global coupling strength and structural connectivity are required for excitatory mean membrane potentials to exceed the excitatory threshold due to weak self-recurrenct excitation. Due to the synchronized excitatory mean membrane potentials across the whole brain region, the mean transition time in the $(C, \delta_{V,Z}) = (0.25, 0.70)$ condition is 1.3 times longer than the mean transition time in the $(C, \delta_{V,Z}) = (0.50, 0.63)$ condition (**Figure 3**).

DISCUSSION

Summary

In this study, we investigated whether there is a model that simultaneously explains two experimentally observed phenomena in the resting state: slow fluctuation manifested by resting-state functional connectivity (rsFC) and fast transient dynamics manifested by EEG microstates. We simulated a neural mass model using the Larter-Breakspear model constrained by the structural connectivity and optimized the model parameters [the global coupling parameter and the local gain parameter (variance of excitatory and inhibitory threshold)] by fitting the simulated rsFC to the experimental rsFC and the simulated microstates to the experimental microstates. As a result, we obtained three key findings: first, the parameter sets with high fitting performance to rsFC overlapped with those with high fitting performance to the microstate; second, two distinct parameter sets were identified within the overlapped parameter region; third, the overlapped parameter sets are much narrower than the parameter sets obtained by fitting only rsFC. In other words, based on these three findings, both fast synchronization transitions and slow BOLD fluctuation changed based on structural connectivity in the overlapped parameter regions. Empirical microstates were similar to simulated microstates in these regions. Thus, fast synchronization transitions correlated with slow BOLD fluctuation based on structural connectivity yielded microstates. These results suggest that adding the microstates as fitting criteria is important for significantly reducing the uncertainty of good model parameters. The maximal cross-correlation coefficient of the rsFC was about 0.45, which is almost the same value as that in previous research (Honey et al., 2009). The maximal cross-correlation coefficient of the microstates was about 0.7, and the mean transition times

of the simulated microstates were about 1.5 or 2.0 times longer than those of the empirical microstates.

Comparison With Previous Research

Since the structural connectivity in this study had stronger interhemispheric connectivity (**Figure S1A**) than in the previous research, the cross-correlation coefficients of rsFC were about 0.05 higher than in the previous research (Honey et al., 2009; Deco et al., 2013).

In the $(C, \delta_{V,Z}) = (0.50, 0.63)$ condition, the dynamics of the excitatory mean membrane potentials resembled the dynamics of previous research in terms of irregular firing (Roberts et al., 2019). The simulated microstates using shuffled structural connectivity did not reproduce either a spatial pattern or a temporal transition (**Figure S4**). In the $(C, \delta_{V,Z}) = (0.25, 0.70)$ condition, the dynamics of the excitatory mean membrane potentials resembled the dynamics of the previous research in terms of regular firing (Honey et al., 2009). Simulated microstates using shuffled structural connectivity reproduced a spatial pattern but not a temporal transition (**Figures S4B**, **S5**). Furthermore, the transition probability matrix of empirical microstates is more similar to the simulated one in the $(C, \delta_{V,Z}) = (0.50, 0.63)$ condition than in the $(C, \delta_{V,Z}) = (0.25, 0.70)$ condition (**Figure S6**). Therefore, the dynamics of the excitatory mean membrane potentials in the $(C, \delta_{V,Z}) = (0.50, 0.63)$ condition resembled the empirical resting-state activity of the human brain. The dynamics of the excitatory mean membrane potentials in the $(C, \delta_{V,Z}) = (0.25, 0.70)$ condition were comparatively unconstrained by the structural connectivity (**Figure 4B**, top). Although these were bandpass-filtered dynamics, perhaps they are related to the phenomena in which the activity spreads to most of the cortex when the cortex is stimulated with transcranial magnetic stimulation (TMS) during sleep (Alkire et al., 2008).

Simulated Microstates

The MS7 and MS11 in each simulated case have a larger occupation ratio than the MS3 in the experiment (**Figure 3**) because the structural connectivity strength between the occipital regions is weak (**Figure S1A**). We assume that the interhemispheric structural connectivity was important for reproducing MS4, considering that MS4 has an equal distribution for the distant interhemispheric positions. Thus, the constraints based on structural connectivity are more important than the signal noise and the conduction delay proportional to the inter-regional distance. Signal noise probably affects temporal transition and the conduction delay proportional to the inter-regional distance probably results in ROIs close to each other being synchronized.

Clinical Application for Cognitive Neuroscience

The development of biomarkers for psychiatric disorders using microstates and rsFC has been investigated (Yamada et al., 2017; D'Croz-Baron et al., 2019). One example is autism spectrum disorder (ASD). ASD patient's rsFC and microstates have different spatial and temporal characteristics compared to healthy individuals. Our results could not reproduce enough spatiotemporal characteristics of rsFC and microstates to compare ASD patients and healthy individuals. However, our results will give hints to solve these problems and enable making hypotheses about the dependence of slow and fast resting-state brain activity on neuronal network parameters.

Effects of Averaging Lead Field

In the Larter-Breakspear model, the numbers of excitatory, and inhibitory neurons are the same in each ROI, and only the structural connectivity differs between the ROIs. For this reason, the high-dimensional lead field was averaged in each ROI to remove the area's influence in each ROI. If the ROIs were divided into smaller sections and the lead field was not averaged for each ROI, the simulated microstates might differ. We note that the contribution rate of the singular value decomposition did not affect the simulated microstates because the cross-correlation coefficient was 0.9995 between the lead fields by singular value decomposition and the mean value.

Limitation of Current Simulation

The occupation ratio of the empirical and simulated microstates differed and the simulated microstate did not reproduce the specific spatial patterns of the empirical microstates. The resting-state alpha waves in the cerebral cortex are empirically affected by the thalamus (Sherman, 2016), and slow fluctuation is affected by the serotonin receptor density (Deco et al., 2018; Shine et al., 2019). However, the Larter-Breakspear model in this study does not include thalamic dynamics, conduction delay proportional to inter-regional distance, signal noise, or serotonin receptor density. Therefore, incorporating these effects in this model might more fully reveal how well the simulated rsFC and microstates reflect the empirical rsFC and microstates in realistic conditions.

Numbers of Participants

Since empirical EEG and rsFC were obtained by two studies aiming at different purposes, the number of participants of EEG and rsFC differ. Empirical rsFC was obtained from 13 participants, sufficient for statistical purposes. Empirical EEG was obtained from only four participants. However, the four microstates identified from this empirical EEG are similar to the four microstates identified by other studies (Michel and Koenig, 2018b). We presume the four microstates will be similar to the microstates identified from 13 participants.

DATA AVAILABILITY STATEMENT

Relevant data and all codes used for the analyses are available from the authors on request.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Ethics Committee of the Advanced Telecommunication Research Institute International, Japan. The

patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

HE and OY: conceptualization, empirical data analysis, simulation programs, and simulation data analysis. OY: funding acquisition. NH: empirical data acquisition. HE, NH, and OY: writing the paper.

FUNDING

This research was supported by contract research with the National Institute of Information and Communications Technology entitled Development of network dynamics modeling methods for human brain data simulation systems (grant #173) and Analysis of multi-modal brain measurement data and development of its application for BMI open innovation (grant #209). OY was partially supported by the contract research Grant Number JP19dm0307009 supported by the Japan Agency for Medical Research and Development (AMED). The funders had no role in the study design, data collection and analysis, decision to publish, or preparation of the manuscript.

ACKNOWLEDGMENTS

We gratefully acknowledge the work of past and present members of our laboratory for helping us to analyze empirical and simulation data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncom.2019.00091/full#supplementary-material

REFERENCES

Abeysuriya, R. G., Hadida, J., Sotiropoulos, S. N., Jbabdi, S., Becker, R., Hunt, B. A. E., et al. (2018). A biophysical model of dynamic balancing of excitation and inhibition in fast oscillatory large-scale networks. *PLoS Comput. Biol.* 14:e1006007. doi: 10.1371/journal.pcbi.1006007

Aihara, T., Takeda, Y., Takeda, K., Yasuda, W., Sato, T., Otaka, Y., et al. (2012). Cortical current source estimation from electroencephalography in combination with near-infrared spectroscopy as a hierarchical prior. *Neuroimage* 59, 4006–4021. doi: 10.1016/j.neuroimage.2011.09.087

Alkire, M. T., Hudetz, A. G., and Tononi, G. (2008). Consciousness and anesthesia. *Science* 322, 876–880. doi: 10.1126/science.1149213

Arthur, D., and Vassilvitskii, S. (2007). K-means++: the advantages of careful seeding. *Proc. Eighteenth Annu. ACM-SIAM Symp. Discret. algorithms* 8, 1027–1025. doi: 10.1145/1283383.1283494

Breakspear, M., Sporns, O., Honey, C. J., and Ko, R. (2007). Network structure of cerebral cortex shapes functional connectivity on multiple time scales. *PSNA.* 104, 10240–10245. doi: 10.1073/pnas.0701519104

Bréchet, L., Brunet, D., Birot, G., Gruetter, R., Michel, C. M., and Jorge, J. (2019). Capturing the spatiotemporal dynamics of self-generated, task-initiated thoughts with EEG and fMRI. *Neuroimage* 194, 82–92. doi: 10.1016/j.neuroimage.2019.03.029

Britz, J., Van De Ville, D., and Michel, C. M. (2010). BOLD correlates of EEG topography reveal rapid resting-state network dynamics. *Neuroimage* 52, 1162–1170. doi: 10.1016/j.neuroimage.2010.02.052

Dale, A. M., Fischl, B., and Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage* 9, 179–194. doi: 10.1006/nimg.1998.0395

D'Croz-Baron, D. F., Baker, M., Michel, C. M., and Karp, T. (2019). EEG microstates analysis in young adults with autism spectrum disorder during resting-state. *Front. Hum. Neurosci.* 13:173. doi: 10.3389/fnhum.2019.00173

Deco, G., Cabral, J., Woolrich, M. W., Stevner, A. B. A., van Hartevelt, T. J., and Kringelbach, M. L. (2017). Single or multiple frequency generators in on-going brain activity: a mechanistic whole-brain model of empirical MEG data. *Neuroimage* 152, 538–550. doi: 10.1016/j.neuroimage.2017.03.023

Deco, G., Cruzat, J., Cabral, J., Knudsen, G. M., Carhart-Harris, R. L., Whybrow, P. C., et al. (2018). Whole-brain multimodal neuroimaging model using serotonin receptor maps explains non-linear functional effects of LSD. *Curr. Biol.* 28, 3065–3074.e6. doi: 10.1016/j.cub.2018.07.083

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., Ponce-Alvarez, A., Mantini, D., Romani, G. L., Hagmann, P., and Corbetta, M. (2013). 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

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.e13. doi: 10.1016/j.neuron.2019.01.017

Friston, K. J., Harrison, L., and Penny, W. (2003). Dynamic causal modelling. *Neuroimage* 19, 1273–1302. doi: 10.1016/S1053-8119(03)00202-7

Fukushima, M., and Sporns, O. (2018). Comparison of fluctuations in global network topology of modeled and empirical brain functional connectivity. *PLOS Comput. Biol.* 14:e1006497. doi: 10.1371/journal.pcbi.1006497

Fukushima, M., Yamashita, O., Knösche, T. R., and Sato, M. (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

Hansen, E. C. A., Battaglia, D., Spiegler, A., Deco, G., and Jirsa, V. K. (2015). Functional connectivity dynamics: modeling the switching behavior of the resting state. *Neuroimage* 105, 525–535. doi: 10.1016/j.neuroimage.2014.11.001

Hodgkin, A. L., and Huxley, A. F. (1952). The components of membrane conductance in L411mV -D. *J. Physiol.* 116, 473–496. doi: 10.1113/jphysiol.1952.sp004718

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. doi: 10.1073/pnas.0811168106

Khanna, A., Pascual-Leone, A., and Farzan, F. (2014). Reliability of resting-state microstate features in electroencephalography. *PLoS ONE* 9:e114163. doi: 10.1371/journal.pone.0114163

Koenig, T., Prichep, L., Lehmann, D., Sosa, P. V., Braeker, E., Kleinlogel, H., et al. (2002). Millisecond by millisecond, year by year: normative EEG microstates and developmental stages. *Neuroimage* 16, 41–48. doi: 10.1006/nimg.2002.1070

Larter, R., Speelman, B., and Worth, R. M. (1999). A coupled ordinary differential equation lattice model for the simulation of epileptic seizures. *Chaos* 9, 795–804. doi: 10.1063/1.166453

Michel, C. M., and Koenig, T. (2018a). EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. *Neuroimage* 180, 577–593. doi: 10.1016/j.neuroimage.2017.11.062

Michel, C. M., and Koenig, T. (2018b). EEG microstates as a tool for studying the temporal dynamics of whole-brain neuronal networks: a review. *Neuroimage* 180, 577–593.

Morris, C., and Lecar, H. (1981). Voltage oscillations in the barnacle giant muscle fiber. *Biophys. J.* 35, 193–213. doi: 10.1016/S0006-3495(81)84782-0

Nakagawa, T. T., Woolrich, M., Luckhoo, H., Joensson, M., Mohseni, H., Kringelbach, M. L., et al. (2014). How delays matter in an oscillatory whole-brain spiking-neuron network model for MEG alpha-rhythms at rest. *Neuroimage* 87, 383–394. doi: 10.1016/j.neuroimage.2013.11.009

Pascual-Marqui, R. D., Michel, C. M., and Lehmann, D. (1995). Segmentation of brain electrical activity into microstates; model estimation and validation. *IEEE Trans. Biomed. Eng.* 42, 658–665. doi: 10.1109/10.391164

Roberts, J. A., Gollo, L. L., Abeysuriya, R. G., Roberts, G., Mitchell, P. B., Woolrich, M. W., et al. (2019). Metastable brain waves. *Nat. Commun.* 10:1056. doi: 10.1038/s41467-019-08999-0

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. doi: 10.1016/j.neuroimage.2015.01.002

Schirner, M., McIntosh, A. R., Jirsa, V., Deco, G., and Ritter, P. (2018). Inferring multi-scale neural mechanisms with brain network modelling. *Elife* 7:e28927. doi: 10.7554/eLife.28927

Schwab, S., Koenig, T., Morishima, Y., Dierks, T., Federspiel, A., and Jann, K. (2015). Discovering frequency sensitive thalamic nuclei from EEG microstate informed resting state fMRI. *Neuroimage* 118, 368–375. doi: 10.1016/j.neuroimage.2015.06.001

Sherman, S. M. (2016). Thalamus plays a central role in ongoing cortical functioning. *Nat. Neurosci.* 19, 533–541. doi: 10.1038/nn.4269

Shine, J. M., Breakspear, M., Bell, P. T., Ehgoetz Martens, K. A., Shine, R., Koyejo, O., et al. (2019). Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. *Nat. Neurosci.* 22, 289–296. doi: 10.1038/s41593-018-0312-0

Smith, S. M., Vidaurre, D., Beckmann, C. F., Glasser, M. F., Jenkinson, M., Miller, K. L., et al. (2013). Functional connectomics from resting-state fMRI. *Trends Cogn. Sci.* 17, 666–682. doi: 10.1016/j.tics.2013.09.016

Tournier, J. D., Calamante, F., and Connelly, A. (2007). Robust determination of the fibre orientation distribution in diffusion MRI: non-negativity constrained super-resolved spherical deconvolution. *Neuroimage* 35, 1459–1472. doi: 10.1016/j.neuroimage.2007.02.016

Van de Ville, D., Britz, J., and Michel, C. M. (2010). EEG microstate sequences in healthy humans at rest reveal scale-free dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 107, 18179–18184. doi: 10.1073/pnas.1007841107

Waberski, T. D., Buchner, H., Lehnertz, K., Hufnagel, A., Fuchs, M., Beckmann, R., et al. (1998). Properties of advanced headmodelling and source reconstruction for the localization of epileptiform activity. *Brain Topogr.* 10, 283–290. doi: 10.1023/A:1022275024069

Yahata, N., Morimoto, J., Hashimoto, R., Lisi, G., Shibata, K., Kawakubo, Y., et al. (2016). A small number of abnormal brain connections predicts adult autism spectrum disorder. *Nat. Commun.* 7:11254. doi: 10.1038/ncomms11254

Yamada, T., Hashimoto, R. I., Yahata, N., Ichikawa, N., Yoshihara, Y., Okamoto, Y., et al. (2017). Resting-state functional connectivity-based biomarkers and functional mri-based neurofeedback for psychiatric disorders: a challenge for developing theranostic biomarkers. *Int. J. Neuropsychopharmacol.* 20, 769–781. doi: 10.1093/ijnp/pyx059

Yuan, H., Zotev, V., Phillips, R., Drevets, W. C., and Bodurka, J. (2012). Spatiotemporal dynamics of the brain at rest - exploring EEG microstates as electrophysiological signatures of BOLD resting state networks. *Neuroimage* 60, 2062–2072. doi: 10.1016/j.neuroimage.2012.02.031

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.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 23 January 2020
doi: 10.3389/fncom.2019.00100

Effects of Cholinergic Neuromodulation on Thalamocortical Rhythms During NREM Sleep: A Model Study

Qiang Li[1,2], *Jiang-Ling Song*[1,2], *Si-Hui Li*[1], *M. Brandon Westover*[2] *and Rui Zhang*[1*]

[1] *Medical Big Data Research Center, Northwest University, Xi'an, China,* [2] *Department of Neurology, Massachusetts General Hospital, Boston, MA, United States*

It has been suggested that cholinergic neurons shape the oscillatory activity of the thalamocortical (TC) network in behavioral and electrophysiological experiments. However, theoretical modeling demonstrating how cholinergic neuromodulation of thalamocortical rhythms during non-rapid eye movement (NREM) sleep might occur has been lacking. In this paper, we first develop a novel computational model (TC-ACH) by incorporating a cholinergic neuron population (CH) into the classical thalamo-cortical circuitry, where connections between populations are modeled in accordance with existing knowledge. The neurotransmitter acetylcholine (ACH) released by neurons in CH, which is able to change the discharge activity of thalamocortical neurons, is the primary focus of our work. Simulation results with our TC-ACH model reveal that the cholinergic projection activity is a key factor in modulating oscillation patterns in three ways: (1) transitions between different patterns of thalamocortical oscillations are dramatically modulated through diverse projection pathways; (2) the model expresses a stable spindle oscillation state with certain parameter settings for the cholinergic projection from CH to thalamus, and more spindles appear when the strength of cholinergic input from CH to thalamocortical neurons increases; (3) the duration of oscillation patterns during NREM sleep including K-complexes, spindles, and slow oscillations is longer when cholinergic input from CH to thalamocortical neurons becomes stronger. Our modeling results provide insights into the mechanisms by which the sleep state is controlled, and provide a theoretical basis for future experimental and clinical studies.

Keywords: thalamocortical neural mass model, cholinergic projection, acetylcholine (ACH), thalamocortical rhythm, NREM sleep

OPEN ACCESS

Edited by:
Julie S. Haas,
Lehigh University, United States

Reviewed by:
Daniel Llano,
University of Illinois at Urbana-Champaign, United States
Qing Yun Wang,
Beihang University, China

***Correspondence:**
Rui Zhang
rzhang@nwu.edu.cn

Received: *09 September 2019*
Accepted: *30 December 2019*
Published: *23 January 2020*

Citation:
Li Q, Song J-L, Li S-H, Westover MB and Zhang R (2020) Effects of Cholinergic Neuromodulation on Thalamocortical Rhythms During NREM Sleep: A Model Study. Front. Comput. Neurosci. 13:100. doi: 10.3389/fncom.2019.00100

1. INTRODUCTION

Sleep plays a pivotal role in mental and physical health. During sleep, the brain alternates between two stages, rapid-eye movement (REM) and NREM. This alternation is reflected in the electrical rhythms generated by the thalamocortical system, evident in the electroencephalogram (EEG). Specifically, the thalamocortical rhythms in NREM sleep include K-complexes and slow oscillations (dominated by low-frequency [0.5, 2]Hz, high amplitude oscillations), as well as spindles (characterized by a waxing and waning waveform in the range [11, 16]Hz). In contrast,

REM sleep exhibits low amplitude activity of higher frequency, which is similar with rhythms in wakefulness, including the alpha oscillation (Rasch and Born, 2013). Studies reveal that spindles and slow oscillations are very helpful in protecting and stabilizing sleep (Roth, 2009; Dang-Vu et al., 2010; Kim et al., 2012).

A number of results in behavioral and electrophysiological experiments suggest that cholinergic neurons, which are distributed in the basal forebrain and brainstem, help shape the oscillatory activity of the thalamocortical network (Bellingham and Funk, 2000; Kobayashi et al., 2003; Boutrel and Koob, 2004; Brown et al., 2012). Meanwhile, it has also been shown that thalamocortical processing is subject to the action of modulatory of acetylcholine (ACH) released by cholinergic neurons (McCormick, 1992). However, a theoretical framework for understanding the mechanisms of cholinergic neuromodulation on thalamocortical rhythms during NREM sleep is still lacking.

Neural mass models (NMMs), which describe the dynamics of neural populations (Wilson and Cowan, 1973; Lopes da Silva et al., 1974), have shown success in clarifying the mechanisms involved in the generation and regulation of oscillation patterns of the "wakefulness-sleep cycle." For example, Robinson et al. (2002) developed a continuum model of large-scale brain electrical activity to understand the transition from the resting background state to the spike-wave state. Suffczynski et al. (2004) proposed a thalamocortical NMM to explain the relation between mechanisms that generate sleep spindles and those for that generate spike-wave activity. More recently, Weigenand et al. (2014) extended a cortical NMM to explore the mechanisms for generating K-complexes and slow oscillations. Cona et al. (2014) presented a new NMM to describe the sleeping thalamocortical system, where thalamic neurons exhibit two firing modes: bursting and tonic.

Although prior (NMM-based) studies have made progress in explaining sleep rhythms, the role of ACH modulation of thalamocortical activities, especially during sleep has not been thoroughly explored. In this work we model the details of cholinergic neuromodulation in the cortex and thalamus, and study the effects of cholinergic projections on thalamocortical rhythms during NREM sleep.

In this study, a new computational model (TC-ACH) is first proposed by integrating a cholinergic neuron population (CH) into classical thalamo-cortical circuitry, with connections between populations in TC-ACH constructed on the basis of the existing mechanisms. We concentrate on the neurotransmitter ACH released by cholinergic neurons, which impacts firing activity of thalamocortical neurons. The proposed model TC-ACH is used to investigate effects of ACH modulation on thalamocortical rhythms during NREM sleep in two ways: (1) measurement of effects of cholinergic modulation in the thalamus; (2) measurement of effects of cholinergic modulation in the thalamocortical system.

The rest of the paper is organized as follows. In section 2, the model framework and mathematical expression of TC-ACH are introduced. Section 3 presents the simulation results. A summary and discussion of our work are given in the last section.

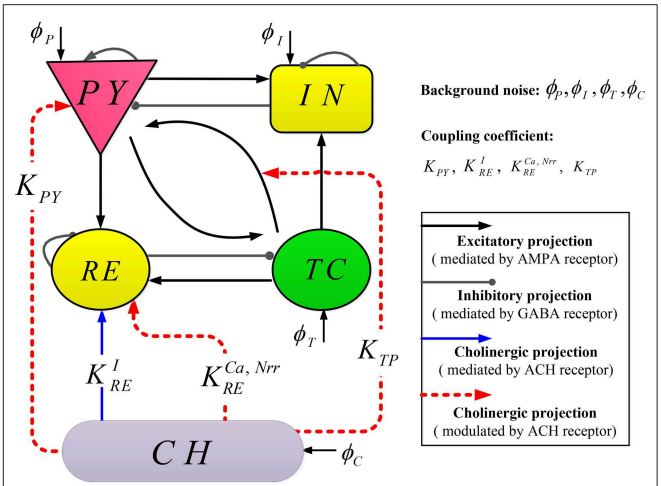

FIGURE 1 | Schematic representation of model TC-ACH.

2. METHODS

2.1. Model Framework

This subsection introduces the topological structure and computational operators of TC-ACH systematically.

2.1.1. Topological Structure

Our computational model TC-ACH is composed of classical thalamo-cortical circuitry as well as a cholinergic neuron population, whose topological structure is illustrated in **Figure 1**.

The thalamo-cortical circuitry consists of two mutually interconnected modules: a cortical module and a thalamic module. Each module comprises two neural populations, where pyramidal cells (PY) and inhibitory interneurons (IN) are included in the cortical module; the thalamic module also includes thalamocortical cells (TC) and thalamic reticular cells (RE). Regarding connections between modules, excitatory neurons (PY or TC) in each module connect to both populations in the other module. Within each module, two populations project to each other, and there exist self-connections within PY, IN, and RE (Costa et al., 2016). Note that in the thalamo-cortical circuitry, the excitatory and inhibitory projections are mediated by AMPA and GABA receptors, respectively.

The cholinergic neuron population (CH) is located in the basal forebrain and brainstem. The neurons in CH release neurotransmitter (ACH) into the synaptic cleft, which changes the activity of thalamocortical neurons. In our proposed model, there exist three modulated cholinergic projections from CH to PY, RE, as well as the connection between TC and PY (represented by red dashed lines in **Figure 1**), and one mediated cholinergic projection from CH to RE (represented by blue the solid line in **Figure 1**). The coupling coefficients of these projections are denoted by K_{PY}, K^{Nrr}_{RE}, K^{Ca}_{RE}, K_{TP}, and K^{I}_{RE}, respectively. Here, all projections involve by ACH.

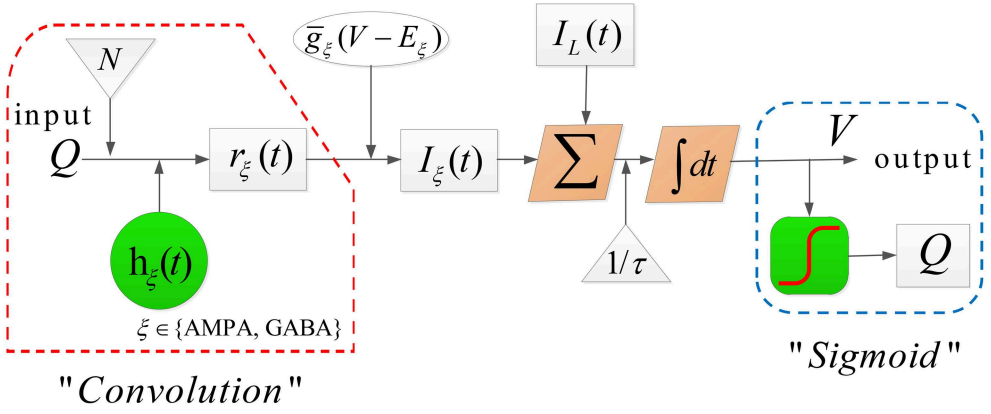

FIGURE 2 | The dynamical evolution in each population in model TC-ACH.

2.1.2. Computational Operators

The dynamical evolution in each population is implemented in two computational blocks.

The first computational block transforms the average membrane potential $V(t)$ into the average firing rate $Q(t)$, formulated by a sigmoid function with the form (Jansen et al., 1993)

$$Q(t) = \frac{Q^{max}}{1 + e^{-(V(t)-\theta)/\sigma}}, \tag{1}$$

where Q^{max}, θ, σ represent the maximal firing rate, the firing threshold and the neural gain, respectively.

In the second computational block, the firing rate $Q(t)$ is first transformed into the fraction of open channels $r_\xi(t)$ by a convolution with an alpha function $h_\xi(t)$, that is,

$$r_\xi(t) = h_\xi(t) \otimes (N \cdot Q(t)), \tag{2}$$

$$h_\xi(t) = \gamma_\xi^2 \cdot te^{-\gamma_\xi t}, t \geq 0. \tag{3}$$

Then the membrane potential $V(t)$ is solved by

$$\begin{aligned} \tau \dot{V}(t) &= -I_L(t) - I_{AMPA}(t) - I_{GABA}(t) \\ &= -\bar{g}_L \cdot (V(t) - E_L) \\ &\quad - \sum_\xi \bar{g}_\xi r_\xi \cdot (V(t) - E_\xi), \end{aligned} \tag{4}$$

which is similar to the classical conductance-based form of Hodgkin and Huxley (1952) with one leak and two synaptic currents (say, $\xi \in \{AMPA, GABA\}$) (Weigenand et al., 2014). Moreover, for the convenience of calculating the convolution $\otimes$, $r_\xi(t)$ can be equivalently obtained by solving the following second-order differential equation

$$\ddot{r}_\xi(t) = \gamma_\xi^2 \cdot (N \cdot Q(t) - r_\xi(t)) - 2\gamma_\xi \dot{r}_\xi(t). \tag{5}$$

In (2)-(5), E_ξ and E_L denote the reversal potential of the synaptic current and leak current, respectively, $\bar{g}_\xi$ represents the synaptic input rate that scales r_ξ and E_ξ, g_L is the maximal conductivity of the leak current conductance, τ is the membrane time constant, N stands for the connectivity constant, and γ_ξ is the rate constant of synaptic response. The detailed dynamical evolution in each population is illustrated in **Figure 2**.

2.2. Mathematical Expression

This subsection formulates the proposed model TC-ACH mathematically in the logic of five populations included in the model.

2.2.1. Cholinergic Neuron Population (CH)

We first consider the cholinergic neuronal population, which is an essential part of TC-ACH in studying the effects of cholinergic modulation on thalamocortical rhythms during NREM sleep. Due to the special role of ACH, here, we apply the concentration of ACH ($[ACH]$) rather than the firing rate applied in Equation (1) to complete the transformation from the membrane potential, with Destexhe et al. (1994)

$$[ACH](t) = \frac{[ACH]^{max}}{1 + e^{-(V_c(t)-\theta_c)/\sigma_c}}, \tag{6}$$

where $[ACH]^{max}$ is the maximum concentration of ACH.

Moreover, CH receives an external input current I_{ext}. Therefore, the membrane potential $V_c(t)$ in CH obeys

$$\tau_c \dot{V}_c(t) = -I_L^c(t) - I_{ext}(t), \tag{7}$$

here, I_{ext} has the form (Rudolph et al., 2004)

$$I_{ext} = g_{ext}(t)(V_c(t) - E_c), \tag{8}$$

$$\dot{g_{ext}}(t) = \frac{-(g_{ext}(t) - g_0(t))}{\tau_{ext}} + \sqrt{\frac{2\sigma_{ext}^2}{\tau_{ext}}}\phi_C(t), \tag{9}$$

where g_0 is the average conductance, τ_{ext} is the time constant, σ_{ext} is the noise standard deviation (SD) value, and $\phi_C(t)$ denotes the independent Gaussian white noise process of unit SD and zero mean.

2.2.2. Populations PY and IN in the Cortical Module

According to Clearwater et al. (2008), we know that ACH acts on PY both pre-synaptically and post-synaptically over time scales from milliseconds to minutes.

In the post-synaptic case, ACH acts on PY through muscarinic receptors, which can cause a certain suppression of voltage-gated potassium channels (McCormick, 1992). Hence, the M-type potassium current I_M^p generated by the inactivation of these channels will change the intrinsic excitability of PY. Motivated by this observation, the membrane potential $V_p(t)$ in PY is then formulated by

$$\tau_p \dot{V}_p(t) = -I_L^p(t) - I_{AMPA}^p(t) - I_{GABA}^p(t) - I_M^p(t), \tag{10}$$

where $I_L^p, I_{AMPA}^p, I_{GABA}^p$ are calculated by Equations (2)–(4). The newly added current I_M^p is described as follows (Rich et al., 2018)

$$I_M^p(t) = g_M(t) \cdot (V_p(t) - E_M) z(t), \tag{11}$$

where $z(t)$ represents the unitless gating variable of the ionic current conductance satisfying

$$\dot{z}(t) = \frac{\overline{z}(t) - z(t)}{\tau_z}. \tag{12}$$

Here, $\overline{z}(t) = \frac{1}{1+e^{(-V_p(t)-39)/5}}$ stands for the voltage-sensitive steady-state activation function and τ_z is the time constant. Note that g_M in Equation (11) is not a constant anymore, but varies in term of $[ACH]$, that is,

$$g_M(t) = g_M^* + K_{PY} \cdot [ACH](t), \tag{13}$$

where g_M^* is a constant with the nominal value listed in **Table 1**.

In the pre-synaptic case, ACH acts on PY to modulate the properties of thalamocortical synapses via nicotinic receptor, which can result in changing the thalamocortical synaptic connection strength from TC to PY (denoted by N_{tp}) (Gil et al., 1997). Furthermore, as summarized by Kimura (2000), the thalamocortical connection strength will increase when the value of $[ACH]$ becomes larger. Hence, we apply the way in Clearwater et al. (2008) to reformulate the connection strength N in Equation (5) as

$$N_{tp}(t) = (N_{tp}^* - N_{tp}^\infty) e^{\frac{-K_{TP} \cdot [ACH](t)}{\mu}} + N_{tp}^\infty. \tag{14}$$

Obviously, there have

$$N_{tp}(t) = \begin{cases} N_{tp}^\infty, & \text{as } [ACH](t) \to \infty; \\ N_{tp}^*, & \text{as } [ACH](t) \to 0. \end{cases}$$

where N_{tp}^∞ and N^* represent the connectivity values under two extreme situations, and μ is a concentration constant.

Because in the cortical module IN contributes less than PY to the effects of cholinergic modulation on thalamocortical rhythms (Picciotto et al., 2012), in this paper, the dynamic evolution in IN is assumed to be not affected by $[ACH]$. That is to say, the mathematical formulation of IN obeys Equations (1)–(5).

TABLE 1 | Description and nominal values of parameters in model TC-ACH.

Symbol	Description	Value	Unit
C_m	Membrane capacitance in the HH model	1	$\mu F/cm^2$
Q_p^{max}	Maximal firing rate	$30 \cdot 10^{-3}$	ms
Q_i^{max}		$60 \cdot 10^{-3}$	
Q_t^{max}, Q_r^{max}		$400 \cdot 10^{-3}$	
$[ACH]^{max}$	Maximum [ACH] in the synaptic cleft	1	mM
θ	Firing threshold	-58.5	mV
σ_p	Inverse neural gain	4.7	mV
$\sigma_i, \sigma_t, \sigma_r$		6	
θ_c	[ACH] threshold	-40	mV
σ_c	The steepness of the sigmoid	4	mV
σ_{ext}	The noise standard deviation	12	nS
γ_e	Synaptic rate constant	$70 \cdot 10^{-3}$	ms^{-1}
γ_g		$58.6 \cdot 10^{-3}$	
γ_r, γ_c		$100 \cdot 10^{-3}$	
$N_{pp}, N_{ip}, N_{pi}, N_{ii}$	Connectivity constant	120, 72, 90, 90	/
N_{tp}^*, N_{tp}^∞		2.5, 5.5	
N_{rp}, N_{rt}, N_{tr}		2.6, 3, 5	
$\overline{N}_{rr}$		25	
N_{pt}, N_{it}		2.5	
τ_p, τ_i	Membrane time constant	30	ms
τ_t, τ_r, τ_c		20	
τ_z, τ_{ext}		75, 2.73	
$\overline{g}_{x \in \{AMPA,GABA\}}$	Input rate of synaptic channel	1	ms
g_M^*, g_T^t, g_h	Conductivity of ion channel	1.6, 3, 0.62	mS/cm^2
g_{LK}		0.18 – 0.55	
$\overline{g}_T^r$	Maximum calcium conductivity	3	mS/cm^2
g_{inc}	Conductivity scaling of h-current	2	/
g_{ACH-n}	Conductance mediated by nAChR	0.15	mS/cm^2
g_0	Average conductance	12.1	nS
E_L^p, E_L^i	Nernst reversal potential	-64	mV
E_L^t, E_L^r, E_L^c, E_c		-70	
E_{LK}, E_{Ca}, E_h		$-100, 120, -40$	
E_{AMPA}, E_{ACH-n}		0	
E_{GABA}		-70	
μ	Concentration constant	1	mM
a, b, c	Positive constant	0.05, 0.58, 0.42	/
ω_{max}	Maximum conductance	0.96	mS/cm^2
P_d	Apparent dissociation constant	0.028	mM
h	Hill coefficient of ACH binding to the receptors	1.8	/
v	Axonal rate constant	$120 \cdot 10^{-3}$	ms^{-1}

(Continued)

TABLE 1 | Continued

Symbol	Description	Value	Unit
ϕ_0	Mean background noise	0	
ϕ_P^{sd}, ϕ_I^{sd}	Standard deviation of cortical background noise	$120 \cdot 10^{-3}$	ms^{-1}
ϕ_T^{sd}	Standard deviation of thalamic background noise	$10 \cdot 10^{-3}$	
$K_{i\in\{PY,TP\}}$	Coupling coefficient of projection from CH	1.5 – 5.5	/
$K_{RE}^{j\in\{Ca,Nrr,I\}}$		1 – 9	

2.2.3. Populations RE and TC in Thalamic Module

We see in **Figure 1** that population RE appears to be a central hub in connecting the cortex and population CH. The electrical activity occurring in RE is important for regulating information transmission in the thalamocortical system and the shape of thalamocortical rhythms. Therefore, we first model the discharge activity of RE in this section.

On one hand, recent research indicates that thalamic reticular neurons can be directly excited by activating $\alpha 7$-containing nicotinic ACH receptor (Ni et al., 2016). Consequently, we introduce a new current I_{ACH-n} as an input of RE from CH (corresponding the blue solid line in **Figure 1**).

On the other hand, it has been shown that the potassium leak current and T-type calcium current are essential for the generation of spindle oscillation (Langdon et al., 2012). Based on this fact, two more currents I_{LK} and I_T are considered in the formulation of the membrane potential $V_r(t)$ in RE simultaneously.

Hence, the final equation is formulated by

$$\begin{aligned}\tau_r \dot{V}_r(t) &= -I_L^r(t) - I_{AMPA}^r(t) - I_{GABA}^r(t) - I_{ACH-n}(t) \\ &\quad -C_m^{-1}\tau_r \cdot (I_{LK}^r(t) - I_T^r(t)).\end{aligned} \tag{15}$$

In Equation (15), I_{ACH-n}^r, I_{LK}^r and I_T^r are defined, respectively, as

$$I_{ACH-n}(t) = g_{ACH-n} \cdot (V_r(t) - E_{ACH-n})\omega(t) \tag{16}$$

where $\omega(t)$ is solved by

$$\dot{\omega}(t) = \frac{\omega_{max}}{1 + (\frac{P_d}{K_{RE}^I \cdot [ACH](t)})^h}, \tag{17}$$

$$I_{LK}^r = g_{LK}^r \cdot (V_r(t) - E_{LK}^r), \tag{18}$$

$$I_T^r(t) = g_T^r(t) \cdot (V_r(t) - E_{Ca}^r)(m(t))^2 h(t). \tag{19}$$

Here, ω denotes the proportion of open ion-channels caused by binding of ACH. g_{ACH-n} and E_{ACH-n} stand for the conductance and reversal potential, respectively. P_d and h represent the apparent dissociation constant and Hill coefficient of ACH binding to the receptors, and ω_{max} is the maximum conductance. $m(t)$ and $h(t)$ are activation and inactivation functions of T-type current. Details about Equations (16)–(19) can be seen in Baran et al. (2010), Sethuramanujam et al. (2016), and Destexhe et al. (1993). In addition, the calcium conductance g_T^r in Equation (19) is recognized to be important for generating bursting oscillations in RE, whose value increases with increasing $[ACH]$ (Fisher and Johnston, 1990). Specifically, it is expressed by

$$g_T^r(t) = \overline{g}_T^r \frac{K_{RE}^{Ca} \cdot [ACH](t)}{K_{RE}^{Ca} \cdot [ACH](t) + a} \tag{20}$$

where $\overline{g}_T^r$ and a are positive constants (Omori and Horiguchi, 2004).

In addition, it has been found that the GABAergic projection within RE is decreased by increasing $[ACH]$ (Fisher and Johnston, 1990). This mechanism is then absorbed into defining the self-feedback connectivity of RE in our work and formulated by

$$N_{rr}(t) = (-b \cdot logK_{RE}^{Nrr} \cdot [ACH](t) + c) \cdot \overline{N}_{rr}, \tag{21}$$

where $\overline{N}_{rr}$, b and c are positive constants (Omori and Horiguchi, 2004).

$$\tau_p \dot{V}_p = -I_L^p - I_{AMPA}^p(r_{ep}) - I_{GABA}^p(r_{gp}) - C_m^{-1}\tau_p I_M \tag{22}$$

$$\tau_i \dot{V}_i = -I_L^i - I_{AMPA}^i(r_{ei}) - I_{GABA}^i(r_{gi}) \tag{23}$$

$$\begin{aligned}\tau_t \dot{V}_t &= -I_L^t - I_{AMPA}^t(r_{et}) - I_{GABA}^t(r_{gt}) - C_m^{-1}\tau_t \\ &\quad (I_{LK}^t - I_T^t - I_h)\end{aligned} \tag{24}$$

$$\begin{aligned}\tau_r \dot{V}_r &= -I_L^r - I_{AMPA}^r(r_{er}) - I_{GABA}^r(r_{gr}) - I_{ACH-n} \\ &\quad -C_m^{-1}\tau_r(I_{LK}^r - I_T^r)\end{aligned} \tag{25}$$

$$\tau_c \dot{V}_c = -I_L^c - I_{ext} \tag{26}$$

$$\ddot{r}_{ep} = \gamma_e^2(N_{pp}Q_p + N_{tp}\eta_t + \phi_P - r_{ep}) - 2\gamma_e \dot{r}_{ep} \tag{27}$$

$$\ddot{r}_{ei} = \gamma_e^2(N_{pi}Q_p + N_{ti}\eta_t + \phi_I - r_{ei}) - 2\gamma_e \dot{r}_{ei} \tag{28}$$

$$\ddot{r}_{et} = \gamma_e^2(N_{pt}\eta_p + \phi_T - r_{et}) - 2\gamma_e \dot{r}_{et} \tag{29}$$

$$\ddot{r}_{er} = \gamma_e^2(N_{tr}Q_t + N_{pr}\eta_p - r_{er}) - 2\gamma_e \dot{r}_{er} \tag{30}$$

$$\ddot{r}_{gp} = \gamma_g^2(N_{ip}Q_i - r_{gp}) - 2\gamma_g \dot{r}_{gp} \tag{31}$$

$$\ddot{r}_{gi} = \gamma_g^2(N_{ii}Q_i - r_{gi}) - 2\gamma_g \dot{r}_{gi} \tag{32}$$

$$\ddot{r}_{gt} = \gamma_r^2(N_{rt}Q_r - r_{gt}) - 2\gamma_r \dot{r}_{gt} \tag{33}$$

$$\ddot{r}_{gr} = \gamma_r^2(N_{rr}Q_r - r_{gr}) - 2\gamma_r \dot{r}_{gr} \tag{34}$$

$$\ddot{\eta}_p = \nu^2(Q_p - \eta_p) - 2\nu\dot{\eta}_p \tag{35}$$

$$\ddot{\eta}_t = \nu^2(Q_t - \eta_t) - 2\nu\dot{\eta}_t \tag{36}$$

Next, we turn to modeling the discharge activity of TC. Different from RE, there exists one more h-type current in TC, which is responsible for the waxing and waning structure of spindle rhythms in thalamus. Hence, the membrane potential $V_t(t)$ in TC is formulated by

$$\begin{aligned}\tau_t \dot{V_t}(t) &= -I_L^t(t) - I_{AMPA}^t(t) - I_{GABA}^t(t) - C_m^{-1} \cdot \tau_t \cdot (I_{LK}^t(t) \\ &\quad -I_T^t(t) - I_h(t)),\end{aligned} \tag{37}$$

where I_{LK}^t, I_T^t are calculated by Equations (18)–(19), and I_h is described as Destexhe et al. (1996)

$$I_h(t) = g_h \cdot (V_t(t) - E_h)(m_{h1}(t) + g_{inc} \cdot m_{h2}(t)). \tag{38}$$

Here, g_h and E_h represent the conductance and reversal potential, g_{inc} is the conductivity scaling. The details of functions $m_{h1}(t)$ and $m_{h2}(t)$ can be found in Destexhe et al. (1993).

2.2.4. Full Mathematical Expression of TC-ACH

There is, to emphasize, one key point in our model. On account of the long range afferent, there exist conduction delays between thalamus and cortex module. According to Costa et al. (2016), this delay is approximated as a convolution with the alpha function $h(t)$. In this case, Equation (5) is then reformulated by

$$\ddot{r}_\xi(t) = \gamma_\xi^2(N \cdot (h(t) \otimes Q(t)) - r_\xi(t)) - 2\gamma_\xi \dot{r}_\xi(t) \qquad (39)$$

during four transmissions "PY → RE," "PY → TC," "TC → PY," and "TC → IN."

The full mathematical expression of TC-ACH is then formulated by Equations (22)–(36) (see page 5). Note that the model output is V_p, which can be viewed as the simulated EEG signals.

Remark 1. The description and nominal values of all parameters included in the model TC-ACH are listed in **Table 1**, whose values are the nominal ones reported in Costa et al. (2016), Omori and Horiguchi (2004), Clearwater et al. (2008), Rich et al. (2018), Baran et al. (2010), Sethuramanujam et al. (2016), Bhattacharya et al. (2012), and Rudolph et al. (2004).

3. RESULTS

In this section, we apply the proposed model TC-ACH to verify the effects of cholinergic modulation on thalamocortical rhythms during NREM in two ways: (1) measurement of effects of cholinergic modulation in the thalamus; (2) measurement of effects of cholinergic modulation in the thalamocortical system.

All of the numerical simulations are performed in MATLAB R2017b (MathWorks,USA), using a stochastic Runge-Kutta method of 4th order (Rößler, 2010) with a step size of 0.1ms.

3.1. The Oscillatory Phenomena Due to Cholinergic Modulation in Thalamus

The thalamus is believed to act as a "pacemaker" for thalamocortical rhythms, and is able to independently generate multiple brain rhythms during sleep (Hughes and Crunelli, 2005; Li et al., 2017). Hence, oscillatory phenomena due to cholinergic modulation in thalamus are first explored, where only the thalamic module and projections from CH to RE are considered. Here, the model output is V_t.

We first show different thalamic oscillations and transitions between them caused by cholinergic projections. To this end, we apply one-dimensional bifurcation analysis for several key parameters, represented by g_{LK}, K_{RE}^I, K_{RE}^{Ca}, and K_{RE}^{Nrr}. The bifurcation diagram is obtained by plotting the stable local minimum, as well as top three maximum values of V_t over changes in each of four parameters. All simulations are executed for 40 s and those minimum and maximum values are obtained from the latter stable 30 s of the time series.

Figures 3A–D shows bifurcation diagrams of V_t over changes in g_{LK}, K_{RE}^I, K_{RE}^{Ca}, and K_{RE}^{Nrr}, respectively. Seen from **Figure 3A**, it reveals that there are four different dynamical states including low firing (I), spindle (II), fast oscillation (III) and slow oscillation (IV) as g_{LK} varies in [0.018, 0.055]. Specifically, when the value of g_{LK} is extremely small, the model exhibits a low firing state and no oscillation behavior can be observed. As g_{LK} becomes a little larger, the model mainly experiences a spindle oscillation pattern for a period, in which multiple pairs of maximum and minimum values are found within each periodic complex. With further growth of g_{LK}, the model exhibits a fast oscillation pattern, in which only one pair of maximum and minimum values emerges within each periodic complex. When strong hyperpolarization through increasing g_{LK} attains, the model moves from a high frequency oscillation pattern into a slow oscillation pattern, in which multiple pairs of maximum and minimum values can also be observed within each periodic complex, but the distances between the top maximum and other maximum values become larger. **Figure 3E** illustrates the obtained thalamic oscillations, which correspond to four dynamical states ($g_{LK} = 0.018, 0.03, 0.045, 0.052$), respectively.

A similar explanation can be given for the cases of K_{RE}^I, K_{RE}^{Ca}, and K_{RE}^{Nrr}, while only two dynamical states (i.e., II and III) are obtained. Moreover, the state transition of K_{RE}^I is different from that of K_{RE}^{Ca} and K_{RE}^{Nrr}. It is apparent that the model has a pattern transition from state III to an increasingly stable state II as K_{RE}^I increases (see **Figure 3B**). By contrast, a transition from state II to state III with the increase of K_{RE}^{Ca} (or K_{RE}^{Nrr}) can be observed in **Figure 3C** (or **Figure 3D**). That is to say, the model appears to demonstrate a stabler spindle oscillation state with the increasing value of K_{RE}^I, while the growth of K_{RE}^{Ca} (or K_{RE}^{Nrr}) leads to reduced spindle rhythms. Prior work has established that spindle rhythms are helpful to protect sleep (Dang-Vu et al., 2010; Kim et al., 2012). Therefore, the obtained results implicate that cholinergic modulation in RE may help promote sleep (or arousal) states.

Next, we check whether our results can be generalized within a certain range of parameters. The above bifurcation analysis allows us to further distinguish different dynamical state regions in the two-parameter space (for example, see **Figure 5A**). Moreover, the power spectral analysis is applied to estimate the dominant frequency (Df) and second dominant frequency (Sec-Df) of neural oscillations from the time series of V_t. **Figure 4** illustrates the box-plots of Df and Sec-Df with respect to three oscillations (fast, slow and spindle). It is clear from **Figure 4** that the Sec-Df has better differentiation capability than Df. Therefore, the corresponding Sec-Df regions are drawn for each pair of parameters (for example, see **Figure 5B**) in the following analysis.

Figures 5A,B illustrate the dynamical state regions and Sec-Df regions with 41×41 grids in the space of $g_{LK} \times K_{RE}^I \in [0.018, 0.055] \times [1, 9]$. As shown in **Figure 5A**, four different state regions are displayed, whose identification is same as above [i.e., low firing (I), spindle (II), fast oscillation (III) and slow oscillation (IV)]. It can be observed that along with K_{RE}^I rising, the model generates more spindle oscillation pattern after lasting a certain fast oscillation period (see the black arrow). By combining the results of frequency analysis shown in **Figure 5B**, we can outline the spindle oscillation region that falls into the 11–14 Hz frequency range. It should be noted that, the obtained result is

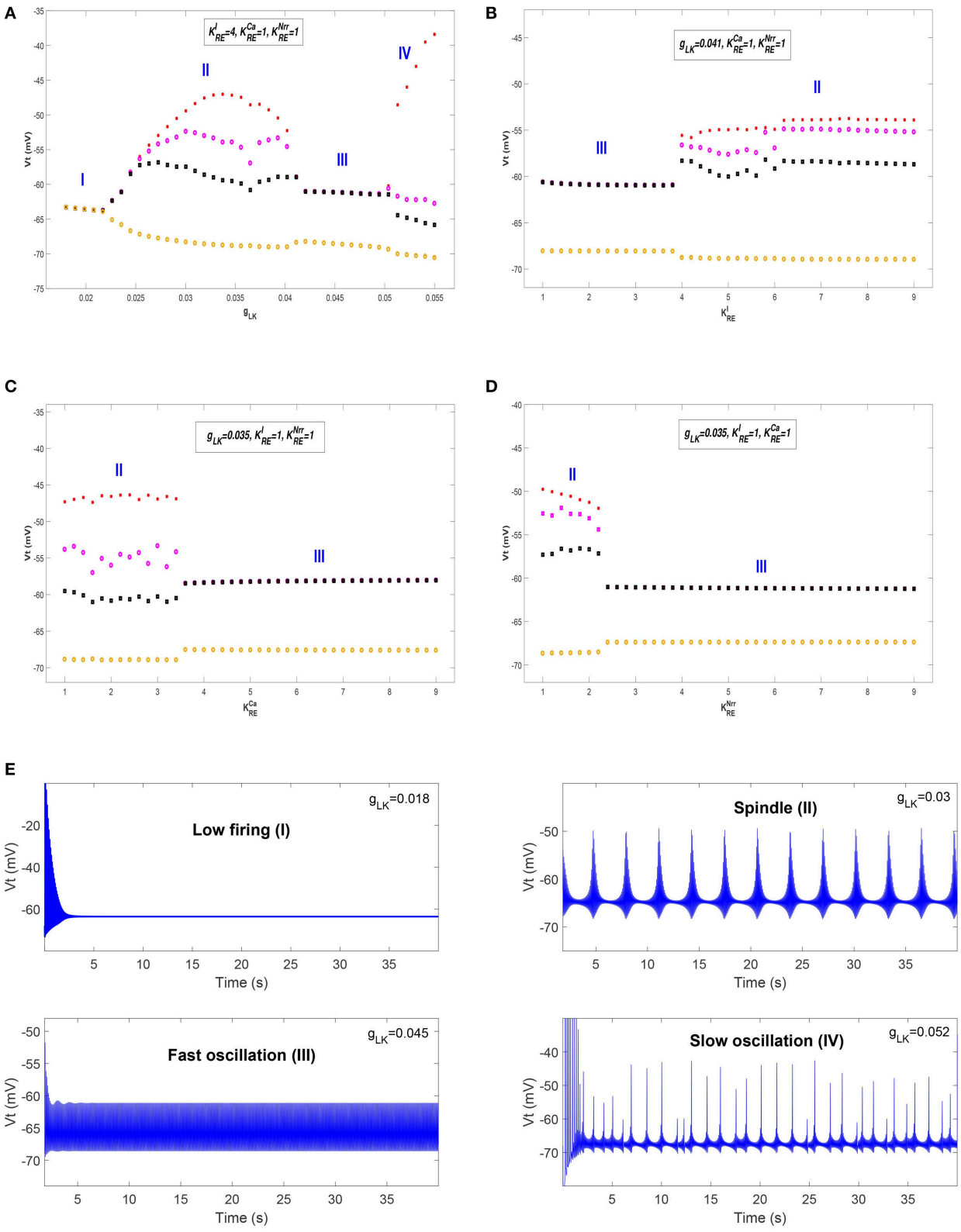

FIGURE 3 | The bifurcation diagrams of V_t over changes in g_{LK}, K^{I}_{RE}, K^{Ca}_{RE}, and K^{Nrr}_{RE}. **(A)** Bifurcation diagram (varying g_{LK}). **(B)** Bifurcation diagram (varying K^{I}_{RE}). **(C)** Bifurcation diagram (varying K^{Ca}_{RE}). **(D)** Bifurcation diagram (varying K^{Nrr}_{RE}). **(E)** Four thalamic oscillations with respect to different values of g_{LK}.

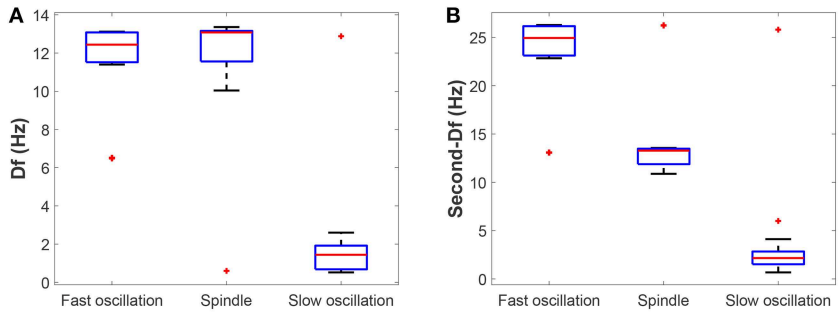

FIGURE 4 | The box-plots of **(A)** Df and **(B)** Sec-Df with respect to three oscillations (fast, slow, and spindle).

consistent with the frequency range [11, 16]Hz characteristic of sleep spindles.

Similar results are obtained in the spaces of $g_{LK} \times K_{RE}^{Ca}$ and $g_{LK} \times K_{RE}^{N_{rr}}$, which are illustrated in **Figures 5C–F**. We observe that increasing K_{RE}^{I} promotes the generation of spindle oscillations in our model (see the black solid arrow in **Figure 5A**), while increasing K_{RE}^{Ca} (or $K_{RE}^{N_{rr}}$) suppresses generation of spindle oscillations (see the black dash arrow in **Figures 5C,E**). These results are in line with the results mentioned above.

Additionally, in order to demonstrate the interactions among parameters K_{RE}^{I}, K_{RE}^{Ca}, and K_{RE}^{Nrr}, the dynamical state regions and Sec-Df regions for each pairwise combination of them are shown in **Figure 6**. As expected, both state and Sec-Df analysis in three panels (K_{RE}^{I}, K_{RE}^{Ca}), (K_{RE}^{I}, K_{RE}^{Nrr}), and (K_{RE}^{Ca}, K_{RE}^{Nrr}) provide the same evidence as the above description, that is, the spindle oscillation is generated as K_{RE}^{I} increases (or as K_{RE}^{Ca}, K_{RE}^{Nrr} decreases).

3.2. The Oscillatory Phenomena Due to Cholinergic Modulation in the Thalamocortical System

Here we concentrate on the effects of cholinergic modulation on thalamocortical rhythms during NREM sleep, including K-complexes, spindles, and slow oscillations. We fix the values of the four parameters to be $g_{LK} = 0.034$, $K_{RE}^{I} = 5.2$, $K_{RE}^{Ca} = 1.5$, and $K_{RE}^{Nrr} = 2.1$. Based on the results obtained in the proceeding subsection, we know that spindle oscillations can be generated in the thalamus module with such settings.

We first show the dynamical behavior of our TC-ACH model with variation of K_{PY}, which is the strength of the cholinergic projection from CH to PY. In this case, we assume that there is no cholinergic projection from CH to "PY-TC connection" ($K_{TP} = 0$), and we apply linear ramps to increase or decrease the parameter value and observe the effects of continuous changes in K_{PY} on the model. **Figure 7B** illustrates the model output V_p as K_{PY} varies in [2.3, 5] linearly (see **Figure 7A**). In the beginning, when $K_{PY} = 2.3$, the model exhibits K-complex oscillations. Then we can find that a small ascent of K_{PY} drives a transition from K-complexes to slow oscillations. Furthermore, with the gradual increase of K_{PY}, the slow oscillation is kept for a short period until K_{PY} reaches its higher threshold. At this time, the slow oscillation is replaced by the α-like oscillation. As K_{PY} ramps down, the model transitions back into its original state gradually (that is, from α-like to slow oscillations, and then from slow oscillations to K-complexes). The enlarged graphs of three different oscillations in 20 s are displayed in **Figures 7E–G**.

On the basis of the above observations, we conclude that K-complexes and slow oscillations can be triggered by weakening the strength of the cholinergic projection from CH to cortex, while the emergence of α-like activity, characteristic of REM, requires stronger cholinergic input. Here, if we roughly consider that K-complexes and slow oscillations mainly emerge during NREM sleep and α-like rhythms mostly appear during REM sleep, the obtained results in our work are consistent with the conclusion in Lena et al. (2005), that is, the concentration of ACH is lowest during NREM sleep and highest during wake and REM sleep.

The spectral analysis and the sample entropy extracted from the model output V_p also support this conclusion from another point of view. **Figure 7C** illustrates the spectrogram of V_p. During the first and last 40 s where K-complexes and slow oscillations are emerging, the spectral power of V_p mainly falls within the [0, 5]Hz range; while in the middle stage where α-like rhythms begin to appear, it increases to [0, 10]Hz. These results obviously conform with the frequency ranges of certain sleep rhythms. Besides, the sample entropy extracted from V_p also shows a similar tendency. In detail, it is larger in the α-like stage, but smaller in the K-complex and slow wave oscillation stage (see **Figure 7D**). Sample entropy measures signal complexity, hence this observation is consistent with the conclusion that the complexity of cortical rhythms are decreased as the deeper sleep state, but increased during REM and wake (Bruce et al., 2009; See and Liang, 2011).

Next, we show the dynamical behavior of constructed model TC-ACH with variation of K_{TP}, which is the strength of the cholinergic projection from CH to the connection between PY and TC. Here, K_{PY} shows the trend as in **Figure 7**, and K_{TP} is set to be 1.5, 3.5, 5.5, respectively, in simulations. **Figure 8** illustrates the model output V_p under three different cases.

In the first case where the connectivity between TC and PY is regulated by a lower strength of cholinergic projection from CH (i.e., $K_{TP} = 1.5$), we observe in **Figure 8A1** that some weaker spindle rhythms emerge when the value of K_{PY} is relatively

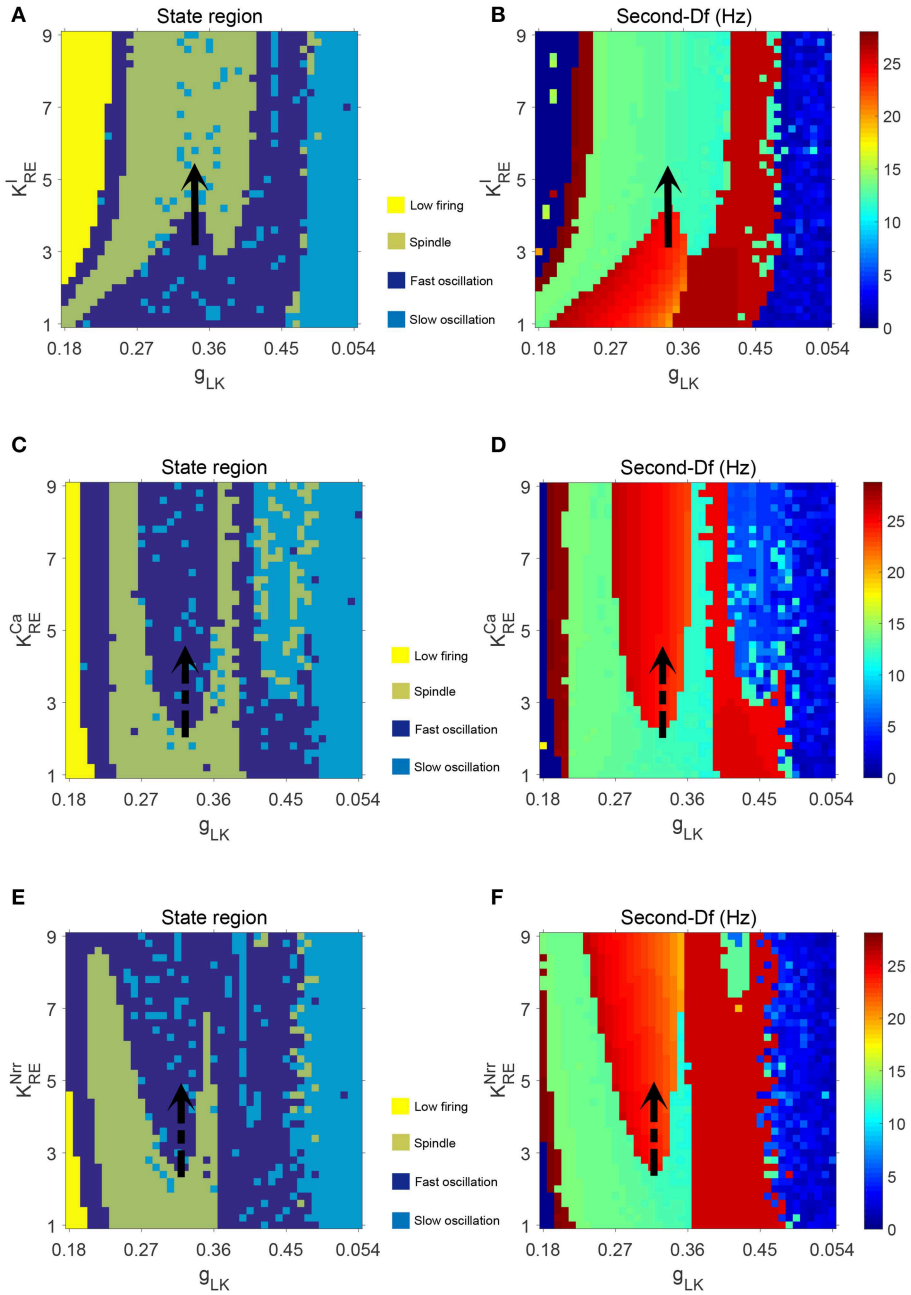

FIGURE 5 | The dynamical state regions **(A,C,E)** and Sec-Df regions **(B,D,F)** in three panels (g_{LK}, K^{I}_{RE}), (g_{LK}, K^{Ca}_{RE}), $(g_{LK}, K^{N_{rr}}_{RE})$.

small. It can be observed more clearly in **Figure 8A2** where the enlarged graph exhibits spindle rhythms (with red circle) and K-complex oscillations in [20, 40] s. Meanwhile, in **Figures 8A3,A4**, we see that the duration of slow oscillation is much longer, and α-like activity is much less by comparing the results shown in **Figure 7**.

This conclusion can be further verified under other two cases. It can be seen in **Figures 8B1–B4** that more spindle rhythms appear as K_{PY} gradually increases and the α-like activities have been largely replaced by slow waves in the case where $K_{TP} = 3.5$. Furthermore, when K_{TP} is set to be 5.5, the spindle rhythms run through the model output from the beginning to end. We can see from **Figure 8C4** that there are still spindle rhythms when K_{PY} attains its maximum value. These observations reveal that the strength of cholinergic projections from CN to "PY-TC connection" play an important role in promoting the spindle rhythms and prolonging the duration of the NREM state in the thalamocortical system.

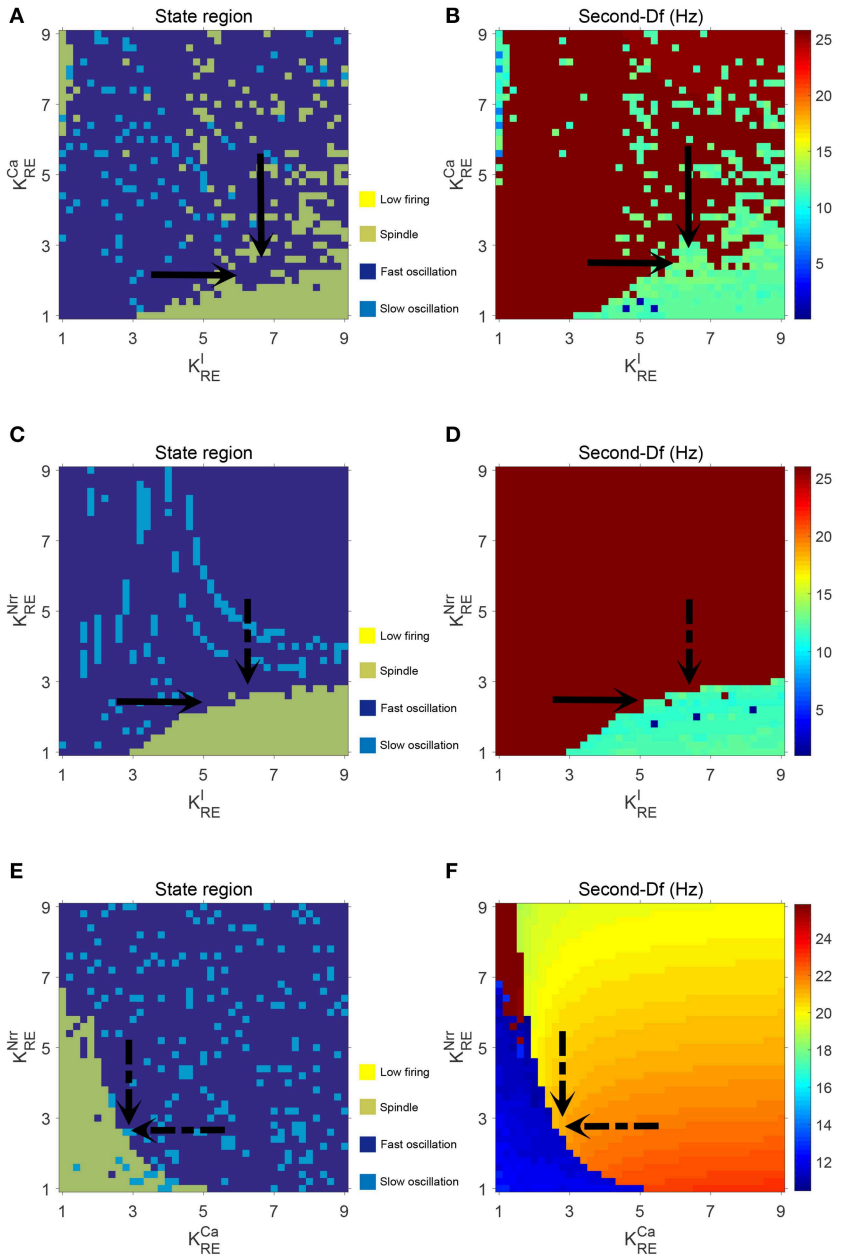

FIGURE 6 | The dynamical state regions **(A,C,E)** and Sec-Df regions **(B,D,F)** in three panels (K^{I}_{RE}, K^{Ca}_{RE}), (K^{I}_{RE}, K^{Nrr}_{RE}), and (K^{Ca}_{RE}, K^{Nrr}_{RE}).

4. SUMMARY AND DISCUSSION

In this paper, we first proposed a novel computational model (TC-ACH) by integrating a neuron population CH into classical thalamo-cortical circuitry (including populations PY, IN, TC, and RE). The connections between five populations are built in accordance with the established mechanisms. Our model considers the neurotransmitter ACH released by neurons in CH, which alters discharge activities of thalamocortical neurons. For simplicity, we represent these additions to the classic model by four cholinergic projections, where coupling coefficients $K^{Ca}_{RE}, K^{N_{rr}}_{RE}, K^{I}_{RE}$ represent different projections from CH to thalamus, and K_{PY} and K_{TP} represent the projections from CH to cortex and thalamocortical system, respectively. On the basis of established model framework of TC-ACH, the corresponding mathematical expression has been formulated in the logic of five populations systematically, where the average membrane potential $V(t)$ is solved to simulate the rhythms generated by each population.

Next, we applied the developed model TC-ACH to study the effects of ACH modulation on thalamocortical rhythms during NREM sleep in two ways:

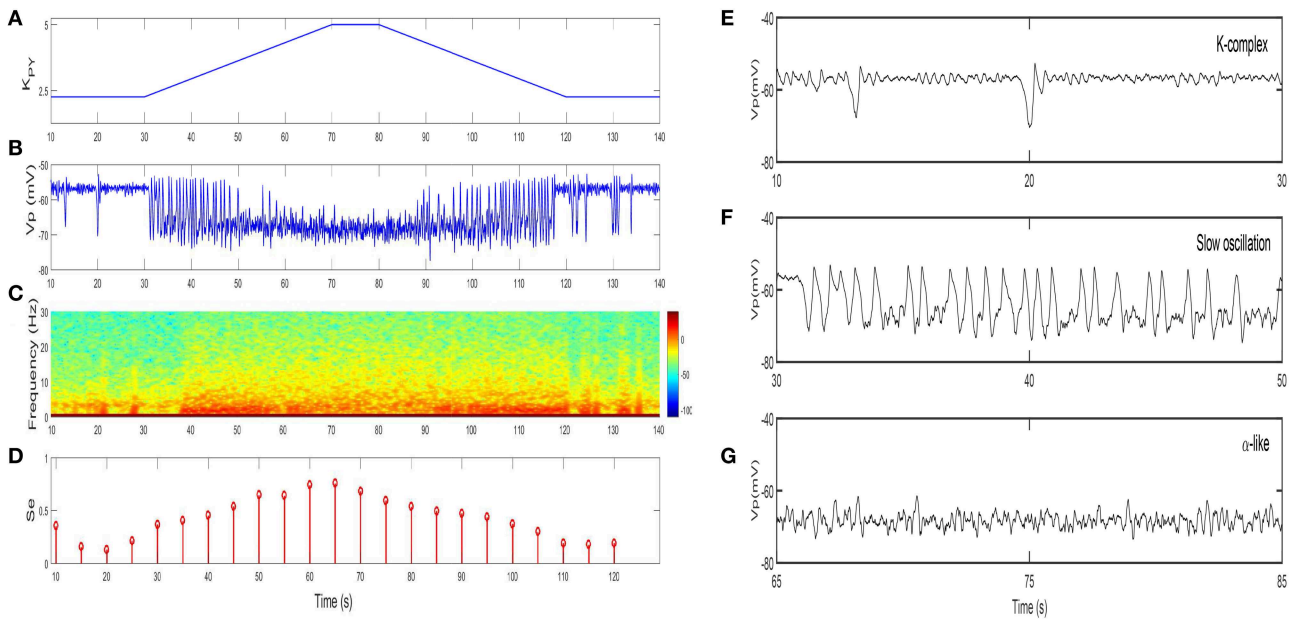

FIGURE 7 | The dynamical behavior of model TC-ACH with the variation of K_{PY}: **(A)** the variation of K_{PY} over time t; **(B)** the model output V_p; **(C)** the spectrogram of V_p; **(D)** the sample entropy of V_p; **(E)** the enlarged model output in [10,30]s (including K-complex); **(F)** the enlarged model output in [30,50]s (including slow oscillation); **(G)** the enlarged model output in [65,85]s (including α-like oscillation).

(1) *Measurement of effects of cholinergic modulation in the thalamus.* In this case, only the thalamic module and projections from CH to RE were considered. Simulation results suggest that cholinergic projection activity is a key factor in modulating oscillation patterns in the thalamic module. Specifically, the model appears to be a stabler spindle oscillation state with the increasing value of K_{RE}^{I}, while the growth of K_{RE}^{Ca} (or K_{RE}^{Nrr}) leads to reduced spindle rhythms. Moreover, with variation of the potassium leak conductance g_{LK}, which is dramatically modulated by the concentration of ACH (McCormick, 1989, 1992), four different dynamical states including the low firing, spindle, fast oscillation, and slow oscillation can be obtained and transited.

(2) *Measurement of effects of cholinergic modulation in the thalamocortical system.* In this case, the dynamical behavior of our TC-ACH model was studied by varying K_{PY} and K_{TP}, respectively. Simulation results show that the K-complex and slow oscillations can be triggered by weakening strength of K_{PY}, while emergence of α-like activity requires stronger input. Furthermore, when there exists cholinergic input from CH to "PY-TC connection" (that is, $K_{TP} \neq 0$), we found that the duration of oscillation patterns during NREM sleep including K-complexes, spindles and slow oscillations is longer. Additionally, when K_{TP} is relatively larger, more spindle rhythms appear and α-like activities are largely replaced by slow waves.

It should be noted that a number of electrophysiological experiments have investigated cholinergic modulation of cortex, thalamus or thalamocortical system, respectively (McCormick and Prince, 1986; McCormick, 1989, 1992; Clarke, 2004; Mesulam, 2004; Hasselmo and Giocomo, 2006; Beierlein, 2014). However, few studies correlate ACH modulation with rhythmic activities (Steriade et al., 1993; Steriade, 2004), where the model-based work to study in theory the effects of cholinergic modulation on thalamocortical rhythms during sleep is far less. More than that, what few existing studies are not comprehensive enough in studying the effects of ACH modulation on thalamocortical rhythms, especially during sleep. For example, the effects of ACH modulation are considered only on cortex or thalamus separately, but not on the whole thalamocortical system (Omori and Horiguchi, 2004; Li et al., 2017); the mechanism regarding ACH modulation is only modeled as a certain parameter, but not as a whole neuron population (Li et al., 2017); the model is constructed at the microscopic level, which cannot relate directly to thalamocortical rhythms at the mesoscopic level (Omori and Horiguchi, 2004). Therefore, in order to overcome such limitations, we have constructed a novel computational model (TC-ACH) by incorporating a cholinergic neuron population into the classical thalamo-cortical circuitry at the mesoscopic level. By thus doing, a deeper understanding of the role of cholinergic modulation on thalamocortical system will be got, and further, a critical insight into the mechanisms controlling sleep state may be found.

Besides five projections applied in the modeling (represented by $K_{RE}^{Ca}, K_{RE}^{Nrr}, K_{RE}^{I}$, K_{PY}, and K_{TP}), there still exist cholinergic modulators on other projections, such as corticoreticular and corticothalamic projections. Castro-Alamancos and Calcagnotto (2001) demonstrated that the corticothalamic

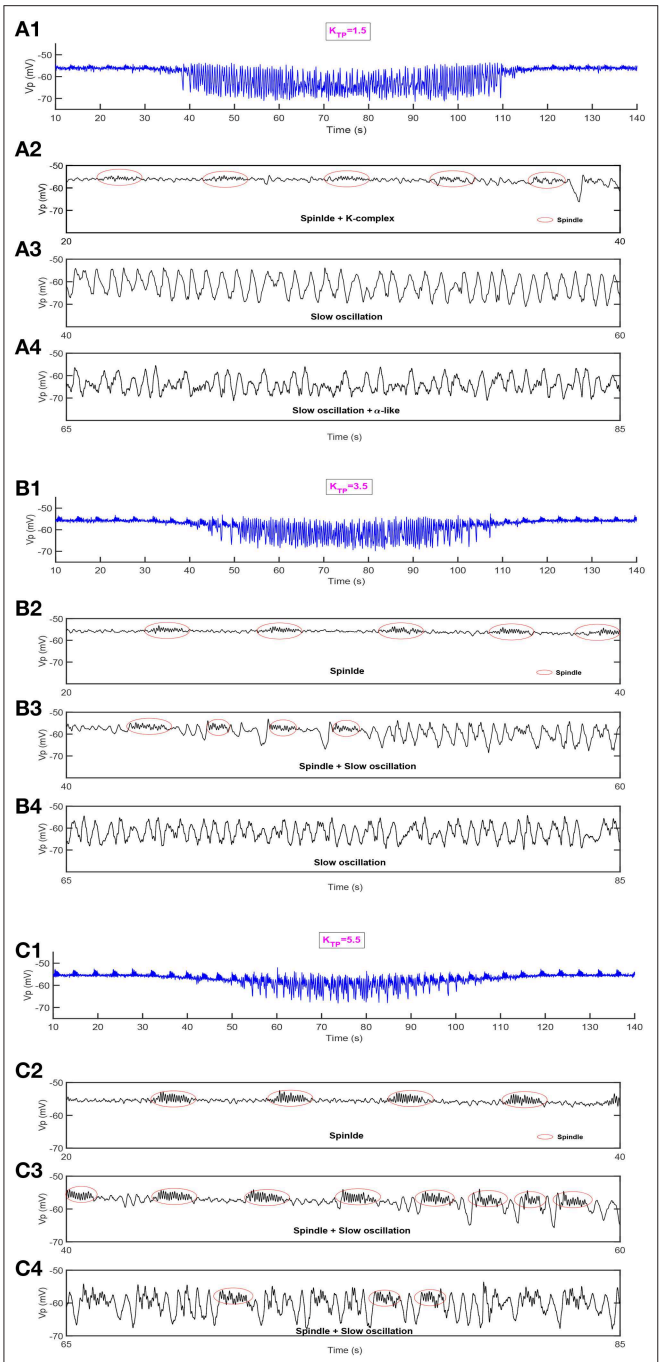

FIGURE 8 | The dynamical behavior of model TC-ACH in the cases where $K_{TP} = 1.5, K_{TP} = 3.5, K_{TP} = 5.5$. **(A1,B1,C1)** The model output Vp; **(A2,B2,C2)** the enlarged model output in [20,40]s; **(A3,B3,C3)** the enlarged model output in [40,60]s; **(A4,B4,C4)** the enlarged model output in [65,85]s.

activity can be filtered by cholinergic activation during arousal high-pass according to the experiments performed *in vitro* and *in vivo*. In addition, as indicated in Itier and Bertrand (2002), the cholinergic modulation on corticoreticular projection may induce the generation of more complex brain rhythms (such as spike and wave) during sleep. However, the quantitative description relating such cholinergic mechanisms to the generation and transition of typical rhythms during NREM sleep is very limited. Therefore, we hope these mechanisms could be further considered with more attempts, to study the cholinergic modulation of thalamocortical rhythms during sleep in different cases.

Another point to stress here is the synaptic connection ways in intra-RE. Previous electrophysiological studies indicated that the thalamic reticular neurons are functionally connected through chemical/electrical synapses (Sanchez-Vives et al., 1997; Landisman et al., 2002; Shu and McCormick, 2002; Long et al., 2004; Deleuze and Huguenard, 2006; Lam et al., 2006). A potentially paradigm-shifting question has been presented as to whether chemical synapses between thalamic reticular neurons are altogether absent in certain mammals or degenerate as a function of increasing age (Landisman et al., 2002; Cruikshank et al., 2010; Hou et al., 2016). However, there still have other works supporting the existence of GABAergic intrareticular synapses. For example, one of the most recent model-based study shows that intrareticular synapses, both chemical and electrical, manifest certain effects on the signal propagation and oscillation (Brown et al., 2019). In our modeling work, by means of considering the chemical synapses in intra-RE population, the measurement of effects of cholinergic modulation in thalamocortical system has been completed well. In contrast, the electrical synapses do not work effectively under the same situation. Consequently, we roughly hypothesize that GABAergic synapses may contribute much more than electrical synapses to the effects of cholinergic modulation in thalamocortical rhythms, at least at the mesoscopic level. But on the other hand, we certainly acknowledge that the electrical intrareticular synapses play important roles vis-à-vis thalamic signaling. Therefore, an important extension of the current work would be the modeling of electrical synapses between reticular neurons from the microscopic point of view, such as the model-based works in Pham and Haas (2018) and Brown et al. (2019).

An important limitation of our TC-ACH model is that it considers only ACH modulation. However, it is known that thalamocortical rhythms during sleep are also directly affected by other neuromodulators, such as noradrenalin (NE), serotonin (5-HT), histamine (HA), and dopamine (DA) from the hypothalamus and brainstem, whose concentrations vary over the night (Lena et al., 2005). Therefore, some apparent questions need to be answered: whether our method can be further developed to shed light on other sleep-related neuromodulators? If so, how to model the corresponding mechanisms and explore their effects on thalamocortical rhythms during sleep? Fortunately, based on the progressive mathematical description of sleep regulatory networks (Kumar et al., 2012; Booth et al., 2017), it may well be possible to carry our model further by constructing a new thalamocortical NMM which contains various sleep-related neuron populations (releasing NE, 5-HT, HA, DA); this is a topic we plan to pursue in future work. In

addition, such model-based research can make a contribution to understanding sleep related pathological conditions, such as sleep-related epilepsy.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

QL, J-LS, S-HL, MW, and RZ designed and performed the research as well as wrote the paper.

FUNDING

This work was supported by the National Natural Science Foundation of China under Grant 61473223, the Innovative Talents Promotion Plan of Shaanxi Province under Grant 2018TD-016, the Key Industry Innovation Chain (Group) of Shaanxi Province under Grant 2019ZDLSF02-09-02. MW was supported during this research by the Glenn Foundation for Medical Research and the American Federation for Aging Research through a Breakthroughs in Gerontology Grant; through the American Academy of Sleep Medicine through an AASM Foundation Strategic Research Award; and by grants from the NIH (1R01NS102190, 1R01NS102574, 1R01NS107291, 1RF1AG064312).

REFERENCES

Baran, I., Iftime, A., and Popescu, A. (2010). Diffusion–convection effects on drug distribution at the cell membrane level in a patch-clamp setup. *BioSystems* 102, 134–147. doi: 10.1016/j.biosystems.2010.09.003

Beierlein, M. (2014). Synaptic mechanisms underlying cholinergic control of thalamic reticular nucleus neurons. *J. Physiol.* 592, 4137–4145. doi: 10.1113/jphysiol.2014.277376

Bellingham, M. C., and Funk, G. D. (2000). Cholinergic modulation of respiratory brain-stem neurons and its function in sleep–wake state determination. *Clin. Exp. Pharmacol. Physiol.* 27, 132–137. doi: 10.1046/j.1440-1681.2000.03192.x

Bhattacharya, B. S., Coyle, D., Maguire, L. P., and Stewart, J. (2012). "Kinetic modelling of synaptic functions in the alpha rhythm neural mass model," in *International Conference on Artificial Neural Networks*, eds A. E. P. Villa (Berlin; Heidelberg: Springer), 645–652. doi: 10.1007/978-3-642-33269-2_81

Booth, V., Xique, I., and Diniz Behn, C. G. (2017). One-dimensional map for the circadian modulation of sleep in a sleep-wake regulatory network model for human sleep. *SIAM J. Appl. Dyn. Syst.* 16, 1089–1112. doi: 10.1137/16M1071328

Boutrel, B., and Koob, G. F. (2004). What keeps us awake: the neuropharmacology of stimulants and wakefulness promoting medications. *Sleep* 27, 1181–1194. doi: 10.1093/sleep/27.6.1181

Brown, J. W., Taheri, A., Kenyon, R. V., Berger-Wolf, T., and Llano, D. A. (2019). A computational model of intrathalamic signaling via open-loop thalamo-reticular-thalamic architectures. *bioRXiv*. 574178. doi: 10.1101/574178

Brown, R. E., Basheer, R., McKenna, J. T., Strecker, R. E., and McCarley, R. W. (2012). Control of sleep and wakefulness. *Physiol. Rev.* 92, 1087–1187. doi: 10.1152/physrev.00032.2011

Bruce, E. N., Bruce, M. C., and Vennelaganti, S. (2009). Sample entropy tracks changes in EEG power spectrum with sleep state and aging. *J. Clin. Neurophysiol.* 26:257. doi: 10.1097/WNP.0b013e3181b2f1e3

Castro-Alamancos, M. A., and Calcagnotto, M. E. (2001). High-pass filtering of corticothalamic activity by neuromodulators released in the thalamus during arousal: *in vitro* and *in vivo*. *J. Neurophysiol.* 85, 1489–1497. doi: 10.1152/jn.2001.85.4.1489

Clarke, P. B. (2004). Nicotinic modulation of thalamocortical neurotransmission. *Prog. Brain Res.* 145, 253–260. doi: 10.1016/S0079-6123(03)45017-6

Clearwater, J., Rennie, C., and Robinson, P. (2008). Mean field model of acetylcholine mediated dynamics in the thalamocortical system. *J. Theor. Biol.* 255, 287–298. doi: 10.1016/j.jtbi.2008.08.010

Cona, F., Lacanna, M., and Ursino, M. (2014). A thalamo-cortical neural mass model for the simulation of brain rhythms during sleep. *J. Comput. Neurosci.* 37, 125–148. doi: 10.1007/s10827-013-0493-1

Costa, M. S., Weigenand, A., Ngo, H.-V. V., Marshall, L., Born, J., Martinetz, T., et al. (2016). A thalamocortical neural mass model of the EEG during NREM sleep and its response to auditory stimulation. *PLoS Comput. Biol.* 12:e1005022. doi: 10.1371/journal.pcbi.1005022

Cruikshank, S. J., Urabe, H., Nurmikko, A. V., and Connors, B. W. (2010). Pathway-specific feedforward circuits between thalamus and neocortex revealed by selective optical stimulation of axons. *Neuron* 65, 230–245. doi: 10.1016/j.neuron.2009.12.025

Dang-Vu, T. T., McKinney, S. M., Buxton, O. M., Solet, J. M., and Ellenbogen, J. M. (2010). Spontaneous brain rhythms predict sleep stability in the face of noise. *Curr. Biol.* 20, R626–R627. doi: 10.1016/j.cub.2010.06.032

Deleuze, C., and Huguenard, J. R. (2006). Distinct electrical and chemical connectivity maps in the thalamic reticular nucleus: potential roles in synchronization and sensation. *J. Neurosci.* 26, 8633–8645. doi: 10.1523/JNEUROSCI.2333-06.2006

Destexhe, A., Babloyantz, A., and Sejnowski, T. J. (1993). Ionic mechanisms for intrinsic slow oscillations in thalamic relay neurons. *Biophys. J.* 65, 1538–1552. doi: 10.1016/S0006-3495(93)81190-1

Destexhe, A., Bal, T., McCormick, D. A., and Sejnowski, T. J. (1996). Ionic mechanisms underlying synchronized oscillations and propagating waves in a model of ferret thalamic slices. *J. Neurophysiol.* 76, 2049–2070. doi: 10.1152/jn.1996.76.3.2049

Destexhe, A., Mainen, Z. F., and Sejnowski, T. J. (1994). Synthesis of models for excitable membranes, synaptic transmission and neuromodulation using a common kinetic formalism. *J. Comput. Neurosci.* 1, 195–230. doi: 10.1007/BF00961734

Fisher, R., and Johnston, D. (1990). Differential modulation of single voltage-gated calcium channels by cholinergic and adrenergic agonists in adult hippocampal neurons. *J. Neurophysiol.* 64, 1291–1302. doi: 10.1152/jn.1990.64.4.1291

Gil, Z., Connors, B. W., and Amitai, Y. (1997). Differential regulation of neocortical synapses by neuromodulators and activity. *Neuron* 19, 679–686. doi: 10.1016/S0896-6273(00)80380-3

Hasselmo, M., and Giocomo, L. (2006). Cholinergic modulation of cortical function. *J. Mol. Neurosci.* 30, 133–135. doi: 10.1385/JMN:30:1:133

Hodgkin, A. L., and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* 117, 500–544. doi: 10.1113/jphysiol.1952.sp004764

Hou, G., Smith, A. G., and Zhang, Z.-W. (2016). Lack of intrinsic gabaergic connections in the thalamic reticular nucleus of the mouse. *J. Neurosci.* 36, 7246–7252. doi: 10.1523/JNEUROSCI.0607-16.2016

Hughes, S. W., and Crunelli, V. (2005). Thalamic mechanisms of EEG alpha rhythms and their pathological implications. *Neuroscientist* 11, 357–372. doi: 10.1177/1073858405277450

Itier, V., and Bertrand, D. (2002). Mutations of the neuronal nicotinic acetylcholine receptors and their association with adnfle. *Neurophysiol. Clin.* 32, 99–107. doi: 10.1016/S0987-7053(02)00294-0

Jansen, B. H., Zouridakis, G., and Brandt, M. E. (1993). A neurophysiologically-based mathematical model of flash visual evoked potentials. *Biol. Cybernet.* 68, 275–283. doi: 10.1007/BF00224863

Kim, A., Latchoumane, C., Lee, S., Kim, G. B., Cheong, E., Augustine, G. J., et al. (2012). Optogenetically induced sleep spindle rhythms alter sleep architectures in mice. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20673–20678. doi: 10.1073/pnas.1217897109

Kimura, F. (2000). Cholinergic modulation of cortical function: a hypothetical role in shifting the dynamics in cortical network. *Neurosci. Res.* 38, 19–26. doi: 10.1016/S0168-0102(00)00151-6

Kobayashi, Y., Inoue, Y., and Isa, T. (2003). Cognitive function of the brainstem cholinergic system. *Int. Congr. Ser.* 1250, 397–406. doi: 10.1016/S0531-5131(03)00967-1

Kumar, R., Bose, A., and Mallick, B. N. (2012). A mathematical model towards understanding the mechanism of neuronal regulation of wake-nrems-rems states. *PLoS ONE* 7:e42059. doi: 10.1371/journal.pone.0042059

Lam, Y.-W., Nelson, C. S., and Sherman, S. M. (2006). Mapping of the functional interconnections between thalamic reticular neurons using photostimulation. *J. Neurophysiol.* 96, 2593–2600. doi: 10.1152/jn.00555.2006

Landisman, C. E., Long, M. A., Beierlein, M., Deans, M. R., Paul, D. L., and Connors, B. W. (2002). Electrical synapses in the thalamic reticular nucleus. *J. Neurosci.* 22, 1002–1009. doi: 10.1523/JNEUROSCI.22-03-01002.2002

Langdon, A. J., Breakspear, M., and Coombes, S. (2012). Phase-locked cluster oscillations in periodically forced integrate-and-fire-or-burst neuronal populations. *Phys. Rev. E* 86:061903. doi: 10.1103/PhysRevE.86.061903

Lena, I., Parrot, S., Deschaux, O., Muffat-Joly, S., Sauvinet, V., Renaud, B., et al. (2005). Variations in extracellular levels of dopamine, noradrenaline, glutamate, and aspartate across the sleep–wake cycle in the medial prefrontal cortex and nucleus accumbens of freely moving rats. *J. Neurosci. Res.* 81, 891–899. doi: 10.1002/jnr.20602

Li, G., Henriquez, C. S., and Fröhlich, F. (2017). Unified thalamic model generates multiple distinct oscillations with state-dependent entrainment by stimulation. *PLoS Comput. Biol.* 13:e1005797. doi: 10.1371/journal.pcbi.1005797

Long, M. A., Landisman, C. E., and Connors, B. W. (2004). Small clusters of electrically coupled neurons generate synchronous rhythms in the thalamic reticular nucleus. *J. Neurosci.* 24, 341–349. doi: 10.1523/JNEUROSCI.3358-03.2004

Lopes da Silva, F., Hoeks, A., Smits, H., and Zetterberg, L. (1974). Model of brain rhythmic activity. *Biol. Cybernet.* 15, 27–37. doi: 10.1007/BF00270757

McCormick, D. A. (1989). Cholinergic and noradrenergic modulation of thalamocortical processing. *Trends Neurosci.* 12, 215–221. doi: 10.1016/0166-2236(89)90125-2

McCormick, D. A. (1992). Neurotransmitter actions in the thalamus and cerebral cortex and their role in neuromodulation of thalamocortical activity. *Prog. Neurobiol.* 39, 337–388. doi: 10.1016/0301-0082(92)90012-4

McCormick, D. A., and Prince, D. A. (1986). Acetylcholine induces burst firing in thalamic reticular neurones by activating a potassium conductance. *Nature* 319:402. doi: 10.1038/319402a0

Mesulam, M.-M. (2004). The cholinergic innervation of the human cerebral cortex. *Prog. Brain Res.* 145, 67–78. doi: 10.1016/S0079-6123(03)45004-8

Ni, K.-M., Hou, X.-J., Yang, C.-H., Dong, P., Li, Y., Zhang, Y., et al. (2016). Selectively driving cholinergic fibers optically in the thalamic reticular nucleus promotes sleep. *Elife* 5:e10382. doi: 10.7554/eLife.10382

Omori, T., and Horiguchi, T. (2004). Dynamical state transition by neuromodulation due to acetylcholine in neural network model for oscillatory phenomena in thalamus. *J. Phys. Soc. Japan* 73, 3489–3494. doi: 10.1143/JPSJ.73.3489

Pham, T., and Haas, J. S. (2018). Electrical synapses between inhibitory neurons shape the responses of principal neurons to transient inputs in the thalamus: a modeling study. *Sci. Rep.* 8:7763. doi: 10.1038/s41598-018-25956-x

Picciotto, M. R., Higley, M. J., and Mineur, Y. S. (2012). Acetylcholine as a neuromodulator: cholinergic signaling shapes nervous system function and behavior. *Neuron* 76, 116–129. doi: 10.1016/j.neuron.2012.08.036

Rasch, B., and Born, J. (2013). About sleep's role in memory. *Physiol. Rev.* 93, 681–766. doi: 10.1152/physrev.00032.2012

Rich, S., Zochowski, M., and Booth, V. (2018). Effects of neuromodulation on excitatory–inhibitory neural network dynamics depend on network connectivity structure. *J. Nonlinear Sci.* 8, 1–24. doi: 10.1007/s00332-017-9438-6

Robinson, P., Rennie, C., and Rowe, D. (2002). Dynamics of large-scale brain activity in normal arousal states and epileptic seizures. *Phys. Rev. E* 65:041924. doi: 10.1103/PhysRevE.65.041924

Rößler, A. (2010). Runge–kutta methods for the strong approximation of solutions of stochastic differential equations. *SIAM J. Num. Anal.* 48, 922–952. doi: 10.1137/09076636X

Roth, T. (2009). Slow wave sleep: does it matter? *J. Clin. Sleep Med.* 5(2 Suppl):S4.

Rudolph, M., Piwkowska, Z., Badoual, M., Bal, T., and Destexhe, A. (2004). A method to estimate synaptic conductances from membrane potential fluctuations. *J. Neurophysiol.* 91, 2884–2896. doi: 10.1152/jn.01223.2003

Sanchez-Vives, M. V., Bal, T., and McCormick, D. A. (1997). Inhibitory interactions between perigeniculate gabaergic neurons. *J. Neurosci.* 17, 8894–8908. doi: 10.1523/JNEUROSCI.17-22-08894.1997

See, A. R., and Liang, C.-K. (2011). "A study on sleep eeg using sample entropy and power spectrum analysis," in *2011 Defense Science Research Conference and Expo (DSR)* (Singapore: IEEE), 1–4. doi: 10.1109/DSR.2011.6026802

Sethuramanujam, S., McLaughlin, A. J., deRosenroll, G., Hoggarth, A., Schwab, D. J., and Awatramani, G. B. (2016). A central role for mixed acetylcholine/gaba transmission in direction coding in the retina. *Neuron* 90, 1243–1256. doi: 10.1016/j.neuron.2016.04.041

Shu, Y., and McCormick, D. A. (2002). Inhibitory interactions between ferret thalamic reticular neurons. *J. Neurophysiol.* 87, 2571–2576. doi: 10.1152/jn.00850.2001

Steriade, M. (2004). Acetylcholine systems and rhythmic activities during the waking–sleep cycle. *Prog. Brain Res.* 145, 179–196. doi: 10.1016/S0079-6123(03)45013-9

Steriade, M., Contreras, D., Dossi, R. C., and Nunez, A. (1993). The slow (<1 hz) oscillation in reticular thalamic and thalamocortical neurons: scenario of sleep rhythm generation in interacting thalamic and neocortical networks. *J. Neurosci.* 13, 3284–3299. doi: 10.1523/JNEUROSCI.13-08-03284.1993

Suffczynski, P., Kalitzin, S., and Lopes Da Silva, F. (2004). Dynamics of non-convulsive epileptic phenomena modeled by a bistable neuronal network. *Neuroscience* 126, 467–484. doi: 10.1016/j.neuroscience.2004.03.014

Weigenand, A., Costa, M. S., Ngo, H.-V. V., Claussen, J. C., and Martinetz, T. (2014). Characterization of k-complexes and slow wave activity in a neural mass model. *PLoS Comput. Biol.* 10:e1003923. doi: 10.1371/journal.pcbi.1003923

Wilson, H. R., and Cowan, J. D. (1973). A mathematical theory of the functional dynamics of cortical and thalamic nervous tissue. *Kybernetik* 13, 55–80. doi: 10.1007/BF00288786

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.

frontiers in Computational Neuroscience

METHODS
published: 21 February 2020
doi: 10.3389/fncom.2020.00013

ROOTS: An Algorithm to Generate Biologically Realistic Cortical Axons and an Application to Electroceutical Modeling

Clayton S. Bingham[1], Adam Mergenthal[2], Jean-Marie C. Bouteiller[2], Dong Song[2], Gianluca Lazzi[3] and Theodore W. Berger[2]*

[1] Department of Biomedical Engineering, Case Western Reserve University, Cleveland, OH, United States, [2] Department of Biomedical Engineering, University of Southern California, Los Angeles, CA, United States, [3] Department of Electrical Engineering, University of Southern California, Los Angeles, CA, United States

Advances in computation and neuronal modeling have enabled the study of entire neural tissue systems with an impressive degree of biological realism. These efforts have focused largely on modeling dendrites and somas while largely neglecting axons. The need for biologically realistic explicit axonal models is particularly clear for applications involving clinical and therapeutic electrical stimulation because axons are generally more excitable than other neuroanatomical subunits. While many modeling efforts can rely on existing repositories of reconstructed dendritic/somatic morphologies to study real cells or to estimate parameters for a generative model, such datasets for axons are scarce and incomplete. Those that do exist may still be insufficient to build accurate models because the increased geometric variability of axons demands a proportional increase in data. To address this need, a Ruled-Optimum Ordered Tree System (ROOTS) was developed that extends the capability of neuronal morphology generative methods to include highly branched cortical axon terminal arbors. Further, this study presents and explores a clear use-case for such models in the prediction of cortical tissue response to externally applied electric fields. The results presented herein comprise (i) a quantitative and qualitative analysis of the generative algorithm proposed, (ii) a comparison of generated fibers with those observed in histological studies, (iii) a study of the requisite spatial and morphological complexity of axonal arbors for accurate prediction of neuronal response to extracellular electrical stimulation, and (iv) an extracellular electrical stimulation strength–duration analysis to explore probable thresholds of excitation of the dentate perforant path under controlled conditions. ROOTS demonstrates a superior ability to capture biological realism in model fibers, allowing improved accuracy in predicting the impact that microscale structures and branching patterns have on spatiotemporal patterns of activity in the presence of extracellular electric fields.

Keywords: deep brain stimulation (DBS), axons, multi-scale, electrical stimulation (ES), morphology, spatio-temporal analysis

OPEN ACCESS

Edited by:
Sacha Jennifer van Albada, Jülich Research Centre, Germany

Reviewed by:
Henrik Lindén, University of Copenhagen, Denmark
Alexander D. Bird, Frankfurt Institute for Advanced Studies, Germany

****Correspondence:***
Clayton S. Bingham
csb119@case.edu;
clayton.bingham@gmail.com

Received: *27 September 2019*
Accepted: *31 January 2020*
Published: *21 February 2020*

Citation:
Bingham CS, Mergenthal A, Bouteiller J-MC, Song D, Lazzi G and Berger TW (2020) ROOTS: An Algorithm to Generate Biologically Realistic Cortical Axons and an Application to Electroceutical Modeling. Front. Comput. Neurosci. 14:13. doi: 10.3389/fncom.2020.00013

INTRODUCTION

In the study of extracellular electrical stimulation of neural systems, spatial and temporal patterns of activity are strongly influenced by tissue geometry. One established approach to studying this relationship is through morphologically detailed equivalent circuit models of neurons, including axons. While these models are invaluable for many different applications, they are especially useful for prediction of tissue response to extracellular stimulation, where explicit morphologies aid the prediction of activation thresholds under varying biological and stimulating conditions (Ranck, 1975; Nowak and Bullier, 1996, 1998). For biologically realistic network models, a common method involves arranging individual neuron models in virtual space to reconstruct elements of the tissue system being studied (Grill, 1999; McIntyre and Grill, 1999; Howell and McIntyre, 2016; Anderson et al., 2018; Bingham et al., 2018). This approach enables accurate prediction of membrane potentials in response to changes in electric field geometry and gradient (Clark and Plonsey, 1970; **Figure 1**).

Despite an understanding that geometry and topology influence activity, much of the biological realism in these studies is reserved for dendritic rather than axonal arbors. This lack of realism in axonal fibers becomes especially disconcerting when considering both that (i) central nervous system axon terminal arbors are often highly branched and tortuous relative to the mostly straight and long nerves of the periphery, and (ii) under most typical stimulating conditions, axons have shorter chronaxies than somas and dendrites (Ranck, 1975; Johnson and McIntyre, 2008; Rattay et al., 2012). It follows that suprathreshold stimulation events result in coupled local (driven by the injected electric field) and distal (synaptically driven) activity, with substantial realism being necessary to predict the spatiotemporal pattern of the resulting response in totality. Deliberate arrangement of neuronal structures is also useful for accurate model-based prediction of tissue–tissue interactions due to electric fields arising from endogenous current sources (Anastassiou and Koch, 2015). Accurate estimation of local field potentials (LFPs) and, therefore, predictions of the region-specific impact of ephaptic coupling are sensitive to the degree of biological realism implemented in a model system (Bingham et al., 2018). Lastly, network models that lack explicit axonal structures may have unrealistic conduction delays between connected populations of neurons, leading to potential prediction errors (Kim et al., 2019). While delays may be trivially added to network connections once they are known, biologically appropriate behaviors must first be estimated. Therefore, geometrical and anatomical realism may also be necessary to study emergent network activity such as co-oscillatory activity in hippocampal networks (Whittington et al., 1997; Fries, 2005).

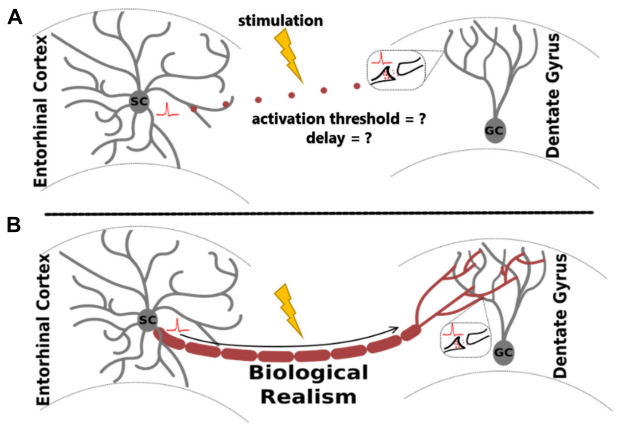

FIGURE 1 | Adding biological realism to axon models for the study of extracellular electrical stimulation allows more accurate analysis of evoked neural network activity. The proposed algorithm, ROOTS, was developed specifically to provide increased realism in fiber models. The utility of ROOTS is to mitigate the challenge portrayed above: without explicit axonal reconstructions, how does one accurately estimate the site of action potential initiation and appropriate orthodromic conduction latencies through the terminal region in the presence of externally applied electric fields? **(A)** corresponds to the case where distance-based pure-delay mechanisms are used. **(B)** corresponds to explicit and biologically accurate axon representations.

Despite the long record of hippocampal observation, axonal morphology is not as scrupulously described as the somas and dendrites of many cell types. With a few exceptions, studies yielding explicit reconstructions through morphometric analysis of neuronal branching have focused on dendritic arbors and overlooked their axons (Desmond and Levy, 1985; Hama et al., 1989; Claiborne et al., 1990). The dearth of robust datasets is exacerbated by the general observation from staining experiments performed in the hippocampus that axonal structures, even of the same cell type, may be less stereotyped than dendritic arbors (Hjorth-Simonsen and Jeune, 1972). Perforant path axon terminal arbors from layer 2/3 entorhinal cortical (EC) spiny stellate cells are not constrained to simple conical, fanned, or star-shaped volumes like so many dendritic arbors (Tamamaki and Nojyo, 1993). Geometric and topological heterogeneity make the prospect of using explicit reconstructions unfeasible for direct use in computational models which require *in situ* cell density at tissue scale. This becomes particularly apparent when considering that the unique geometry of the dentate gyrus, which changes from septal to temporal poles, requires dramatic inter-region variety in the volume, orientation, and contour of afferent EC axons.

Despite the general absence of morphometrics for axons, the shape of terminal fields, distribution of synaptic spines, and a rough measure of anatomic domains of axon terminal fields from either histology or other imaging methods provide data from which minimally functional axons can be grown. Explicit dendritic reconstructions, spine counting, and anterograde and retrograde staining experiments provide information regarding synaptic targets. Distributions of synaptic targets combined with a knowledge of the general path and origin of a fiber is sufficient to generate a functional structure that captures the tertiary conformation and local divergence of the axon terminal arbor. When representing a neural process as a graph, nodes placed at synapses and edges to connect between them and the soma can effectively reconstruct a functional dendritic arbor (Türetken et al., 2011). Likewise, presynaptic boutons provide nodes that

can be connected to each other and the soma with a deliberate arrangement of edges to form a spatial network or graph. In other words, a minimally functional axon connects a parent cell to synaptic targets.

Beyond minimally functional morphologies, sophisticated models of realistic neuronal branching have been proposed with primary application to either generating unique and artificial dendritic trees or reconstructing them from a series of images. There are two chief types of generative models: (i) stochastic (Rozenberg and Salomaa, 1980) or (ii) greedy (Cuntz et al., 2007, 2010; Budd et al., 2010; Budd and Kisvarday, 2012). Stochastic models operate by sampling distributions of branching statistics extracted from experimentally measured neurons, while greedy graph-based models are much more frequently used to reconstruct three-dimensional dendritic trees from stacks of manually labeled images (Türetken et al., 2011). Although each of these approaches have been useful under certain conditions, neither has been properly adapted for use in the generation of virtual biologically realistic axons.

Consequently, we present in this paper a new graph-based algorithm for generating biologically realistic tree representations of axon terminal arbors. The proposed model is inspected for its utility in studying extracellular electrical stimulation of cortical tissue through analysis of the impact of arbor topography and morphometry on activation thresholds and, by extension, spatiotemporal patterns of activity in the hippocampus. The results of this work comprise (i) a quantitative analysis of the generative algorithm proposed, (ii) presentation and quantitative/qualitative description of generated fibers, (iii) comparison to leading alternative methods, (iv) demonstration of a method to reduce spatial complexity of axonal arbors while maintaining accurate prediction of neuronal response to extracellular electrical stimulation, and (v) an extracellular electrical stimulation strength–duration study. The value of these studies is twofold: (i) establishing the novelty and utility that this modeling system yields and (ii) determining if stimulation–response recruitment order (i.e., large before small) for straight, long, large-diameter, and myelinated peripheral fibers is similarly true of small, highly branched, and unmyelinated cortical fibers.

MATERIALS AND METHODS

The study presented here focuses on accurately capturing the emergent spatial features of spiny stellate EC axons within the dentate gyrus in Sprague-Dawley rats. The model is designed to be flexible to the inclusion of novel morphometric criteria as new experimental data become available but, at present, it is clearly important that (i) fibers are constrained to 1–1.5 mm within the septotemporal axis (Tamamaki and Nojyo, 1993), (ii) laminar organization along the transverse axis is inviolate, with medial and lateral EC axons confined to the middle and outer thirds of the dentate molecular layer, respectively, and (iii) axons synapse with a pronounced *en passant* connective schema, where most pre-synaptic boutons are non-terminal. Many more features and their sources are detailed in **Table 1**. A greedy, graph-based, or ruled-optimum ordered tree system (ROOTS) was developed to control these features. First, we will explain the development of constraints and inputs to the method and then explain, in detail the functions of the method.

TABLE 1 | Principle features **(left)** of entorhinal cortical axons found in the dentate gyrus perforant path and the studies which reported them **(right)**.

Feature	References
Strictly laminar dentate perforant path	Hjorth-Simonsen and Jeune, 1972; Witter, 2007
En passant; mostly non-terminal boutons	Witter, 2007; Bindocci et al., 2017
Distribution of bifurcation angles $\approx 80.3 \pm 35.7°$	Budd and Kisvarday, 2012
0.1 μm fiber diameter, ≈0.7 μm boutons	Tamamaki and Nojyo, 1993
Primary bifurcation at/near crest, envelopes entire transverse of dentate, continues to CA3	Hjorth-Simonsen and Jeune, 1972; Schwartz and Coleman, 1981; Tamamaki and Nojyo, 1993; Witter, 2007
≈17,700 synapses per EC axon	Desmond and Levy, 1985; Hama et al., 1989; Claiborne et al., 1990
Myelination – mixed, though clearest images show no myelination below the crest of dentate	Tamamaki and Nojyo, 1993

These features, when combined with the topography of spines in the outer and middle thirds of the dentate gyrus provide guiding parameters for construction of artificial tree models. The proposed method utilized these as criteria for generation of dentate perforant path fibers that provide a test-case for ROOTS as a method.

Volume and Nodal Constraints

In the construction of a spatial ordered tree, the number and topography of target nodes strongly determines the emergent features of the resulting graph. Therefore, the selection of nodes is an important step in the successful generation of biologically appropriate axonal trees. This process was executed using the Kjoenigsen rat hippocampal atlas slice at −3.34 from Bregma to segment model boundaries (Kjonigsen et al., 2008). Past efforts have elucidated the approximate number and spatial distribution of synaptic targets within the dentate perforant path (Claiborne et al., 1990; Bingham et al., 2016; Hendrickson et al., 2016). The approximate number and density of granule cells found within the dentate region of a 1.5 mm extruded slice was calculated based on density measurements reported in the literature (Gaarskjaer, 1978; Patton and McNaughton, 1995). Spine counts and the laminar topology of entorhinal–dentate connections were used to create a pool of synaptic targets and the number of perforant path arbors from which afferent connections might be formed (Desmond and Levy, 1985; Hama et al., 1989; Claiborne et al., 1990). An abbreviated table derived from Hendrickson et al. (2016) can be found below to summarize these parameters as they were used in this study:

To summarize: the number of synaptic spines in the outer and middle third of hippocampal granule cell dendritic arbors the size, number, and density of granule cells, and the number/density of EC cells contributing to the perforant path provide the necessary arithmetic for deducing the number of synapses made within each axon terminal field (**Table 2** and **Figure 2**). When combined with observational

TABLE 2 | Parameters describing the entorhinal cortical–dentate gyrus topology used to design the topography of synaptic targets for axon fiber growing described in later sections.

Entorhinal cortical–dentate gyrus topological parameters	
Granule cell # spines: middle 1/3:	1050–1200
Granule cell # spines: outer 1/3:	1100–1300

This table provides a range of spine-counts per dentate granule cell with which the perforant path fibers may synapse (Claiborne et al., 1990; Desmond and Levy, 1985). Medial entorhinal cortical fibers synapse in the middle 1/3 and lateral entorhinal cortical fibers synapse in the outer third of the granule cell dendritic arbors.

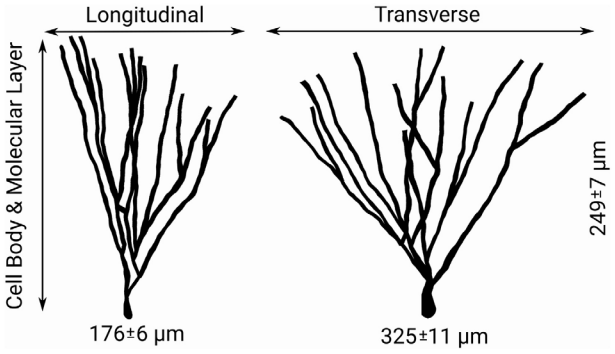

FIGURE 2 | Features of granule cell arbors – including the distributions of longitudinal, transverse, and dentate normal lengths of arbors shown above – and afferent connectivity provides guiding information for both the sampling of target nodes for generation of axon morphologies and the simplification of generated morphologies (Claiborne et al., 1990).

data regarding the septotemporal range of these axons, this information provided an approximate volume throughout which nodes could be distributed and a plausible synthetic terminal arbor could be grown. These same parameters were used to construct much larger, more complex, and previously validated mechanistic models of a rat dentate hippocampus; therefore, additional details can be found in Bingham et al. (2018) and Hendrickson et al. (2016).

Constraining Patterns of Axon Branching

While many of the possible branching patterns of a functional arbor are constrained by the topography of synaptic targets, there remain as many as $n^{(n-2)}$ trees that span a set of targets (n), few of which are biologically plausible (Cayley, 1889). It follows that encouraging biological realism requires additional non-trivial steps to constrain branching features of generated topologies. **Figure 3** is an algorithm flow diagram and pseudocode (provided in **Supplementary Figure S1**) for proper preprocessing and successful execution of ROOTS. This process generated the trees analyzed in later sections of this manuscript. In brief, ROOTS seeks to minimize the quantity of membrane required to span a set of synaptic targets while satisfying user-specified branching criteria. These criteria (at the time of writing) include: branch extension angle (meander) and length ("Extension Criteria"), and bifurcation angle and length ("Bifurcation Criteria"). The method was also designed in a manner that allows additional global criteria (branch order, number of bifurcations, total length, etc.) to be designed and applied within ROOTS. The process by which this is accomplished involves serially considering sorted (by source–target distance or "Likely Path"; according to **Figure 3**) open points by alternating between branch extension (appending points to an existing branch) and bifurcation (beginning of a new branch). If in the process of extending a branch it is found that no points satisfy branch extension criteria, then the algorithm switches to bifurcation. If a bifurcation can be created according to bifurcation criteria, then the algorithm switches back to extending the newly begun branch. This iterative process continues until either extension and bifurcation criteria cannot be satisfied or no open points remain. Model inputs (synaptic targets, branching and bifurcating criteria, and global criterial) dually exert control over the emergent spatial/geometric features of the entire terminal field and the branching patterns that develop as the fiber is constructed.

While the core of this algorithm is capable of growing axons with highly particular geometries, it is limited in the cases where axons execute acute turns without forming connections that cross the resulting sulcus. An additional rule can be used to mitigate this case-specific flaw: a relative location sensitive dynamic source point and reference angle (**Figure 3B**). The principle difference between simplified and dynamic source updating modes (**Figures 3A** vs. **3B**) is the way in which points are sorted and, therefore, the order in which points are considered for inclusion into the tree and the reference angles that are subsequently calculated. The "likely path sort" components in **Figure 3B**, effectively, allow fibers to be grown along manifold surfaces where the simplified algorithm is for efficient growth of conical or star-shape fibers. This approach requires additional preprocessing – replacing the initial sorting of points with respect to a single source point with a more sophisticated sort and a dictionary of relative source-points and reference directions. This new process is performed by (i) finding spatial clusters of synaptic targets using K-means, (ii) fitting a mesh to the spatial clusters using Delaunay triangulation, and (iii) discovering the most likely path to any cluster center from the origin source-point through the constructed mesh (Hartigan and Wong, 1979; Chew, 1989). The most likely path, found using Dijkstra's shortest path algorithm, is then used as a lookup table where the edges leading to the cluster within which any new point being considered may be found are used to constrain angles of branch extension and bifurcation (Chen, 2003). Execution of these three new steps results in a path directed sorting of targets. It should be noted that K-means is the least critical new component and is only used to regularize the triangulation and reduce the complexity for the steps that follow (between 2500 and 3000 clusters were used in this implementation to have the desired effect). This new rule allows successful execution of acute turns, where the fiber bends backward toward the source point without crossing the resulting cleft or forming any cycles. Critically, these acute turns are executed without relaxing constraining parameters for branching and bifurcating.

While there are many clustering algorithms and many sophisticated meshing algorithms, K-means and Delaunay algorithms were selected because of their speed and reliability.

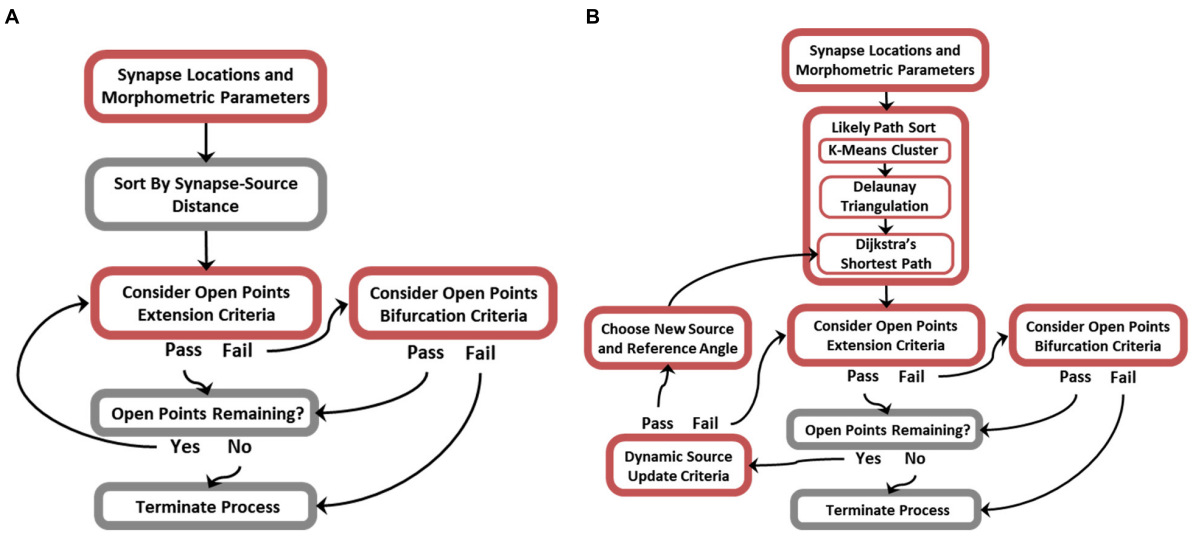

FIGURE 3 | Flow diagrams describing the ROOTS algorithm. For greater efficiency, the generative algorithm was designed to operate in one of two modes: **(A)** Without or **(B)** with dynamic source-point updating criteria. This feature enables generated morphologies to successfully conform to acute turns of varying radius. Disabling this feature when not needed improves efficiency.

Because dynamic sourcing (**Figures 3B** vs. **3A**) adds considerable computational burden, it is valuable to select preprocessing methods which do not add to this burden needlessly. **Supplementary Figures S2–S4** show development of this stage of the algorithm and its three constituents: K-means clustering, Delaunay triangulation, and Dijkstra's shortest path methods.

Synaptic Boutons

In recognition that axons have complex surfaces and are not just a series of smooth and simple pipes, additional functions were written to allow complexities such as synaptic boutons to be added to an already grown fiber. These boutons, as they are seen in the dentate perforant path at non-terminal presynaptic densities, are described by Tamamaki as "periodic varicosities" (Tamamaki and Nojyo, 1993). Each bouton is ≈5 μm long, ≈0.7 μm in diameter, with varying distances between, depending on the topography of the synaptic targets. Because neither the exact distribution of inter-bouton distances nor multisynapse formation behavior is known in this tissue system, it was assumed that they were uniformly distributed throughout the terminal axon arbors every 25 μm, with each bouton being 5 μm in length and 0.7 μm in diameter. Because these boutons are non-terminal they are likely to be actively conducting. Lacking experimental evidence to frustrate this assumption, sodium, calcium, and potassium channel densities and conductances were implemented with the same parameters in bouton compartments as in non-bouton axonal regions.

Following development and testing of this algorithm, fibers were exported for simulation in environments such as NEURON (Hines and Carnevale, 2003). Later in this paper, extracellular electrical stimulation studies performed with the generated morphologies are presented to demonstrate the maturity of this analysis pipeline.

Arbor Simplification and Computational Complexity

Simulation of detailed neuronal models is computationally expensive. While this study was enabled by non-competitive access to a 4,040 processors computing cluster, there remain concerns about impractical and unnecessary model complexities. With as many as 17,700 possible synaptic connections made by each EC arbor, it became clear that an approach to morphology simplification would be necessary to reduce the computational burden of both generating and simulating arbors in NEURON. While the exact number is not known, because of the *en passant* nature of the perforant path fibers it is likely that a passing fiber synapses more than once with any target granule cell. This provides the opportunity to generate trees using fewer nodes due to the relative co-locality of synaptic connections between an arbor and any given granule cell. To examine this assumption, we generated arbors with 8,850 (two synapses per target cell) or 5,900 (three synapses per target cell) target cells rather than 17,700 target cells and used a single node from each to guide arbor growth. Perforant path fibers were generated using each of these node counts and then each was also line simplified. The Ramer–Douglas–Peucker (RDP) algorithm for line simplification was used on each of these trees to determine the minimum node count required to approximate full complexity fibers (Saalfeld, 1999). Using a fraction of the typical dentate granule cell dendritic arbor height and width (e.g., 20% or ∼45 μm) to set a maximum RDP-epsilon ensured that path deviations would be much less likely to make otherwise probable EC–DG connections anatomically impossible.

Alternative Generative Models

To test the functionality of this algorithm in comparison with others that already exist, we attempted to create satisfactory

arbors with two commonly used alternative tools which were originally designed to grow dendritic arbors: the TREES Matlab toolbox (Cuntz et al., 2010), and L-NEURON (Ascoli and Krichmar, 2000; Scorcioni and Ascoli, 2005).

The TREES toolbox, like ROOTS, depends upon carefully selected nodes or points to grow a graph; therefore, spanning trees generated by this method can use the same set of synaptic or cellular targets as those utilized by the proposed system. Unlike ROOTS, however, the only parameter that can be adjusted to improve the performance of the resulting morphology is the balancing factor (BF). This BF represents the weighting of priority for path length versus conductive delay in the spanning tree that is generated. To study this approach, a new topology grown via TREES utilizing the same control points as a ROOTS fiber. The BF was calibrated by minimizing a multi-objective function (MOF), formulated as an unweighted sum of independently normalized mean root square error for each of the following morphometrics: Euclidean distance/path-length (BF), branch order, bifurcation angle, and total path length. The BF of the arbor generated by TREES was set at the value that minimized the difference between the MOF scores of the arbors generated by each system. To maximize similarity between these two morphologies through calibration of the TREES BF, direct (Euclidean) vs. path length ratios were calculated for each carrier node, then a histogram of these data was fit with a kernel density estimate (KDE via Gaussian smoothing). This process was repeated for path length, branching order, branch length, and branching angle. The KDEs for direct vs. path length and branching angle for the TREES axon were subtracted from those for the ROOTS fiber. These differences were independently normalized and the RMSE was calculated. Each of these metrics were summed without any weighting. The TREES BF was then calibrated through minimizing the summed normalized root mean-squared error of these differences (U) according to Eq. 1.

$$U = \mathrm{dBF} + \mathrm{dBO} + \mathrm{dPL} + \mathrm{dBL} + \mathrm{dBA} \tag{1}$$

where each term represents the normalized summed difference of BF (dBF), branching order (dBO), total path length (dPL), branch length (dBL), and bifurcation angle (dBA). Unlike TREES and the proposed algorithm, L-NEURON doesn't rely on a preselection of target nodes to construct a topology, but rather depends upon measures of branching structure of a tree, or morphometrics. Fitted distributions to these measurements are then stochastically sampled to grow a tree. This approach is intended to capture branching patterns without much regard for the emergent spatial features of a tree. Because of the lack of an extensive database of EC axons from which to take branching measurements, we resorted to using the companion tool L-MEASURE to extract morphometrics from one of our own generated arbors to gauge the feasibility of using stochastic methods to generate axonal arbors when such a dataset does become available (Scorcioni et al., 2008). Extracted morphometrics were then fed to L-NEURON to generate a morphometrically equivalent arbor.

Virtual arbors were ultimately evaluated based on their ability to capture known spatial features of EC axon terminal fields, including a complex geometry which conforms to the contours of the dentate gyrus.

Strength–Duration Relationship in Response to Extracellular Electrical Stimuli

To understand how fiber geometry in the hippocampus gives rise to patterns of activity, fibers with varying patterns of diameter were simulated in response to a range of current source–arbor distances and stimulus amplitudes. Images of spiny stellate fibers from Tamamaki and Nojyo (1993) and more evidence from Bindocci et al. (2017) show continuous fibers with "periodic varicosities," or non-terminal synaptic boutons on the *en passant* fibers. Because fiber diameter has been shown to influence both conduction velocity and excitability, it was important to explore the response characteristics of fibers with this sub-micron variation in diameter (Clark and Plonsey, 1970). A fiber was generated by ROOTS and instantiated with one of three patterns of diameter in the simulation environment, NEURON 7.6.2, so that its behavior could be simulated (Hines and Carnevale, 2003). Hodgkin–Huxley membrane biophysics under *in vivo* temperature conditions were inserted in all compartments and d-lambda rules were used to determine appropriate space constants for compartmentalization of the fibers (Hines and Carnevale, 2003). All other biophysical features were borrowed from nodal biophysics described in Johnson and McIntyre (2008). An itemized table of biophysical properties can be found in **Supplementary Figure S5**. All morphological features for these fibers, other than the deliberate variations in diameter, remained as presented in **Table 1**.

Monopolar point-source stimuli were used in all stimulation experiments presented in this article. Electrodes were placed in one of two locations near the primary bifurcation of the perforant path at distances of 100 and 500 μm from the nearest neuronal compartment.

Following model construction, two sets of analysis were performed. The first comprised a strength–duration study of complex arbors. Square-wave pulses of anodic or cathodic charge polarities (+ and −, respectively) with widths between 0.025 and 1.4 ms were delivered at each of the two distances. The extracellular voltage throughout the model space was estimated using an analog of Coulomb's law with material resistivity of 3.8 Ω-m (Eqs 2 and 3) (Clark and Plonsey, 1970; Holt and Koch, 1999; Bingham et al., 2018).

$$\phi(x, y, z) = \frac{I_0}{4\pi^* \sigma_i^* r_i} \tag{2}$$

where Φ is the field potential resulting from a current source, I. The conductance, σ, is the inverse of resistivity. Radial distance, r, is found by Eq. 3:

$$r_i = \sqrt{(x_i - x_0)^2 + (y_i - y_0)^2 + (z_i - z_0)^2} \tag{3}$$

By adding new voltage sources in series with the circuital elements representing each section of membrane, extracellular potentials calculated via Eqs 2 and 3 were applied to

neuronal compartments within the NEURON model using the "extracellular" mechanism. Stimulation was delivered over a range of current amplitudes designed to cover, at a minimum, rheobase to twice rheobase for each set of stimulating conditions.

The second analysis examined the same models and stimulating setup but focused on the temporal distribution of the response at rheobase. The time it took each compartment in the fiber to reach action potentiation was recorded, used to populate a histogram, and then Gaussian smoothed to present a KDE. These plots were used to identify the impact of boutons on conduction velocity throughout the complex arbor.

Comparison With Implied Axon Conduction Latency Estimates

In order to demonstrate the differences between the relatively sophisticated axonal models grown via ROOTS and more simple axon representations, a simple distance-based delay was estimated based on the average conductance velocities of the ROOTS fibers and visualized using KDEs after the manner previously described. The distance-based conduction delay mechanism just explained is described in figures and relevant results sections as a "pure-delay" mechanism because no explicit cable model is used to approximate the delay. To clearly demonstrate the differences that are not readily seen in the KDE, pair-wise differences (residuals) of compartmental APs using ROOTS and the pure-delay mechanism were measured and plotted.

Data/Model Sharing

It is important to us that the method by which functional cortical axons were generated be highly accessible to other members of the computational neuroscience community, especially those studying neuromodulation for the treatment of diverse neurological disorders. Therefore, the graph-based algorithm used in this study has been compiled in a user-friendly manner and distributed to the Python Package Index under the name *Neural Roots* (*Roots*)[1,2] along with concise documentation. Neural Roots was written with Python 3.6.9 and has a few non-standard dependencies, including: SciPy, NetworkX (Hagberg et al., 2005), Mayavi (Ramachandran, 2001), Shapely, and Pandas (McKinney, 2010) (searchable in PyPi, the Python Package Index). While some backward compatibility is likely, this has not been extensively tested. The model is also being prepared for submission to ModelDB where it will be available alongside other elements presented in this paper.

RESULTS

Results comprise three parts: (i) a qualitative and quantitative assessment of the proposed algorithm, (ii) a comparative analysis with alternative methods, and (iii) the presentation of strength–duration curves and an analysis of EC fiber recruitment order and temporal distribution of the response to extracellular stimulation.

[1]https://github.com/bingsome/roots

[2]https://pypi.org/project/Roots/

Algorithm Assessment and Validation

There are two chief loops within the core ROOTS algorithm: branch extension and bifurcation. As the algorithm proceeds, the complexity of branch extension decreases while that of bifurcation increases, this can be seen in **Figure 4** as a linear/slightly supra-linear shape of each individual line. Bifurcation criteria become more difficult to satisfy as the axon graph becomes more complete, resulting in exponential increases in time-to-completion when larger and larger axons (more nodes to connect) are generated. Many aspects of these trends are dependent upon the spatial topography of the nodes themselves. For example, if nodes fall along a straight enough line, a single branch extension loop with no bifurcations will complete the axon graph and the time-to-completion will be perfectly linear with respect to the number of nodes.

Arbor Simplification and Computational Complexity

Consideration of the complexity of the resulting fiber is essential because simulation of just a few of these fibers at full complexity could be highly taxing on a single-processor computer. **Supplementary Figure S6** demonstrates the relationship between simulation efficiency and fiber complexity using the NEURON engine. Simulation efficiency is reported as the ratio of clock-time to simulated time. As fiber complexity increases, the amount of processing time (clock-time) required to complete an otherwise controlled simulation also increases linearly. Data for **Supplementary Figure S6** were collected on an Acer Aspire TC-885-UR17 desktop computer.

Ramer–Douglas–Peucker line simplification was performed on axons with two or three synapses per target granule cell (8,850 or 5,900 nodes). **Figure 5** demonstrates that axons with as few as 2,944 and 2,097 target nodes can be used to approximate the behavior of an arbor with maximum complexity, 8,850 and 5,900 target granule cells, respectively. These simplifications can be made while allowing, at most, 38–42 μm deviations from original contours and <5% reduction in total path length.

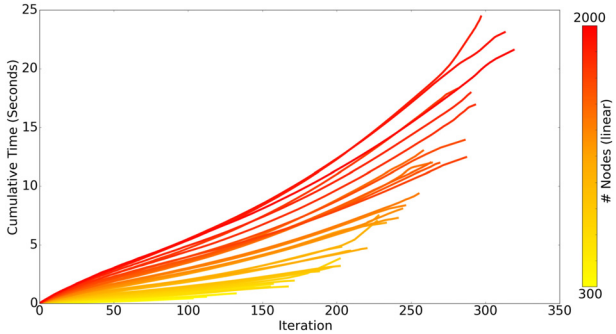

FIGURE 4 | Each line represents the time-course of ROOTS execution to generate an axonal graph with dentate perforant path topography using different numbers of carrier-points, ranging from 300 to 2000. Each iteration of the algorithm includes one loop for branch extension and another for bifurcation. While the above plot presents performance while growing a specific type of arbor, method performance is highly dependent upon input parameters and may be faster or slower for other model types.

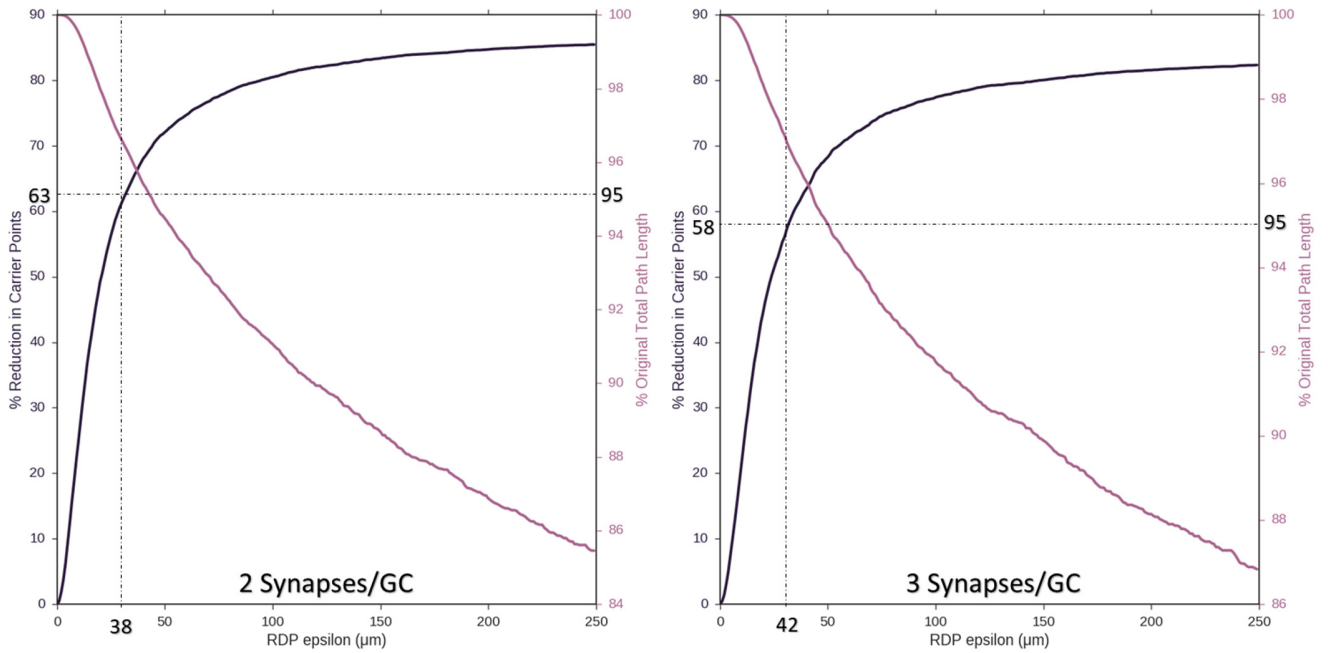

FIGURE 5 | As an approach to reducing computational complexity, Ramer–Douglas–Peucker (RDP) line simplification was studied over a range of epsilon (the maximum distance a simplified curve can be removed from an the original), for reduction in node count (**Left** *y*-axis) and change in total path length (**Right** *y*-axis). Simplification was performed for the cases where two or three synapses were made with each targeted cell (~8,850 or 5,900 starting nodes, left and right plots). Above 95% path length is preserved by setting epsilon to 38 and 42 μm, resulting in reduction of required number of nodes to 2,944 and 2,097, respectively (~60%).

Example graph models of EC axon terminal fields which incorporate these simplifications generated by the proposed algorithm are presented in **Figure 6**. The algorithm described in **Figure 3B** yielded fibers that captured the known features of layer 2/3 EC spiny stellate axons which make up the dentate perforant path (**Table 1**). These features include but were not limited to: distribution of bifurcation angles of approximately 80±34°; a septal–temporal range of between 1 and 1.5 mm; laminar organization with the MEC and LEC in the outer and middle thirds of the dentate molecular layers, respectively; saturation of branching order at a reasonable level to encourage *en passant* synapsing; presentation of both DG and CA3 terminal fields with the preservation of a fissure between the two fields; and finally, the tree structure passed through a plausible topography of DG and CA3 arbor domains which enables *in situ* levels of connectivity.

Alternative Generative Models

A morphology generated via our own graph-based algorithm (**Figure 3B**) was selected and morphometrically described via L-Measure and then, using the stochastic system called L-Neuron, attempted to regenerate a morphometric equivalent. This exercise failed to return a morphology which could conform to the topography of the molecular layer of the dentate gyrus. The chief reason being that purely stochastic methods are unable to prevent excursion of fibers beyond natural boundaries in the volume of tissue system being modeled. The general stochastic method is, therefore, unsuitable for generation of anatomically appropriate

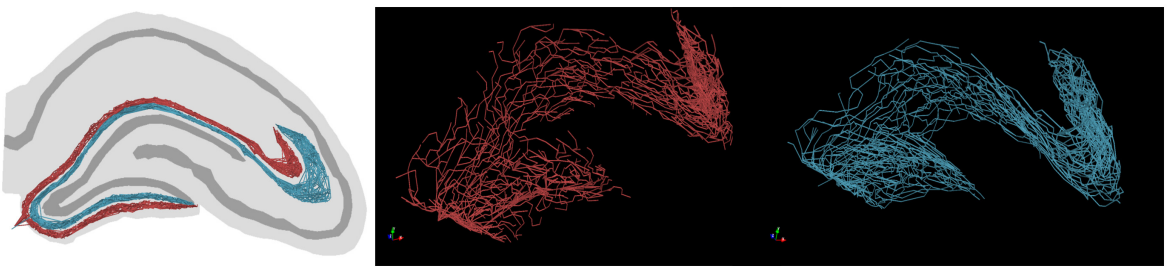

FIGURE 6 | Example graph models (red, LEC; teal, MEC) of entorhinal cortical axon terminal fields, generated by the proposed algorithm. These axons have features expected in *in situ* fibers, namely: *en passant* terminal topography, laminar architecture (i.e., MEC/LEC are spatially segregated within the dentate molecular layer), and fibers proceed to the ends of each supra/infrapyramidal blade of the dentate gyrus and proceed to a terminal field in CA3/2/1.

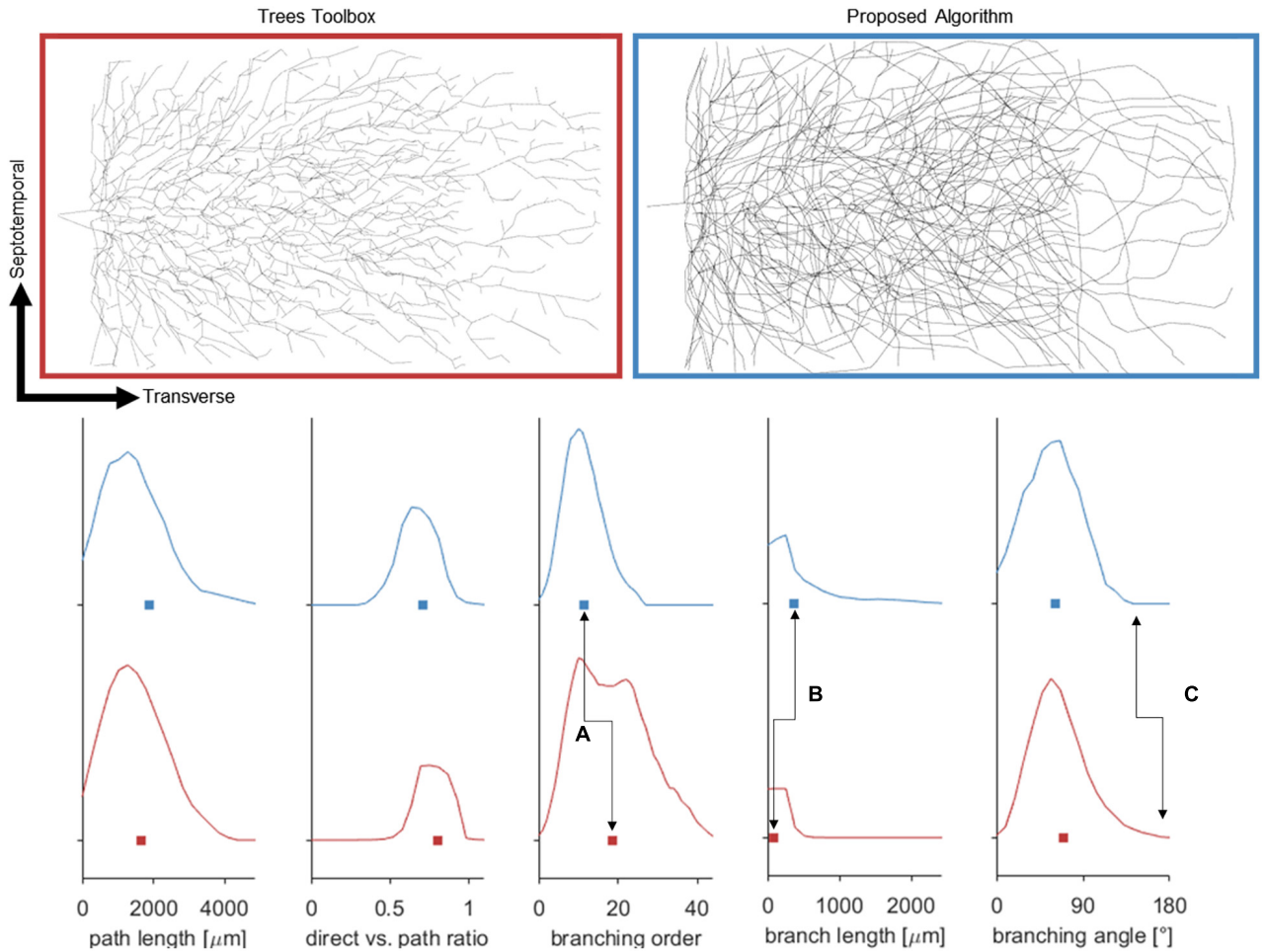

FIGURE 7 | (Top) Dorsal aspect of the enclosed blade (suprapyramidal) portion of an entorhinal cortical axon generated by one of two methods: Trees Matlab toolbox (BF = 5), and the proposed algorithm. Balancing factor, and connection threshold for the Trees toolbox were selected to align these distributions aligned favorably. **(Bottom)** Fundamental statistical distributions comparing the two methods in terms of nodal path length, direct distance vs. path distance, branching order, branch lengths, and bifurcation angles. The plotted curves represent probability densities and the box represents the mean. Despite matching many features of the proposed algorithm, the Trees generated arbor does not capture the *en passant* nature of the perforant path [seen in branching order (A), and length (B)] and has lingering bifurcations which have extreme angles (C).

axonal morphologies without further modification to constrain emergent spatial characteristics.

Attempts were also made to generate accurate morphologies using the TREES MatLab toolbox. These efforts yielded trees superior to those generated by the L-Neuron method, though still deficient in important ways. An axon generated by TREES was selected for quantitative and qualitative comparison with that yielded by earlier analysis using the ROOTS system (Arbor Simplification and Computational Complexity). Each tree was constructed using the same set of target nodes; therefore, all differences in patterns of branching arise from differences between the two algorithms. The TREES BF was then calibrated through minimizing a MOF, attempting to find the best possible match between TREES and ROOTS arbors. This process yielded a functional axon from the TREES algorithm, though differences remained that could not be resolved through manipulation of the BF alone (**Figure 7**). Despite a high degree of similarity in direct (Euclidean) vs. path length, total path length, and branching angle distributions, TREES resulted in a higher than expected number of terminal branchlets. This shifted the branching order and branch length distributions to the right and left, respectively. These shifts demonstrate difficulty for the TREES approach in appropriately capturing the *en passant* connection schema which typifies EC axons of the perforant path, where terminal branchlets are reportedly rare. **Figure 8** (branch length vs. branch order) further highlights this conformational difference between the two fibers and provides a useful comparison to experimental measurements. In this figure, the black line denotes the branch order at which 99% of the total path length was achieved in layer two spiny stellate axons reported by Budd et al. (2010). While the comparison with Budd's report is useful, it is made with hesitancy because of the variability of branching that occurs within the long stretch of fiber between the entorhinal cortex and the extensively branching terminal arbor. The difficulty TREES has in controlling

branch order arises from the paradigm of control implemented in the system. More specifically, the balancing of total path length versus conduction time (BF) does not provide for fine control of the shape and branching properties of an arbor within a volume. The significance of this limitation of TREES with respect to ROOTS is particularly clear when considering the implications for extracellular stimulation where it is thought that terminations and thorough-fare fibers behave differently in the presence of extracellular electric fields (Joucla and Yvert, 2009).

An additional challenge not satisfied by the TREES approach is that of performing acute turns. A portion of the perforant path innervates the CA3/CA2 region of the hippocampus and therefore must be able to extend from the end of the enclosed blade of the dentate into these other domains. The algorithm presented in **Figure 3B** (dynamic source) describes how updating reference angles allow the fiber to bend around complex anatomies successfully without violating morphometric constraints (**Figures 6**, **9**). An attempt to replicate this feat using the TREES Matlab toolbox is presented in **Supplementary Figure S7**.

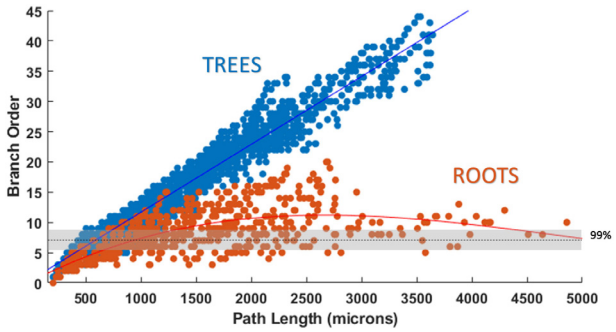

FIGURE 8 | Plotting path length vs. branch order makes it clear that ROOTS more accurately captures the *en passant* nature of the perforant path than does TREES. Each dot represents the branch order of a NEURON section. 99% of total axon path length of entorhinal cortical spiny stellate cells should be achieved with fiber of branch order no greater than ≈7 (Budd et al., 2010). Increased levels of early terminations in TREES is a geometric challenge that may have significant implications for extracellular stimulation (Joucla and Yvert, 2009).

Strength–Duration Relationship in Response to Extracellular Electrical Stimuli

To determine the importance of including boutons when simulating extracellular electrical stimulation of axon models, strength–duration curves for three diameter patterns of a ROOTS arbor (generated via the algorithm described in **Figure 3B**), where diameters correspond to bouton and inter-bouton sizes (according to **Figure 9**), were estimated and presented in **Figure 10**. When stimulating with anodal pulses, these results agreed with previous reports that larger diameter fibers have shorter chronaxies than small, and that this difference is exaggerated by large electrode-fiber distances and longer stimulation pulses. For cathodal impulses of such small and highly branched fibers, the model yielded negligible opportunity for selective activation of fibers by diameter. This protocol was repeated for a biologically realistic arbor, complete with boutons. The recruitment pattern of the boutoned fiber was no different in a cathodal field but under anodal conditions had activation thresholds between those of the 0.7 and 0.1 μm uniform diameter arbors. The differences in threshold between the boutoned fiber and the 0.7 μm fiber decreased with increasing electrode–fiber distance. These differences dissolved as the pulse-width approached 700 μs.

The impact of boutons on temporal dynamics are presented in **Figure 11**. The time of first action potential is plotted along the horizontal axis for each neuronal compartment in the biologically realistic arbor with the vertical axis representing probability density (frequency). A fiber with each of three patterns of fiber diameters (uniform 0.1, 0.7, or boutoned) were simulated in response to 1 ms stimuli at rheobase amplitude. Activity was initiated near the electrode and then actively conducted to other compartments, including those beyond the effective volume of the extracellular electric field where artifact voltage dropped to tens of microvolts. In **Figure 11** are plotted, as a KDE, the time-course of threshold activity of all compartments in each case. As the average fiber diameter is increased, the shape of the KDE shifts to the left, indicating that activity in the arbor is occurring with less delay following stimulation. This shift is explained by the increased conduction velocity throughout the arbor which results from changing compartment diameters. Average conduction velocities were 0.88, 0.53, and 0.12 m/s in the 0.7 μm, boutoned, and 0.1 μm fibers, respectively. The boutoned fiber nearly mirrored temporal patterns of activity of the uniform 0.7 μm fiber, except at very close electrode–fiber distances.

Comparison With Implied Axon Conduction Latency Estimates

When compared with a pure-delay mechanism, as in **Figure 12** left, ROOTS fibers exhibited a smoother distribution of compartmental Aps, though the KDEs for the pure-delay mechanism had similar shape and length. When exploring the impact of varying the conductance velocity on pair-wise compartmental AP latencies between the pure-delay mechanism and each of the ROOTS fiber geometries, the role of fiber topology on the precise latencies of conduction to each compartment is more obvious. Especially when comparing the pure-delay mechanism estimate to the 0.1 μm ROOTS fiber behavior (**Figure 12** right), very large differences in AP latencies emerge. This demonstrates how sensitive this test is to conduction velocity alone but also illustrates the importance of accurate representations of fiber microstructure features and arbor topology. Even as the differences are reduced in comparisons with larger and faster conducting fibers, errors of several milliseconds remain. Even should a modeler be so lucky to correctly select a proper conduction velocity (not to mention correctly predicting the site of action potential initiation), differences are difficult to eliminate without more

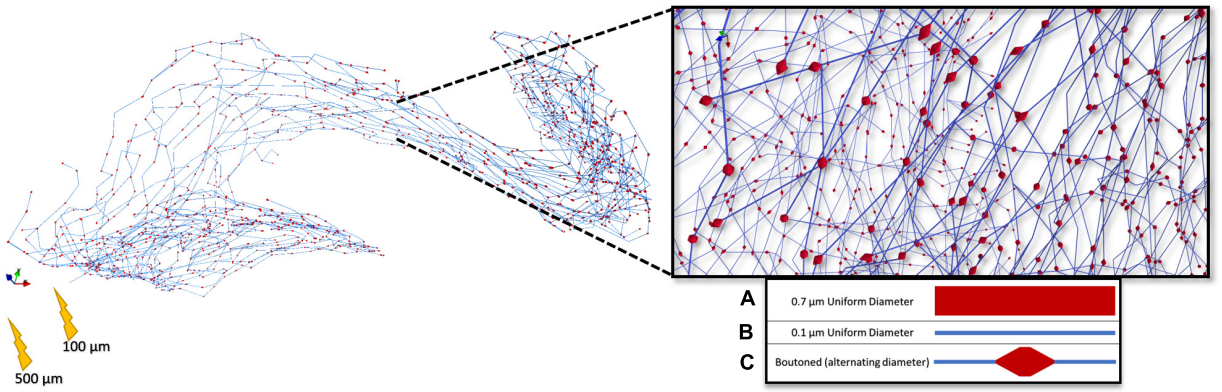

FIGURE 9 | Rendering of a highly branched perforant path axon with presynaptic boutons dispersed evenly throughout the terminal field. 5 µm boutons were spaced 25 µm apart according to Tamamaki and Nojyo (1993). This arbor, with three different representations (A–C) of diameter, was then used in a set of stimulation experiments to study the impact of microstructure geometry on emergent patterns of spatiotemporal patterns of activity.

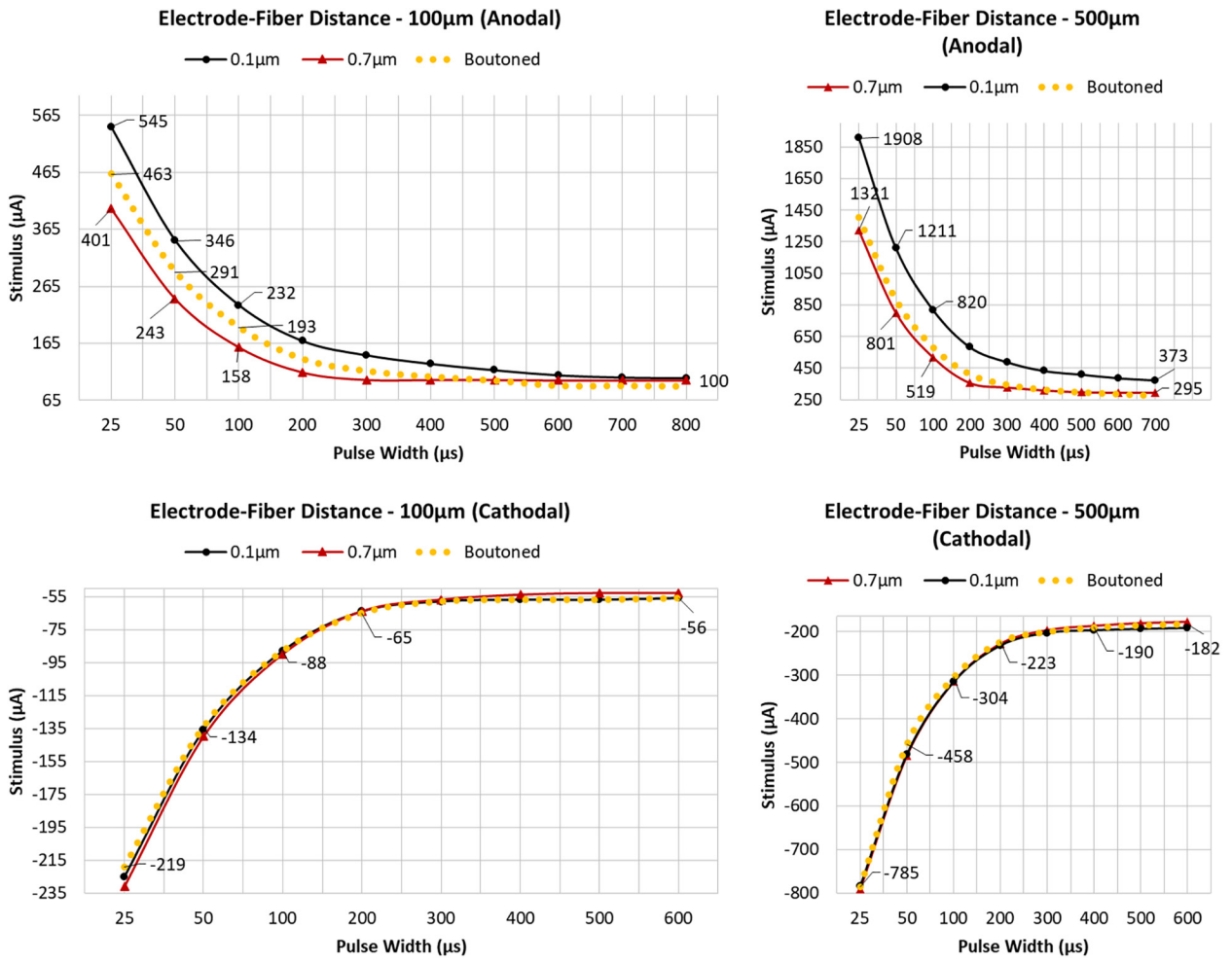

FIGURE 10 | Arbors of three patterns of diameter corresponding to stem and bouton diameters observed by Tamamaki and Nojyo (1993), and a fiber with appropriately distributed boutons, were simulated in response to cathodal pulses (25 µs–1.25 ms) and a range of stimulus amplitudes from distances of 100 and 500 µm **(Left, Right)** Tamamaki and Nojyo (1993). Larger fibers have shorter chronaxies than small and this difference is exaggerated by large electrode–fiber distances and longer pulse-widths. 0.7, 0.1, and Boutoned legend keys correspond to A, B, and C in the **Figure 9** legend.

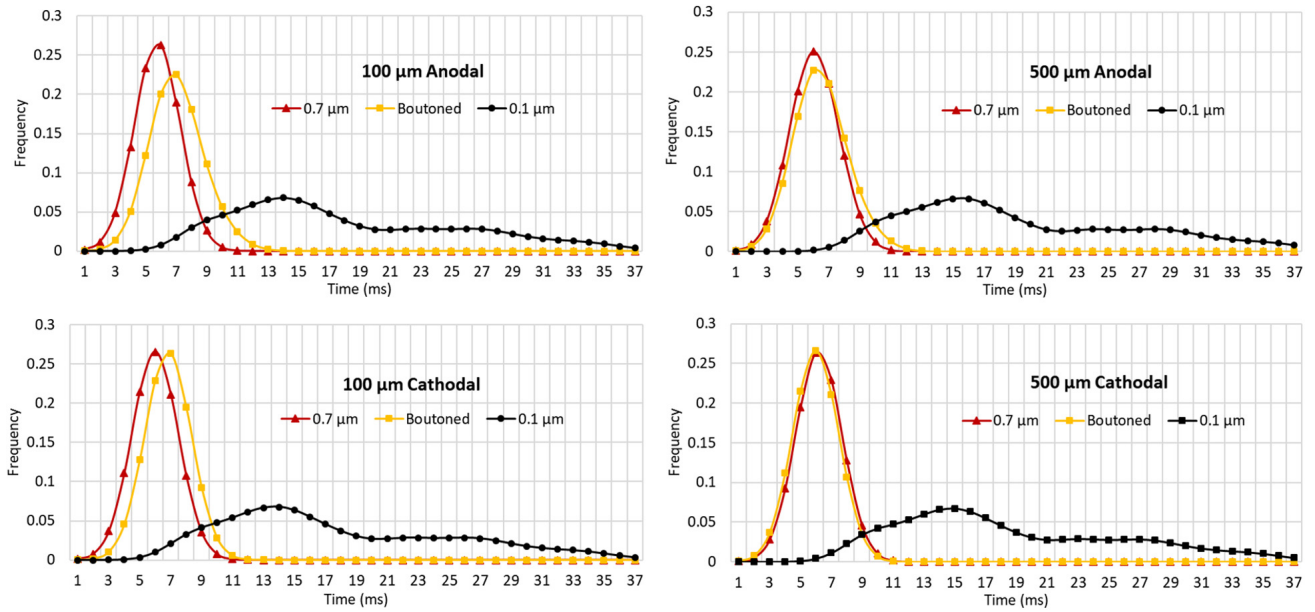

FIGURE 11 | Kernel density estimates (1.5 ms Gaussian kernel) for compartmental action potential delays in perforant path arbors with a diameter of 0.7 and 0.1 µm, and a boutoned fiber when stimulated at ˜rheobase amplitude for 1 ms. From **top** to **bottom**, **left** to **right**, showing two electrode-fiber distances and anodal and cathodal stimulation. Excepting at very close distances, boutoned fibers had temporal activation features best approximated by a fiber with uniform 0.7 µm diameter. 0.7, 0.1, and Boutoned legend keys correspond to A, B, and C in the **Figure 9** legend.

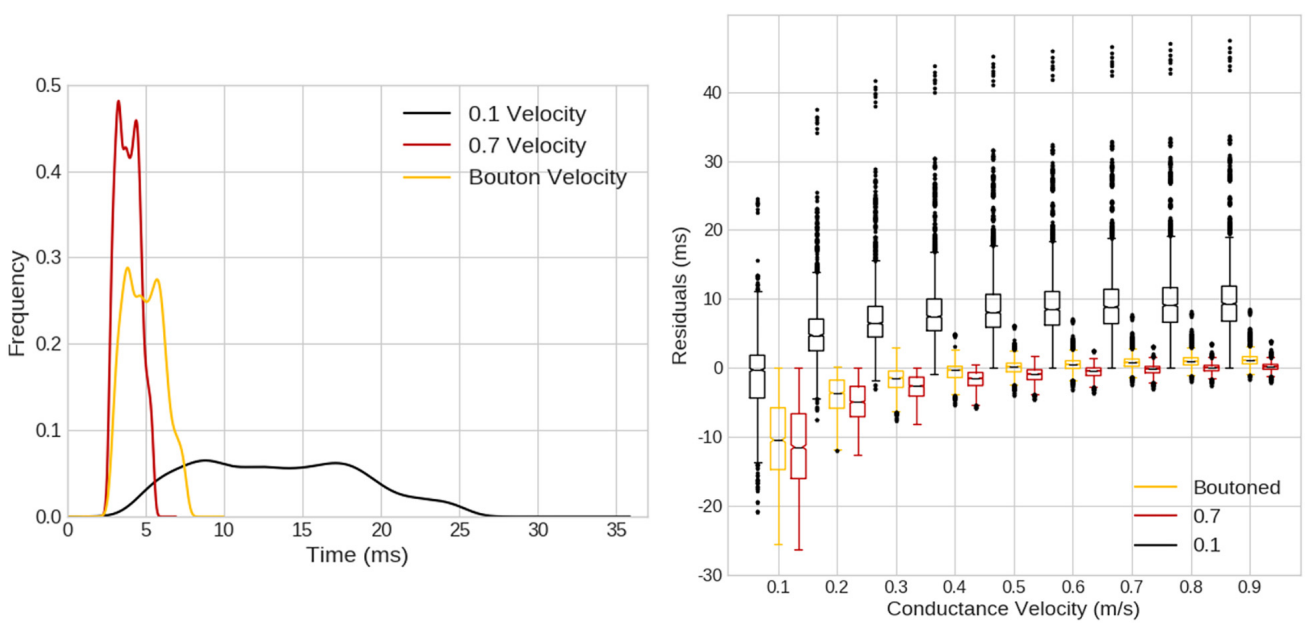

FIGURE 12 | (Left) Kernel density estimates (1.5 ms Gaussian kernel) for expected compartmental action potential delays calculated based on the Euclidean distance from the site of action potential initiation (obtained from ROOTS fibers simulated in **Figures 10**, **11**) to compartment centroids. Three cases are visualized based on the average conductance velocity calculated from fiber behavior in the ROOTS explicit fiber at threshold (100 µm-case). Average velocities were 0.88, 0.53, and 0.12 m/s, for 0.7 µm, boutoned, and 0.1 µm ROOTS fibers, respectively. **(Right)** Pair-wise compartmental latencies were compared between a distance-based delay mechanism and explicit fiber representation via ROOTS across a range of conduction velocities plotted along the *x*-axis. As the conduction velocity moved nearer to the average velocity of the ROOTS fiber, average residuals decreased. Error was largest when comparing the smallest ROOTS fiber, but differences were still on the order of milliseconds near ideal conduction velocity for the faster conducting ROOTS fibers.

sophisticated methods of approximating the path length to each compartment.

DISCUSSION

Together with increased computing power, more robust repositories of electrophysiological and histological information have created opportunities for neural models of greater complexity. Despite these ever-improving sources of data, biologically realistic neural modeling has outpaced experimental studies in many areas of inquiry, providing a testbed for new hypotheses as well as a wealth of preliminary data to aid the design of superior *in vitro* or *in vivo* studies. However, the mismatch in biological realism between experimental and computational modeling, and continuing ambitions to advance computational neuronal modeling, creates a need for sound approaches to generate functional neuronal models that can be refined as the quantity and quality of experimental data improve. The modeling approach presented in this paper demonstrates this utility.

General Application of ROOTS

This manuscript has demonstrated the application of ROOTS to a single fiber system, but this process could be followed for the generation of appropriate models from other anatomical regions with different morphological features. This is true because the criteria used by ROOTS are generally applicable to axon terminal arbors that form a spatial network by connecting parent cells to synaptic targets with a unique branching pattern. The degree of accuracy that can be obtained with ROOTS is directly linked to the quality of anatomical data available to constrain the method. There are three basic tasks that must be accomplished to generate usable models: identify the topography of synaptic targets, identify the branching properties of terminal arbors, and identify the volume from which arbors begin. While these tasks present a challenge, none are as difficult as describing and extracting enough explicit reconstructions from histology to fully recapitulate a large-scale pathway model.

Identifying the topology of synaptic targets, at its most basic, means combining knowledge of the spatial range of terminal arbors with an estimate of the number of terminal boutons expected for each fiber. Identifying branching properties of terminal arbors requires reduction of experimental observations of these terminal arbors to a statistical distribution of bifurcation angles, at the very least. Knowledge of the number of bifurcations, total fiber length, meander angles, and inter-bifurcation lengths is also useful. Lastly, identifying the volume from which arbors initially bifurcate is important in order to force a fundamental directionality on generated arbors. Each of these steps yields important parameters that ROOTS expects and uses to encourage realism in the morphologies that are output by the software. Should one of these specific data not be available, the modeler should not be discouraged from attempting to use ROOTS anyway, as each statistic is merely one part of the puzzle ROOTS seeks to solve. Generated fibers can still be a useful stand-in and can always be augmented as superior data become available.

Limitations and Alternative Methods

Despite the demonstrated functionality of the model presented herein, it is not without limitations. First among potential limitations is biological realism, which may still be limited due to lack of accurate and well described experimental measurements of fibers from hippocampal tissue samples. The clearest limitation (though simultaneously the chief motivation for development) of this modeling approach and the study presented in this paper is the lack of an extensive dataset describing EC axons in the perforant path, including rich morphometrics. Existing explicit reconstructions have readily identifiable errors and, therefore, should not be used as solitary sources of branching morphometrics or be virtualized and used in stimulation models without sophisticated automated and/or manual repair (Tamamaki and Nojyo, 1993; Budd et al., 2010). The issues include slicing artifacts which distorts both distances along or across serial slices and bifurcation angles where branches span multiple slices; angle and distance mismeasurements due to imaging along a single axis and, therefore, failing to correct for the impact of rotation of neural processes with respect to observer perspective; and failures of automated algorithms which may result in orphaned sections or even cycles in the final tree (Quilichini et al., 2010). Despite the unavailability or verifiability of these data, the modeling methodology presented in this study represents the most detailed and sophisticated functional model of layer 2/3 EC spiny stellate dentate perforant path axons to date. In general, the utility of this modeling approach is most apparent for cases in which explicitly reconstructed morphologies are sparse, poorly described, and the fibers to be modeled have complex geometry and branching structure. Further, the parameterization of the model is such that, should this data later become available, the generative morphologies could be updated to reflect new knowledge of *in situ* morphometrics.

While the present study espouses a ROOTS, there are many algorithms that might achieve topological characteristics reminiscent of neuronal branching.

Although existing stochastic models may be effective for dendritic arbor generation, they are not naturally adapted to the construction of axonal morphologies because they fail to reliably conform to pre-determined geometries or volumes that are irregular or highly non-symmetrical (Rozenberg and Salomaa, 1980). The ability to conform to a predetermined volume or geometry is a particularly important feature for axons in the hippocampus, which often have lamellar organization between layers and laminar organization within their terminal fields. For stochastic models to accomplish this, new pre-processing intensive and potentially inefficient volumetric constraints would be required, adding to algorithmic complexity (e.g., randomly walking line that must also connect arbitrary points).

When guided by serial histological sections, spanning-tree algorithms provide an efficient approach to reconstructing spatial trees. However, without some modification, traditional

minimum-spanning-tree algorithms can result in spatial error and distorted branching patterns (Budd et al., 2010). Using higher resolution images forces significant biological realism onto the resulting graph, making error more manageable. However, minimum-spanning-trees generated from points sampled at random from a volume representing an axon terminal field are less biologically realistic due to the extreme path lengths that result. To moderate this outcome, some investigators have proposed a spanning-tree algorithm which balances the minimization of membrane against path length (Cuntz et al., 2007, 2010; Budd et al., 2010; Budd and Kisvarday, 2012). However, optimizing this balance is known to be an np-hard problem due to the direct conflict introduced by including these two variables in the same loss function (Hu, 1974; Alpert et al., 1995; Khuller et al., 1995; Wu et al., 2000; Gastner and Newman, 2006). Another challenge to using path-length/membrane-minimization balancing algorithms is that it only allows indirect control over important tree morphometrics such as bifurcation angle, branch extension angle, and inter-bifurcation length. Prominent implementations of this approach do not allow explicit morphometric thresholds to be set and, therefore, extreme branching patterns remain possible. Inflexibility and lack of sufficient parameterization represent significant limitations of these prior efforts because branching patterns directly impact spatiotemporal patterns of activity.

It is plausible that established stochastic methods, known to be highly proficient at recapitulating tree morphometrics, could be modified to allow directedness and improved conformation to terminal field volumes. Despite these feasible and time-worthy alternatives, the model presented in this study has demonstrated sufficient performance in terms of accuracy, flexibility, and computational complexity.

While many alternatives to RDP for tree simplification exist, this method was preferable to other node cluster-and-merge methods (e.g, Kruskal's algorithm) because these alternatives have the potential to introduce large differences in branching patterns, creating additional challenges in evaluating the equivalence of simplified and unsimplified fibers. Evaluation of the RDP simplification approach could be done in a very straightforward fashion because line simplification doesn't result in changes in branching patterns—simplified arbors were compared to the reference case on the basis of path length and subjective evaluation of an acceptable RDP-epsilon, or the maximum distance between original and simplified contours.

Analysis in this paper involving extracellular field estimations used analytical methods that fail to completely account for anisotropy and heterogeneous resistivity of the tissue volume or full-wave propagation of electric fields. More sophisticated field estimation techniques (e.g., finite element, finite volume, boundary element, or admittance methods) that account for complex impedance or by filtering point-source estimations to account for amplitude dampening and phase-shift of stimuli could have been used but were deemed unnecessarily complex for the principle questions under examination in this article (Bossetti et al., 2007; Al-Humaidi, 2011). It should be noted, however, that rheobase predictions result from very long stimulus pulses, where >95% of the frequency domain signal falls <1 kHz; it follows that tissue capacitance is only a marginal source of error in these estimates of EC axon chronaxie. With respect to resistivity, this exercise is further justified because resistivity measurements in the region performed by López-Aguado et al. (2001) show nearly uniform resistivity throughout the molecular layer of the hippocampus; further, because the current source in these experiments was placed relatively far from resistive boundaries, current shunting distortions should be minimized in our estimations of electric fields in the volume occupied by EC fibers (Bingham et al., 2018). While future work not limited by these assumptions will be performed, doing so here falls outside the scope of the present study.

The study of strength versus duration of stimulus in the measurement of activation thresholds as performed in this paper does not provide conclusive data which concretely establishes the connection between fiber size and excitability for hippocampal networks. This is due to the realistic elements still missing from the models used, including non-uniform diameters of fibers, irregular bouton geometry and volumes, varying bouton topography, and an electrophysiological study of any differences in channel density and dynamics between boutoned and non-boutoned axon regions. Despite these limitations, it seems important to consider what impact boutons may have on fiber excitability within the terminal fields of small, highly branched, and unmyelinated axon fibers. Importantly, these results imply that action potentials in extracellular anodally stimulated axon terminal fields are initiated in boutons with lower input resistance than nearby portions of the fiber and that failing to consider the impact of boutons on conduction velocity will likely result in meaningful temporal errors. However, thorough experimental work is needed to confirm this possibility. It is further valuable to recognize that for long pulses delivered at more distant locations uniformly large diameter fibers approximate the vastly more complex fully boutoned fibers quite well and may provide a viable approach to reducing the computational demands of models that include detailed axon terminal arbors.

While much remains to be done, this study represents a step forward for detailed computational modeling of complex neuronal systems. Where previous models were either focused on peripheral axons with less complex arbors or used sophisticated methods to generate dendritic trees but neglected axons altogether, the model presented here demonstrates an approach to constructing functional axonal morphologies that can be used for diverse applications, including extracellular electrical stimulation of the cortex. It should be noted that while ROOTS is itself general, expert knowledge of the tissue system being virtually recapitulated is required: branching criteria and synaptic or cellular targets must be provided to the tool. Critical morphometric parameters include bifurcation angles, branch extension angle, and internode length; these must be independently determined by users. Despite these requirements, ROOTS presents many opportunities to increase the sophistication of model-based studies of neural tissue system dynamics and central nervous system neuromodulating devices such as deep brain stimulation or potential hippocampal memory prostheses (McIntyre, 2009; Hampson et al., 2018). ROOTS could

also be used to support models of optical or pharmaceutical neuromodulation (Foutz et al., 2012; Bouteiller and Berger, 2017). The model framework could easily be extended to incorporate high resolution imaging data or combined with more complex volume conduction (or diffusion, diffraction, etc.) models to study tissue–electrode interactions at smaller spatial scales (Lujan et al., 2013). In addition to supporting prosthesis design, ROOTS facilitates model-based exploration of the effect of diseases which remodel axons on proper function of hippocampal tissue (e.g., multiple sclerosis) (Michailidou et al., 2015). More generally, ROOTS supports efforts to create networks of neuronal models for the study of biologically plausible spatiotemporal patterns of activity.

AUTHOR SUMMARY

As computer technology matures, constructing virtual models of brain parts has become an increasingly valuable approach to understanding how patterns of activity emerge in different neuronal structures. Many efforts to model populations of neurons have emphasized the implementation of biological realism for cell bodies and dendrites while settling for simplistic representations of axons. This neglect leads to potentially large errors in predictions of when and where synaptically driven activity in a neural circuit might occur. To address this concern, we have developed a novel algorithm called ROOTS to generate biologically realistic axon models for use in computer simulations of the brain. With realistic axons in place, such models can be used to predict how different regions of the brain respond to stimulation from implanted electrodes as part of a prosthetic device. This improvement in the realism of tissue models of the brain will provide superior support to on-going work to reveal the mechanisms of brain disorders and to optimize devices or drugs that are used to treat them.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

CB designed the models and algorithms, conducted the analysis, and prepared the manuscript. AM and J-MB provided support for the design, analysis, and manuscript preparation. J-MB, DS, GL, and TB oversaw the project scope, project conception, and worked to maintain funding to support the project.

FUNDING

This work was supported by the National Institutes of Health (5U01EB025830-07). No other grants directly funded this work.

ACKNOWLEDGMENTS

We thank Romil Nathan for testing the ROOTS system and suggesting various improvements to the codebase in the preparation of software distribution.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncom.2020.00013/full#supplementary-material

FIGURE S1 | Pseudocode of the core algorithm as presented in this study. Other components such as source point reassignment, and topology simplification procedures are executed in series with (before or after) the steps outlined here.

FIGURE S2 | Clustering is performed via K-means clustering method. This regularizes the spatial properties of the mesh which will comprise the volume of the arbor and reduces computational complexity of subsequent steps.

FIGURE S3 | Delaunay triangulation to construct a surface from spatially clustered synaptic targets. This yields a network from which ideal paths to potential target zones can be calculated.

FIGURE S4 | Plotting of each "most likely path" between nodes in the Delaunay triangulated network of k-means cluster centers. Paths are constructed via Dijkstra's algorithm. These paths are provided as inputs to the ROOTS algorithm to loosely guide branching behavior.

FIGURE S5 | A brief table of biophysical parameters used in simulation of NEURON models. The only explicitly varying parameter (those not differing due only to changing compartment areas) was extracellular axial resistance. This followed the pattern implemented in Johnson and McIntyre (2008) to ensure this value was sensitive to fiber diameter.

FIGURE S6 | Entorhinal cortical axons may form as many as 17,700 synapses with granule cells. However, growing axons with this number of nodes is slow and simulating axons with this complexity is computationally prohibitive when attempting to simulate *in situ* scale/density tissue models.

FIGURE S7 | Example graph models (red, LEC; teal, MEC) of entorhinal cortical axon terminal fields, generated by the TREES toolbox method. These axons were generated with the addition of a terminal field in CA3/2/1. For very deep folds with small clefts, the TREES MST method has difficulty suppressing trans-cleft connections that short the cortical circuit without inappropriate manipulation of the number and density of targets, further exacerbating the proliferation of terminal branchlets.

VIDEO S1 | This video shows the spatiotemporal pattern of activity following stimulation of an axon terminal arbor generated with ROOTS.

REFERENCES

Al-Humaidi, A. (2011). *Resistor Networks and Finite Element Models*. Manchester: The University of Manchester.

Alpert, C., Hu, T., Huang, J., Kahng, A., and Karger, D. (1995). "Prim-Dijkstra tradeoffs for improved performance-driven routing tree design," in *Proceedings of the IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Piscataway, NJ.

Anastassiou, C. A., and Koch, C. (2015). Ephaptic coupling to endogenous electric field activity: why bother? *Curr. Opin. Neurobiol.* 31, 95–103. doi: 10.1016/j.conb.2014.09.002

Anderson, R., Farokhniaee, A., Gunalan, K., Howell, B., and McIntyre, C. (2018). Action potential initiation, propagation, and cortical invasion in the hyperdirect pathway during subthalamic deep brain stimulation. *Brain Stimul.* 11, 1140–1150. doi: 10.1016/j.brs.2018.05.008

Ascoli, G., and Krichmar, J. (2000). L-neuron: a modeling tool for the efficient generation and parsimonious description of dendritic morphology. *Neurocomputing* 32, 1003–1011. doi: 10.1016/s0925-2312(00)00272-1

Bindocci, E., Savtchouk, I., Liaudet, N., Becker, D., Carriero, G., and Volterra, A. (2017). Three-dimensional Ca2+ imaging advances understanding of astrocyte biology. *Science* 356:6339. doi: 10.1126/science.aai8185

Bingham, C., Loizos, K., Yu, G., Gilbert, A., Bouteiller, J., Song, D., et al. (2016). "A large-scale detailed neuronal model of electrical stimulation of the dentate gyrus and perforant path as a platform for electrode design and optimization," in *Proceedings of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Piscataway, NJ, 2794–2797.

Bingham, C., Loizos, K., Yu, G., Gilbert, A., Bouteiller, J.-M., Song, D., et al. (2018). Model-based analysis of electrode placement and pulse amplitude for hippocampal stimulation. *IEEE Trans. Biomed. Eng.* 99:1. doi: 10.1109/TBME.2018.2791860

Bossetti, C. A., Birdno, M. J., and Grill, W. M. (2007). Analysis of the quasi-static approximation for calculating potentials generated by neural stimulation. *J. Neural Eng.* 5:44. doi: 10.1088/1741-2560/5/1/005

Bouteiller, J. M., and Berger, T. W. (2017). "The role of simulations in neuropharmacology," in *Computational Neurology and Psychiatry*, eds P. Érdi, B. Sen Bhattacharya, A. Cochran, (Cham: Springer).

Budd, J., and Kisvarday, Z. (2012). Communication and wiring in the cortical connectome. *Front. Neuroanat.* 6:42. doi: 10.3389/fnana.2012.00042

Budd, J., Kovács, K., Ferecskó, A., Buzás, P., Eysel, U., and Kisvárday, Z. (2010). Neocortical axon arbors trade-off material and conduction delay conservation. *PLoS Comput. Biol.* 6:e1000711. doi: 10.1371/journal.pcbi.1000711

Cayley, A. (1889). A theorem on trees. *Q. J. Pure Appl. Math.* 23, 376–378.

Chen, J. (2003). Dijkstra's shortest path algorithm. *J. Formal. Math.* 15, 237–247. doi: 10.1109/TVCG.2010.57

Chew, L. (1989). Constrained delaunay triangulations. *Algorithmica* 4, 97–108. doi: 10.1007/bf01553881

Claiborne, B., Amaral, D., and Cowan, W. (1990). Quantitative, three-dimensional analysis of granule cell dendrites in the rat dentate gyrus. *J. Comp. Neurol.* 302, 206–219. doi: 10.1002/cne.903020203

Clark, J., and Plonsey, R. (1970). A mathematical study of nerve fiber interaction. *Biophys. J.* 10, 937–957. doi: 10.1016/s0006-3495(70)86344-5

Cuntz, H., Borst, A., and Segev, I. (2007). Optimization principles of dendritic structure. *Theor. Biol. Med. Model.* 4:21.

Cuntz, H., Forstner, F., Borst, A., Häusser, M., Kuhlman, S., and Saggau, P. (2010). One rule to grow them all: a general theory of neuronal branching and its practical application. *PLoS Comput. Biol.* 6:e1000877. doi: 10.1371/journal.pcbi.1000877

Desmond, N. L., and Levy, W. B. (1985). Granule cell dendritic spine density in the rat hippocampus varies with spine shape and location. *Neurosci. Lett.* 54, 219–224. doi: 10.1016/s0304-3940(85)80082-3

Foutz, T. J., Arlow, R. L., and McIntyre, C. C. (2012). Theoretical principles underlying optical stimulation of a channelrhodopsin-2 positive pyramidal neuron. *Am. J. Physiol. Heart Circ. Physiol.* 107, 3235–3245. doi: 10.1152/jn.00501.2011

Fries, P. (2005). A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. *Trends Cogn. Sci.* 9, 474–480. doi: 10.1016/j.tics.2005.08.011

Gaarskjaer, F. B. (1978). Organization of the mossy fiber system of the rat studied in extended hippocampi. I. Terminal area related to number of granule and pyramidal cells. *J. Comparative Neurol.* 178, 49–72.

Gastner, M. T., and Newman, M. E. J. (2006). Shape and efficiency in spatial distribution networks. *J. Stat. Mech.* 2006, P01015–P01015. doi: 10.1088/1742-5468/2006/01/p01015

Grill, W. (1999). Modeling the effects of electric fields on nerve fibers: influence of tissue electrical properties. *IEEE Trans. Biomed. Eng.* 46, 918–928. doi: 10.1109/10.775401

Hagberg, A., Schult, D., and Swart, P. (2005). *Networkx: Python Software for the Analysis of Networks. Mathematical Modeling and Analysis*. New Mexico: Los Alamos National Laboratory.

Hama, K., Arii, T., and Kosaka, T. (1989). Three-dimensional morphometrical study of dendritic spines of the granule cell in the rat dentate gyrus with HVEM stereo images. *J. Electron Microsc. Tech.* 12, 80–87. doi: 10.1002/jemt.1060120203

Hampson, R., Song, D., Robinson, B., Fetterhoff, D., Dakos, A., Roeder, B., et al. (2018). Developing a hippocampal neural prosthetic to facilitate human memory encoding and recall. *J. Neural Eng.* 15:036014. doi: 10.1088/1741-2552/aaaed7

Hartigan, J., and Wong, M. (1979). Algorithm AS 136: a k-means clustering algorithm. *J. R. Statist. Soc.* 28, 100–108.

Hendrickson, P. J., Yu, G., Song, D., and Berger, T. W. (2016). A million-plus neuron model of the hippocampal dentate gyrus: critical role for topography in determining spatiotemporal network dynamics. *IEEE Trans. Biomed. Eng.* 63:576. doi: 10.1109/TBME.2015.2445771

Hines, M., and Carnevale, N. (2003). "The neuron simulation environment," in *The Handbook of Brain Theory and Neural Networks*, 2nd Edn, ed. M. A. Arbib, (Cambridge, MA: MIT Press).

Hjorth-Simonsen, A., and Jeune, B. (1972). Origin and termination of the hippocampal perforant path in the rat studied by silver impregnation. *J. Comp. Neurol.* 144, 215–231. doi: 10.1002/cne.901440206

Holt, G., and Koch, C. (1999). Electrical interactions via the extracellular potential near cell bodies. *J. Comput. Neurosci.* 6, 169–184.

Howell, B., and McIntyre, C. (2016). Role of soft-tissue heterogeneity in computational models of deep brain stimulation. *Brain Stimul.* 10, 46–50. doi: 10.1016/j.brs.2016.09.001

Hu, T. (1974). Optimum communication spanning trees. *SIAM J. Comput.* 3, 188–195. doi: 10.1137/0203015

Johnson, M., and McIntyre, C. (2008). Quantifying the neural elements activated and inhibited by globus pallidus deep brain stimulation. *J. Neurophysiol.* 100, 2549–2563. doi: 10.1152/jn.90372.2008

Joucla, S., and Yvert, B. (2009). The "mirror" estimate: an intuitive predictor of membrane polarization during extracellular stimulation. *Biophys. J.* 96, 3495–3508. doi: 10.1016/j.bpj.2008.12.3961

Khuller, S., Raghavachari, B., and Young, N. (1995). Balancing minimum spanning trees and shortest-path trees. *Algorithmica* 14:305. doi: 10.1007/bf01294129

Kim, J., Lee, H., Choi, W., and Lee, K. (2019). Encoding information into autonomously bursting neural network with pairs of time-delayed pulses. *Sci. Rep.* 9:1394. doi: 10.1038/s41598-018-37915-7

Kjonigsen, L., Leergaard, T., Kjode, J., Witter, M., and Bjaalie, J. (2008). Interactive histological atlas system for anatomical parcellation of the rat hippocampus proper, fasciola and associated parahippocampal cortex. *Front. Neuroinform.*

López-Aguado, L., Ibarz, J., and Herreras, O. (2001). Activity-dependent changes of tissue resistivity in the CA1 region in vivo are layer-specific: modulation of evoked potentials. *Neuroscience* 108, 249–262. doi: 10.1016/s0306-4522(01)00417-1

Lujan, J. L., Chaturvedi, A., Choi, K. S., Holtzheimer, P. E., Gross, R. E., Mayberg, H. S., et al. (2013). Tractography-activation models applied to subcallosal cingulate deep brain stimulation. *Brain Stimul.* 6, 737–739. doi: 10.1016/j.brs.2013.03.008

McIntyre, C. (2009). Computational modeling of deep brain stimulation. *Neuromodulation* 116, 171–178. doi: 10.1016/b978-0-12-374248-3.00017-3

McIntyre, C., and Grill, W. (1999). Excitation of central nervous system neurons by nonuniform electric fields. *Biophys. J.* 76, 878–888. doi: 10.1016/s0006-3495(99)77251-6

McKinney, W. (2010). "Data structures for statistical computing in python," in *Proceedings of the 9th Python in Science Conference*, Piscataway, NJ.

Michailidou, I., Willems, J., Kooi, E., van Eden, C., Gold, S., Geurts, J., et al. (2015). Complement C 1q-C 3–associated synaptic changes in multiple sclerosis hippocampus. *Ann. Neurol.* 77, 1007–1026. doi: 10.1002/ana.24398

Nowak, L., and Bullier, J. (1996). Spread of stimulating current in the cortical grey matter of rat visual cortex studied on a new in vitro slice preparation. *J. Neurosci. Methods* 67, 237–248. doi: 10.1016/0165-0270(96)00065-9

Nowak, L., and Bullier, J. (1998). Axons, but not cell bodies, are activated by electrical stimulation in cortical gray matter. *Exp. Brain Res.* 118, 477–488. doi: 10.1007/s002210050304

Patton, P. E., and McNaughton, B. (1995). Connection matrix of the hippocampal formation: I. The dentate gyrus. *Hippocampus* 5, 245–286. doi: 10.1002/hipo.450050402

Quilichini, P., Sirota, A., and Buzsáki, G. (2010). Intrinsic circuit organization and theta–gamma oscillation dynamics in the entorhinal cortex of the rat. *J. Neurosci.* 30, 11128–11142. doi: 10.1523/JNEUROSCI.1327-10.2010

Ramachandran, P. (2001). "MayaVi: a free tool for CFD data visualization," in *Proceedings of the 4th annual CFD symposium, Aeronautical Society of India*, Piscataway, NJ.

Ranck, J. (1975). Which elements are excited in electrical stimulation of mammalian central nervous system: a review. *Brain Res.* 98, 417–440. doi: 10.1016/0006-8993(75)90364-9

Rattay, F., Paredes, L., and Leao, R. (2012). Strength–duration relationship for intra-versus extracellular stimulation with microelectrodes. *Neuroscience* 214, 1–13. doi: 10.1016/j.neuroscience.2012.04.004

Rozenberg, G., and Salomaa, A. (1980). *The Mathematical Theory of L Systems*. Cambridge, MA: Academic press.

Saalfeld, A. (1999). Topologically consistent line simplification with the Douglas-Peucker algorithm. *Cartogr. Geograph. Inform. Sci.* 26, 7–18. doi: 10.1559/152304099782424901

Schwartz, S. P., and Coleman, P. D. (1981). Neurons of origin of the perforant path. *Exp. Neurol.* 74, 305–312. doi: 10.1016/0014-4886(81)90169-2

Scorcioni, R., and Ascoli, G. (2005). Algorithmic reconstruction of complete axonal arborizations in rat hippocampal neurons. *Neurocomputing* 65, 15–22. doi: 10.1016/j.neucom.2004.10.105

Scorcioni, R., Polavaram, S., and Ascoli, G. (2008). L-Measure: a web-accessible tool for the analysis, comparison and search of digital reconstructions of neuronal morphologies. *Nat. Protoc.* 3:866. doi: 10.1038/nprot.2008.51

Tamamaki, N., and Nojyo, Y. (1993). Projection of the entorhinal layer II neurons in the rat as revealed by intracellular pressure-injection of neurobiotin. *Hippocampus* 3, 471–480. doi: 10.1002/hipo.450030408

Türetken, E., González, G., Blum, C., and Fua, P. (2011). Automated reconstruction of dendritic and axonal trees by global optimization with geometric priors. *Neuroinformatics* 9, 279–302. doi: 10.1007/s12021-011-9122-1

Whittington, M., Stanford, I., Colling, S., Jefferys, J., and Traub, R. (1997). Spatiotemporal patterns of ? frequency oscillations tetanically induced in the rat hippocampal slice. *J. Physiol.* 502, 591–607. doi: 10.1111/j.1469-7793.1997.591bj.x

Witter, M. (2007). The perforant path: projections from the entorhinal cortex to the dentate gyrus. *Prog. Brain Res.* 163, 43–61. doi: 10.1016/s0079-6123(07)63003-9

Wu, B., Chao, K.-M., and Tang, C. (2000). Approximation algorithms for some optimum communication spanning tree problems. *Discrete Appl. Math.* 102, 245–266. doi: 10.1016/s0166-218x(99)00212-7

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.

frontiers in Computational Neuroscience

METHODS
published: 22 April 2020
doi: 10.3389/fncom.2020.00034

Topological View of Flows Inside the BOLD Spontaneous Activity of the Human Brain

Arjuna P. H. Don[1], *James F. Peters*[1]*, *Sheela Ramanna*[2] *and Arturo Tozzi*[1,3]

[1] Computational Intelligence Laboratory, University of Manitoba, Winnipeg, MB, Canada, [2] Applied Computer Science, University of Winnipeg, Winnipeg, MB, Canada, [3] Department of Physics, University of North Texas, Denton, TX, United States

Spatio-temporal brain activities with variable delay detectable in resting-state functional magnetic resonance imaging (rs-fMRI) give rise to highly reproducible structures, termed cortical lag threads, that propagate from one brain region to another. Using a computational topology of data approach, we found that persistent, recurring blood oxygen level dependent (BOLD) signals in triangulated rs-fMRI videoframes display previously undetected topological findings, i.e., vortex structures that cover brain activated regions. Measure of persistence of vortex shapes in BOLD signal propagation is carried out in terms of Betti numbers that rise and fall over time during spontaneous activity of the brain. Importantly, a topology of data given in terms of geometric shapes of BOLD signal propagation offers a practical approach in coping with and sidestepping massive noise in neurodata, such as unwanted dark (low intensity) regions in the neighborhood of non-zero BOLD signals. Our findings have been codified and visualized in plots able to track the non-trivial BOLD signals that appear intermittently in a sequence of rs-fMRI videoframes. The end result of this tracking of changing lag structures is a so-called persistent barcode, which is a pictograph that offers a convenient visual means of exhibiting, comparing, and classifying brain activation patterns.

Keywords: Betti Numbers, brain activity, fMRI video, persistence bar code, topological data analysis

OPEN ACCESS

Edited by:
Tatyana Sharpee,
Salk Institute for Biological Studies, United States

Reviewed by:
Tuo Zhang,
Northwestern Polytechnical University, China
Valeri Makarov,
Complutense University of Madrid, Spain

***Correspondence:**
James F. Peters
james.peters3@umanitoba.ca

Received: *30 October 2019*
Accepted: *30 March 2020*
Published: *22 April 2020*

Citation:
Don APH, Peters JF, Ramanna S and Tozzi A (2020) Topological View of Flows Inside the BOLD Spontaneous Activity of the Human Brain. Front. Comput. Neurosci. 14:34. doi: 10.3389/fncom.2020.00034

1. INTRODUCTION

Point clouds are a natural outcome of a topology of data approach in tracking intermittent as well as persistent BOLD signals in different sections of the brain. A point cloud is a collection of sampled pinpointed places in a subregion (Ghrist, 2014). A topology of data circumvents noise in data and focuses on those data that persist over time (Edelsbrunner and Harer, 2010). Computational topology of data provides a practical method in isolating, measuring, and classifying persistent lag structures BOLD signals in each rs-fMRI video frame that are mapped to point clouds in a finite, bounded region in an n-dimensional Euclidean metric space. It has been observed that brain activity in one region of the brain propagates to others with variable temporal delay (Mitra et al., 2015; Matsui et al., 2016; Park et al., 2019), giving rise to brain activity lag (delay) structures. Lag threads are temporal sequences of propagated brain activities (Mitra et al., 2015). Lag structures in fMRI video frames are rich source of point clouds. Triangulated brain point clouds are a source of brain activation area shapes that appear intermittently in different cortical regions during a rs-fMRI video such as the four videos provided by Mitra et al. (2015). Selected barycenters of triangles in a triangulated cortical point cloud are connected to form vortexes covering each brain

activation region with its own distinctive shape. Each vortex in a triangulated rs-fMRI video frame is a collection of connected cycles (see, e.g., the lower half of **Figure 1**) that make it possible to approximate, measure, track and compare brain activation region shapes. A vortical view of brain activity first appeared in Freeman (2009).

Each vertex in a triangulated brain activation region point cloud is represented by a feature vector containing useful information such as activation region area and representative Betti number (Giusti et al., 2015). Traditionally, the more intuitive geometric forms of Betti numbers are counts of cells (vertexes, edges, filled triangles), cycle counts or surface hole counts (Zomorodian, 2001). A particularly useful intuitive form of geometric Betti number is a count of the number of connected vortex cycles covering an activation subregion (Peters, 2020). Less frequently used algebraic Betti numbers (counts of the number of generators in a free Abelian group; Munkres, 2000) also provide insight concerning the inner workings of cycles in triangulated brain activation regions. We first consider the persistence of geometric numbers (vortex cycle counts) over sequences of sequences of triangulated rs-FMRI video frames. Later, in **Appendix B**, the persistence of algebraic Betti numbers is also considered to obtain an alternative view of the changing character of connected cycles in triangulated rs-FMRI lag threads in three regions of the brain.

Tracking the appearances of the Betti number of a brain activation vortex containing a particular number of connected cycles leads to the construction of a persistence barcode (see the top half of **Figure 1**) in which a Betti number appears in a video frame, disappears afterward and possibly reappears one or more times in later video frames. 2D (planar) as well as 3D (volumetric) persistent barcodes provide an easy-to-read means of tracking intermittent BOLD signals in a sequence of rs-fMRI video frames.

The origin of topological data analysis and persistent homology can be traced back to Edelsbrunner et al. (2000, 2001). A common approach is to build a continuous shape (graphs) on top of data to detect complex topological and underlying geometric structures (Carlsson, 2009; Chazal and Michel, 2017). This shape is called a simplicial complex or a nested family of simplicial complexes and the process of shape construction is commonly referred to a filtration (Zomorodian, 2001). One of the fundamental tools in computational topology is persistent homology (Zomorodian and Carlsson, 2005), which is a powerful tool to compute, study, and efficiently encode multiscale topological features of nested families of simplicial

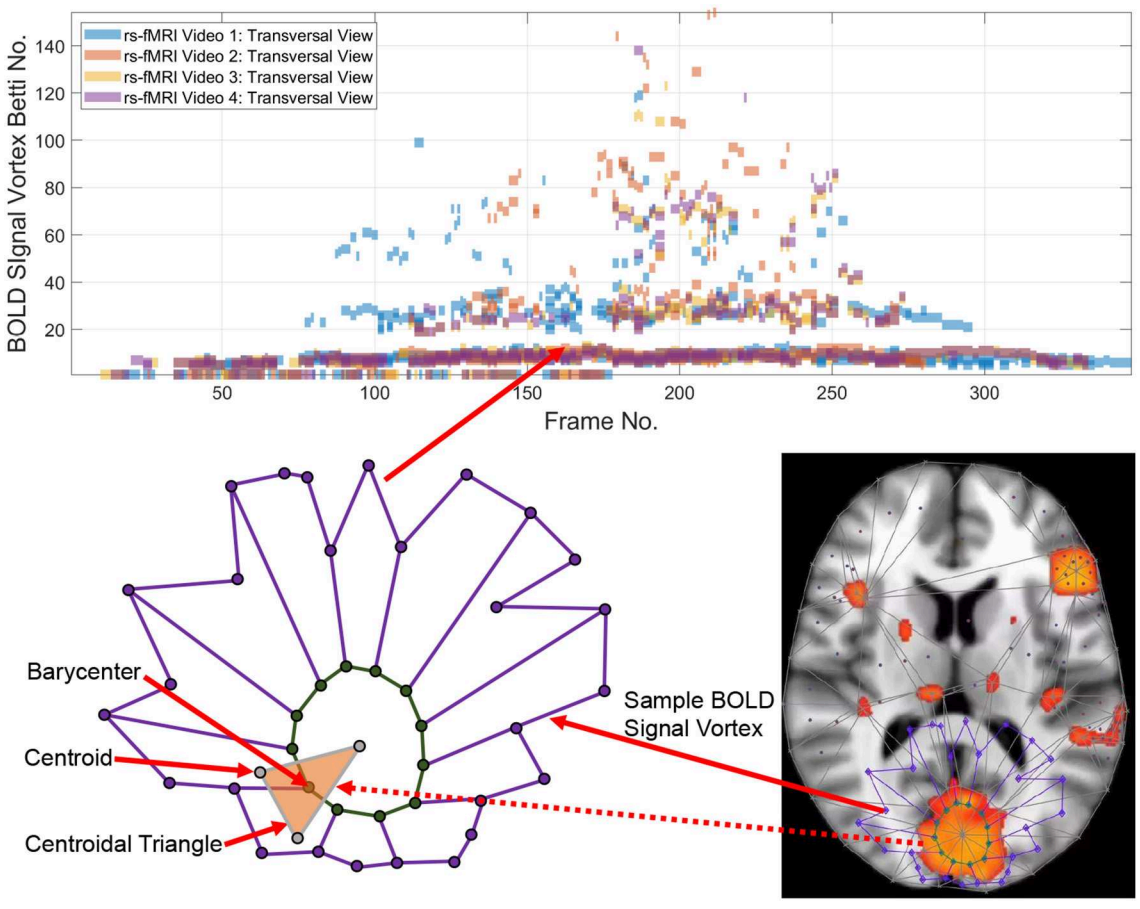

FIGURE 1 | Betti numbers for a rs-fMRI BOLD signal vortex on a Transversal view for four videos.

complexes and topological spaces (Edelsbrunner and Harer, 2008).

Earlier studies of brain networks primarily focused on a graph-theoretic approach where the brain regions and their connections are encoded as a graph (i.e., a network of nodes and edges) and cycles (representing complex behaviors). Such networks were modeled and analyzed with methods such as Q-modularity (Meunier et al., 2009) or with network measures such as betweeness centrality (Bullmore and Sports, 2009). Brain networks with weighted edges where problems of selecting thresholds for edge weights and dealing with sparse edges can be found in Achard and Bullmore (2007) and van Wijk et al. (2010). This has led to application of persistent homology to the problem of determining multiple thresholds derived from more than one network. A brain network can be considered as the 1-skeleton of a simplicial complex, where the 0-dimensional hole is the connected component, and the 1-dimensional hole is a cycle (Chung et al., 2019). The number of k-dimensional holes of a simplicial complex is its k-th Betti number. Persistent homology-based multiscale hierarchical modeling was proposed in Lee et al. (2011b), Petri et al. (2014), Giusti et al. (2015), Sizemore et al. (2018), and Chung et al. (2019) to name a few. Here, graph filtration is used to build these networks in a hierarchically manner. Filtration is the process of connecting edges to form a graph.

Betti numbers computed during this filtration process have been used for further statistical analysis such as Pearson correlation to MRI image data (Chung et al., 2015) and various metrics for similarity and distances assessments (Lee et al., 2011a; Chung et al., 2019). 0-dimensional holes (β_0 or zeroth Betti numbers) have been computed during the graph filtration process and a persistent barcode has been constructed for subsequent statistical analysis (Cassidy et al., 2015; Chung et al., 2017). However, the *cycle* concept is extremely important to the study of behavior diffusion and integration of the brain network. In persistent homology, cycles are measured using the first or one-dimensional Betti Number (β_1). In Chung et al. (2019), the one dimensional Betti numbers are used to measure cycle and the significance of the number of cycles is evaluated using the Kolmogorov–Smirnov (KS) distance.

For geometric representations of rs-fMRI lag threads, an incisive form of statistical analysis of the geometry is given in terms of edge density and Eigen values (Giusti et al., 2015). The results of this form of geometry-based statistical analysis are given in **Appendix A**.

In contrast to earlier approaches, we use computational geometry to detect lag thread shapes in fMRI video frames using a *geometric* Betti number that counts the total number of connected cycles forming a vortex (nested, usually non-concentric, connected cycles) derived from the triangulation of brain activation regions. We build on our previous work in Peters et al. (2017), Don and Peters (2019), and Peters (2020), to evaluate real BOLD resting state rs-fMRI videos from Mitra et al. (2015).

Here, the video frames are processed directly to obtain Betti numbers by triangulating the transversal, sagittal, and coronal sections of the fRMI video frames and constructing vortexes through a process of filtration. Rather than constructing graphs of brain networks and analysing cycles in the network, we analyze cycles by processing the video frames. The vortexes correspond to the changing activation areas in the video frames. The number of vortexes in a frame represents the most relevant areas of change. Higher Betti number values imply that the change is closer to the center of the section of the brain. To the best of our knowledge, this method of detecting cycles in persistent homology is novel and a preliminary version of this paper appeared in Don A. P. et al. (2019).

2. MATERIALS AND METHODS

2.1. Theory

The basic approach is to introduce a geometric representation of brain activation regions in terms of intersecting cycles that are sequences of path-connected vertexes on the barycenters of triangles forming vortices (Don and Peters, 2019; Peters, 2020). Each vortex is a collection of connected cycles called a vortex nerve. Of particular interest are those nerves that have a maximal collection of triangles of a common vertex in the triangulation of a finite, bounded planar region. In our case, the planar region is a rs-fMRI video frame. A vortex nerve results from the triangulation of the sections of each rs-fMRI brain video frame. A centroid (also called a seed point), is used as a vertex in the triangulation of a video frame. A barycenter on a such a triangle is in a high light intensity video frame region between the dark regions, which we refer to as *holes*.

Definition 1 (Hole). A ***hole*** is a collection of contiguous low intensity voxels. The ***centroid of a hole*** is the center of mass of the hole. In an rs-fMRI video frame, the holes are in background regions containing dark (low intensity) voxels and the foreground regions are filled with mainly orange voxels. Sample centroids of brain activation region holes are given in **Figure 1**.

Using Delaunay triangulation (Delaunay, 1934; Yung et al., 2016), each pair of closest neighboring centroids of holes become the vertexes of edges of triangles covering brain activation regions in rs-fMRI videos. Each line segment drawn between closest pairs of centroids becomes the edge of a triangle in the Delaunay triangulation of an rs-fMRI video frame. The concept of a hole is crucial to this work, since edges drawn between barycenters in the interior of adjacent centroidal triangles reveal paths of high intensity voxels between brain activation region holes.

Definition 2 (Barycenter of centroidal triangle). Recall that the median line of a triangle is a line drawn from a vertex to the midpoint of the opposite side of the triangle. The ***barycenter of a centroidal triangle*** is the point of intersection of the median lines of the triangle. A sample barycenter is given in **Figure 1**.

The barycenters of centroidal triangles covering brain activation regions are always between holes. Barycenters are stepping stones in the construction vortex cycles.

Definition 3 (Barycentric cycle). A ***barycentric cycle*** is a sequence of edges drawn between neighboring barycenters of adjacent centroidal triangles.

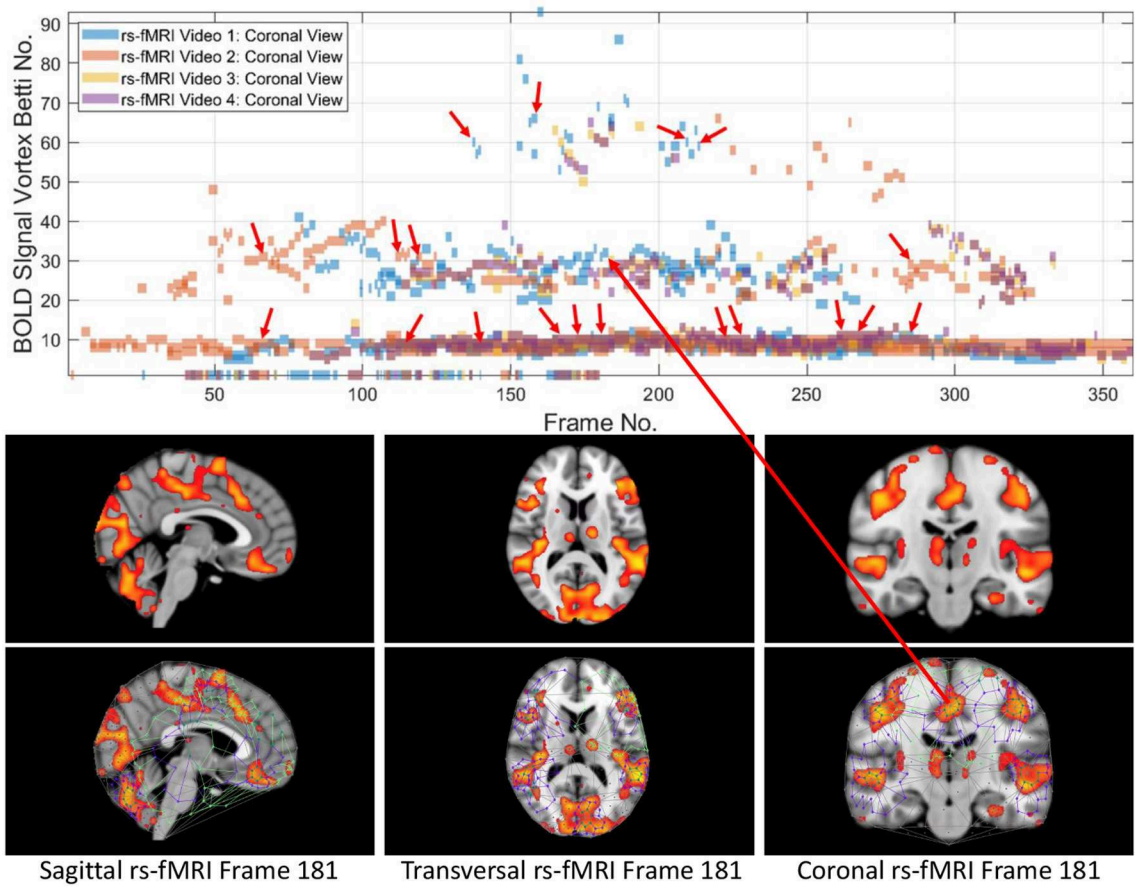

FIGURE 2 | Coronal persistence barcode with Betti numbers as vortex cycle counts in a rs-fMRI video frame (see text for further details).

As a result, connected barycenters model paths for high intensity voxels recorded in a brain activity video frame. Barycentric cycles are the basic building blocks in the construction of local vortexes covering triangulated brain activation regions.

Definition 4 (Local Vortex (briefly, Vortex)). A ***local vortex*** is a collection of nesting, usually non-concentric barycentric cycles. The simplest vortex contains a single barycentric cycle (see, for example, the vortex in **Figure 3**).

2.2. Betti Numbers on a Triangulated Brain Activation Region

Betti numbers provide a computational topology perspective on the structure of brain activation subregions. In our case, the ***Betti number*** β_1 tells us either the number of connected vortex cycles in a vortex on a triangulated brain activation region (***geometric view***). Later, in **Appendix B**, we also introduce and apply the Betti number β_α, which is a count of the number of generators in a free Abelian group representation of an rs-fMRI video frame vortex (***algebraic view***). The focus here is on the persistent geometric Betti numbers across sequences of triangulated video frames. Each such Betti number is mapped to an entry in a persistent barcode (see top half of **Figure 1**). This topology of data pictographs is useful in representing the persistence of the brain activation region shapes found in rs-fMRI brain video sections. From an intuitive perspective, there are three types of geometric Betti numbers, namely, $\beta_0, \beta_1, \beta_2$, introduced in Zomorodian (2001).

Definition 5 (β_0). The ***Betti number*** $\boldsymbol{\beta_0}$ is a count of the total number of elementary cell complexes, which are vertexes, edges and filled triangles attached to each other in a triangulated region.

Definition 6 (β_1). The ***Betti number*** $\boldsymbol{\beta_1}$ is a count of the number cycles.

Each β_1 in a persistent barcode (Ghrist, 2008) represents the number of connected barycentric cycles covering an activation area of the brain.

Definition 7 (β_2). The ***Betti number*** $\boldsymbol{\beta_2}$ is a count of the number of holes. In our case, $\boldsymbol{\beta_2}$ is a count of the number of contiguous low intensity voxels in brain activation regions in an rs-fMRI video.

Definition 8 (C_0 vortex cycle). ***cycle*** $\boldsymbol{C_0}$ is the ***innermost cycle in a vortex***. In our case, the sequence of

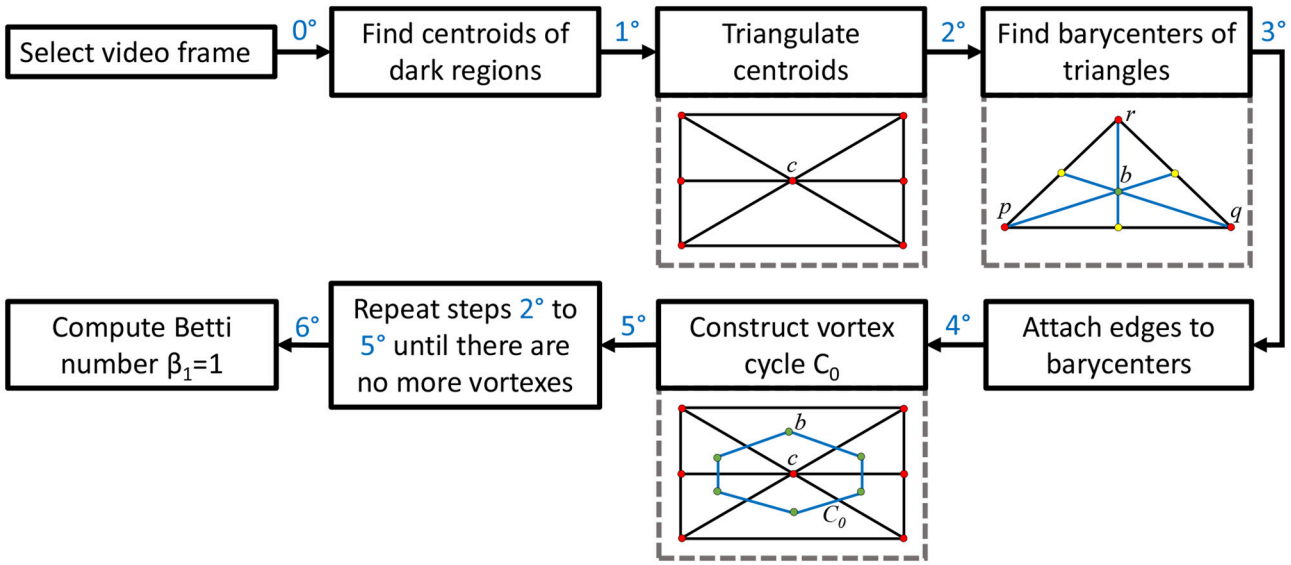

FIGURE 3 | Method flow graph.

edges connected between barycenters of adjacent triangles with a common vertex, collectively called an Alexandroff nerve (Alexandroff, 1965) forms the *cycle* C_0. Notice that cycle C_0 will always be in the interior of a brain activation region containing high intensity voxels.

Definition 9 (C_1). In a vortex with 2 nested cycles, ***cycle*** $\boldsymbol{C_1}$ is the vortex cycle that has only C_0 in its interior. In effect, cycle C_1 is a collection of path-connected vertices on a sequence of edges surrounding C_0. Cycle C_1 usually overlaps the boundary of a high intensity brain activation region in an rs-fMRI video frame.

Definition 10 (Bridge Segment). A ***bridge segment*** is an edge attached between vertices on a pair of neighboring cycles. Let *cycA*, *cycB* be a pair of neighboring cycles (i.e., there are no cycles in between *cycA* and *cycB*) and let p be a vertex on *cycA* and q, a vertex on *cycB*. The edge $\overline{pq}$ is a bridge segment. Edge $\overline{cj}$ between vertices c and j is a bridge segment and there is no bridge segment between vertices c and i in **Figure 5**.

In the main body of this paper, we give barcoded video results for the more intuitive geometric Betti number counts of vortex cycles on triangulated brain activation regions. Later, in **Appendix B**, we give examples of both geometric and algebraic forms of Betti numbers. A repetition of the same β_1 across a sequence of consecutive frames tells us that a similar vortex shape recurs in these frames. The geometric Betti number of a vortex containing two cycles with k bridge segments attached between the pair of vortex cycles equals k + 2 (Don and Peters, 2019).

For example, the Betti number β_0 of the vortex in the lower half of **Figure 1** equals 12 and $\beta_1 = 1$, since there are 12 cycle edges and there is only one cycle in the vortex shown.

2.3. rs-fMRI Lag Threads Having Descriptive Proximity

A pair of objects are descriptively proximal (near each other), provided the objects have the same description (Di Concilio et al., 2018). A ***feature vector*** provides a description of an object. In this work, feature vector (Betti number, area) describes a subregion of a rs-fMRI brain region. Rather than a purely theoretical, abstract approach to descriptive proximity spaces, the focus here is on computational descriptively proximities. Briefly, computational descriptive proximity includes algorithms as well as structures such as set intersection, union and closure and proximity space axioms introduced in Peters (2020). Descriptive proximites provide mathematical framework useful in measuring, comparing, and classifying (1) lag structures and threads across frames in the same video or (2) lag structures and threads across frames in different rs-fMRI videos. For example, in terms of (1), ($\mathcal{B}_t$, inner vortex cycle area) = (100, 100 mm^2) describes a brain activation subregion in the transversal brain section in frames 10 and 75 in **Figure 7**. In terms of (2), Let $\mathcal{B}_s, \mathcal{B}_t, \mathcal{B}_c$ be Betti numbers for the sagittal, transversal and coronal brain regions. The feature vector $(\mathcal{B}_s, \mathcal{B}_t, \mathcal{B}_c)$ is used to describe and compare lag threads in frames across different videos. This is an important advantage that accrues from the application of computational descriptive proximities.

2.4. Methods

This section briefly introduces the method used to derive vortex cycles on triangulated video frames (steps 0 through 5) and their geometric Betti numbers (step 6), which are used to construct persistent barcode for rs-fMRI videos. The fMRI videos (of 688 subjects) used in this work were obtained from the Harvard-MGH Brain Genomics Superstruct Project (Buckner et al., 2014). Each video contains 360 frames that exhibit the propagation of

BOLD signals in the sagittal, transversal and coronal sections of the human brain (see, e.g., the middle row of **Figure 2**). Let K be a rs-fMRI video frame. The steps to obtain triangulation vortexes covering the brain activation regions shown in **Figure 2** are exhibited in the flow graph in **Figure 3**.

$\mathbf{0^0}$ After selecting a video frame, find the centroids (centers of mass) of the dark background regions. Holes are identified by binarizing each video frame.

$\mathbf{1^0}$: Triangulate the centroids of the tiny dark regions inside the brain activation regions in K. A sample centroidal triangle is shown in **Figure 1**.

$\mathbf{2^0}$: Find the barycenter of every centroidal triangle $\triangle$ in K. Each barycenter is a voxel representing a high BOLD signal situated between centroids.

$\mathbf{3^0}$: Connect the barycenters where there is the greatest concentration of centroidal triangles $\triangle$s with a common vertex. Recall that the vertex common to a collection of triangles is an example of an Alexandroff nerve (Alexandroff, 1965) (see, e.g., the collection of centroidal triangles with a common vertex covered by the inner vortex cycle in **Figure 1**). In this work, the focus is on finding maximal Alexandroff nerves.

$\mathbf{4^0}$: Construct vortex cycles on the barycenters of centroidal triangles $\{\triangle\}$ along the boundary of C_0.

$\mathbf{5^0}$: Repeat steps $\mathbf{2^0}$ through $\mathbf{4^0}$ until there are no more vortexes covering subregions containing nonzero BOLD signals. The end result is a collection of connected nesting non-concentric cycles that form a *vortex*. Once these vortexes are generated, the next step is to compute the β_1 for each vortex that has been found in each of the triangulated video frames. Notice that there is usually more than one vortex in video frame.

$\mathbf{6^0}$: Compute Betti number. Count the number of non-single edge (main) cycles in a vortex plus the number of signal edge cycles connected between the main vortex cycles. This process is repeated for each brain section in every frame containing sagittal, trasversal and coronal brain sections in each of the sample rs-fMRI videos.

To construct a Betti number-based persistent barcode, insert a bar in a pictograph (an easy-to-read visualization of brain activity instants accumulated in what is known as a homology barcode Ghrist, 2014), using rs-fMRI video frame number (x-axis) and its corresponding Betti Number (y-axis) (shown in the top half of **Figure 1**).

3. RESULTS

3.1. Part 1. Edge Density and Eigen Value Statistics

Since the focus of this study of brain activation in rs-fMRI videos is on the formation of nesting cycles (vortexes) covering the interior brain activation regions, we consider the edge densities and eigenvalues that reflect the levels of connectedness of these activation region cycles.

3.1.1. Edge Density

In this work, edge density quantifies the connectedness in the vortex representation of a brain activation region in a rs-fMRI video frame. Specifically, ***edge density*** is proportional to the product of the number of vortex bridge segments between cycles $\times$ the number of cycle vertices. Mathematically, we have

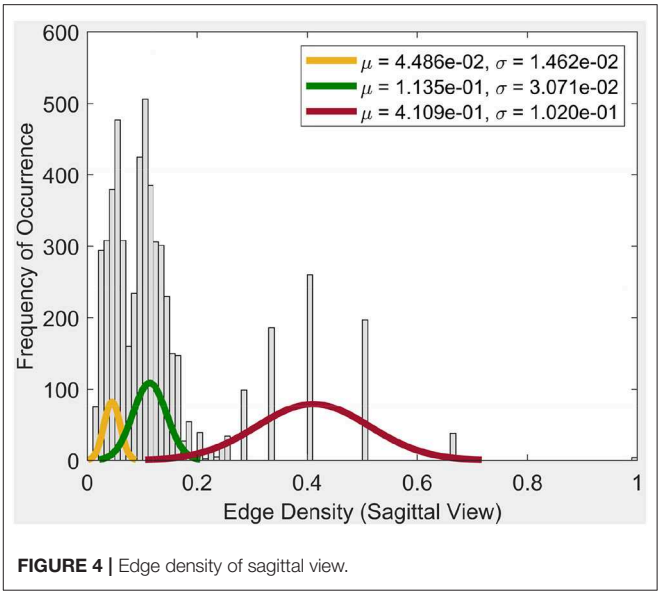

FIGURE 4 | Edge density of sagittal view.

$$E = \text{number of bridge segments between vortex cycles.}$$
$$V = \text{number of vortex cycle vertices.}$$
$$EdgeDensity = 2\left[\frac{E}{V(V-1)}\right].$$

An increase in the number of bridge segments E between vortex cycles results in higher edge density. This can happen in the case where there is more than one bridge segment connected to a vortex cycle. An increase in the number of vortex cycle vertices V with no change in the number of bridge segments, leads to a lower edge density. This occurs whenever the number of dark regions increase in a brain activation region, which leads to an increase in the number of centroidal triangles. This also leads to an increase in the number of barycenters. Physically, each barycenter pinpoints the location of high intensity in a brain activation region. Each vortex cycle vertex represents a barycenter on brain activation region triangle.

Frequency of occurrence in the plot in **Figure 4**, for example, is the number of times a certain edge density appears in the cycles. Whenever a certain edge density appears, the plot will show the frequency of occurrence as 1. If it reappears, the frequency of occurrence will be 2. So it will increase the said value for the occurrence by 1, each time it appears.

From the plot for the Sagittal view in **Figure 4**, it can be seen that each distribution can be represented using three normal distributions. Observe that the left-most normal distribution has a mean $\mu = 0.04$ and a standard deviation $\sigma = 0.014$ but it has a lower edge density compared to the other two (middle, right) distributions. The middle normal distribution has $\mu = 0.11$ and a $\sigma = 0.03$. This distribution has the highest frequency of occurrence of edges. The right normal distribution has a $\mu = 0.41$ and a $\sigma = 0.1$ and has the second highest edge frequency.

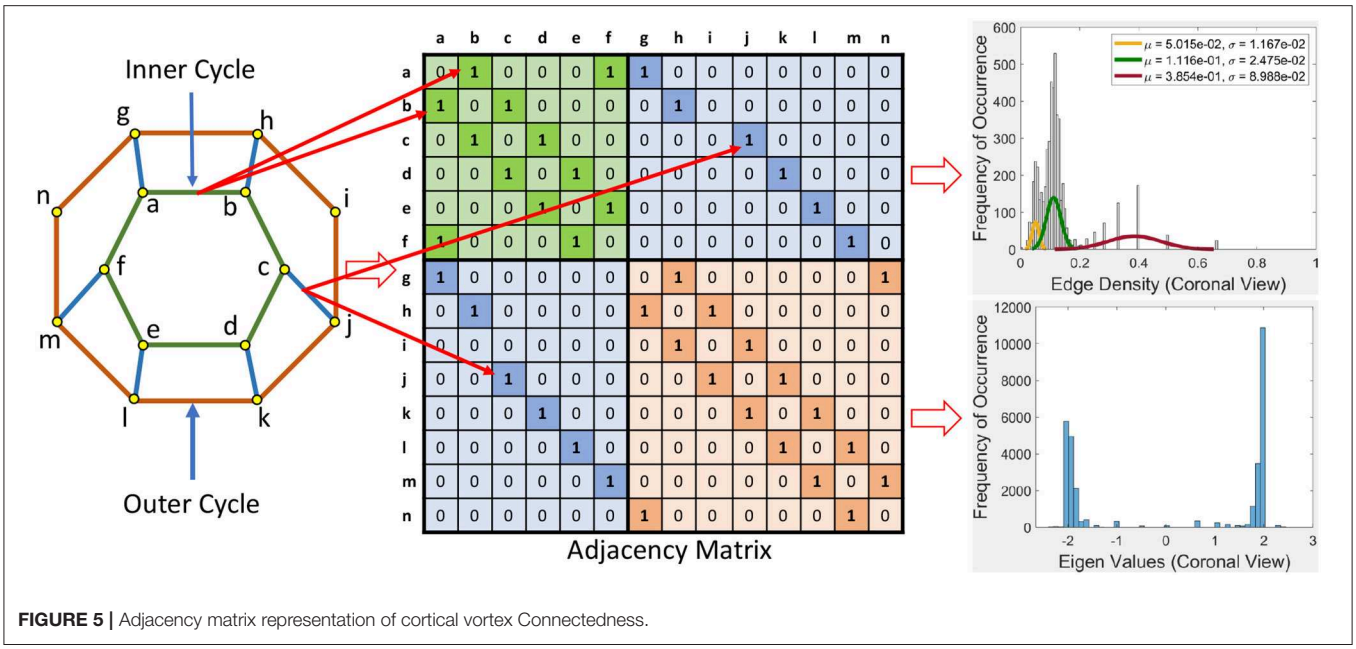

FIGURE 5 | Adjacency matrix representation of cortical vortex Connectedness.

When we consider the cycles that are generated by these edge density values, we can observe the following. In order to get a edge density value close to 0.4, there needs to be single cycle with a lower number of vertices. For example, in the case where there is a single cycle with three vertices (i.e., a triangle), we have the edge density of 0.66. This represents the highest value attainable in vortex cycles. Any single or multiple cycles will have a lower edge density than 0.66. In other words, higher the number of cycles and vertices in a cycle; the lower the edge density. Hence, in **Figure 4**, it can be observed that the highest frequencies of occurrence of edges is for edge densities between 0.1 and 0.2. For further evidence of this, see the plots for transversal and coronal views given in 9 and 10 reported in **Appendix A**.

3.1.2. Eigenvalue Spectrum

This section briefly introduces an eigenvalue spectrum representation of the eigen values derived from the connectivity relations between cortical vortex cycles. A number λ is an ***eigenvalue*** of a linear transformation A (our adjacency matrix), provided there is a vector $x \neq 0$ so that $A(x) = \lambda x$. The vector x called an ***eigenvector***. Eigen values are computed using an ***adjacency matrix*** A, which is a $n \times n$ square matrix where n is the number of vertices in a vortex cycle. Here, an ***eigenvalue spectrum*** is defined by an activation matrix view of the connectedness between vertices in a vortex representation of brain activation regions in rs-fMRI videos. Each adjacency matrix represents the connectivity in vortex cycles. To help visualize the connectivity between vertexes in a vortex, a color-coding scheme is given. That is, the vertices on each vortex cycle edge are color-coded as shown in **Figure 5**. In this color-coding scheme, sub-matrix (a–f) represents inner cycle vertexes (color-coded Green ■). The Sub-matrix (g–n) represents cycle vertexes (color-coded Orange ■). The other two sub matrices represent the bridge segment vertexes (color-coded Blue ■).

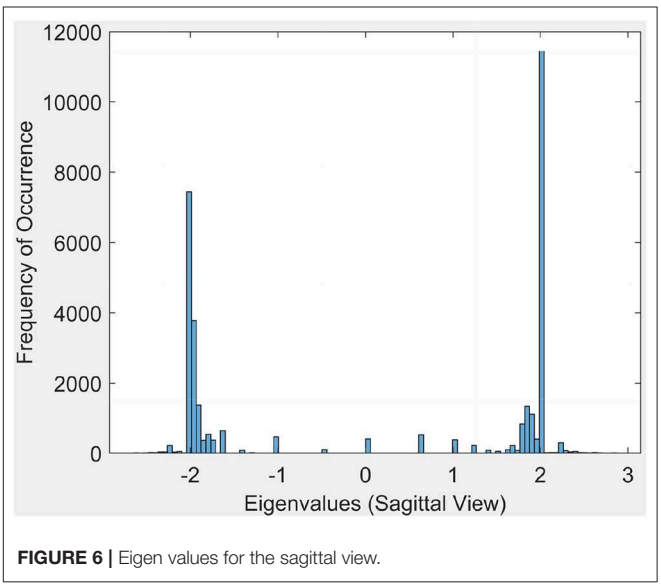

FIGURE 6 | Eigen values for the sagittal view.

A bridge segment between vortex cycles with its ending vertices on neighboring cycles has its connectivity represented by a cell containing 1 in the two sub-matrices (color-coded Blue ■) that are not on the main diagonal as illustrated in **Figure 5**. The bridge segment $\overline{cj}$ is represented by 1s in cells (c,j) and (j,c). If a vortex cycle has an edge between two vertices in locations i and j the cycle, it is represented symmetrically by 1 in a matrix cell (i,j) and a corresponding cell (j,i). For example if there is a edge between vertices 4 and 6 then cells (4, 6) and (6, 4) is 1 in the corresponding adjacency matrix.

Each eigenvalue plot encapsulates results of the sagittal, transversal and coronal views of brain activations recorded in rs-fMRI videos. Each plot represents 12 separate brain section videos = 3 brain sections × 4 original composite videos (i.e., we

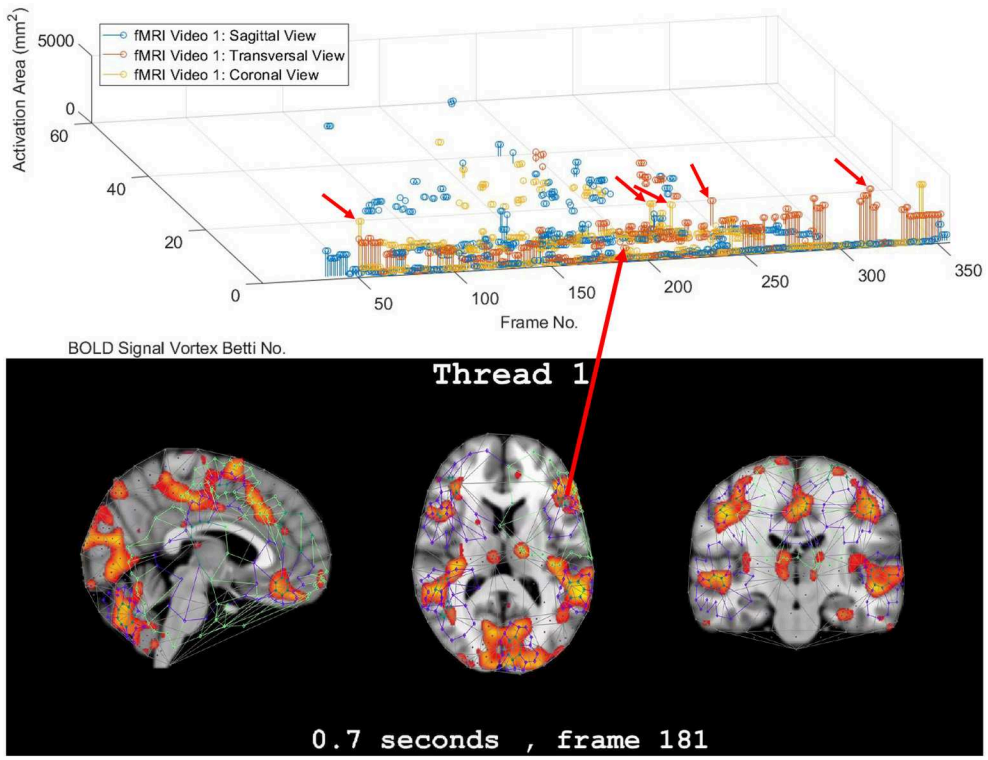

FIGURE 7 | Frame-Betti number-Area for rs-fMRI BOLD signal vortexes on three brain regions.

have extracted three separate brain section videos from each of the four original composite videos). The separated brain section videos made it possible for us to carry out statistical analysis based on the triangulation of the centroids on holes found in the brain activation regions of each brain section in the reported experimental results recorded in the original four videos.

Example 1. In **Figure 5**, connectivity between vertices on vortex cycles are color-coded. For instance, the connectivity between vertexes a and b on the inner cycle in **Figure 5** is represented by a green cell ■ containing a 1 in (b, a) (row b and column a) and by a green cell containing a 1 in (a, b) (row a and column b). ■

Sample results for the eigenvalue spectrum for the sagittal view of brain activation regions is recorded in the plot in **Figure 6**. Plots of the eigenvalue spectrum for the transversal and coronal views are given in **Appendix A**.

3.2. Part 2. Derivation of Persistent Brain Activation Subregion Signature

Each triangulated BOLD signal propagation subregion has a signature defined by the vector (frame number, cycle count Betti number, inner vortex cycle area). This leads to the production of four triangulated rs-fMRI videos available at Don A. et al. (2019), that are part of the University of Manitoba Vortex Signature Project. Vortexes have been derived from each of the triangulation of the BOLD signal activation regions in each of the brain sections in the video frames in the four rs-fMRI videos from Buckner et al. (2014).

A straightforward outcome of the derived vortexes is a rich source of new means of describing individual BOLD signal activation regions as well as a bridge to various forms of descriptions of lag threads. For example, each vortex has a Betti number (count of the number of connected cycles) and various cycle areas. Each of the three brain regions in each frame in the Harvard Brain Genomics rs-fMRI videos has its own vortex and, consequently, its own Betti number. Typically, each video frame will have more than one Betti number derived from vortexes on the multiple brain activation subregions.

In this study, the focus is on the area of the inner cycle of a BOLD signal subregion vortex. This is the case, since each inner cycle lies entirely within the interior of an activation subregion. Hence, an inner cycle area is a reliable approximation of a brain activation area. In sum, geometric Betti numbers and inner vortex cycle area help gauge the extent of an activation subregion. Considered either separately or taken together, a vortex on brain activation subregion provides the basis for a subregion signature, i.e., a distinctive characteristic of a brain activation subregion in a rs-fMRI video frame. For example, (frame number, Betti number) = (140, 60), (200, 60), (210, 60) provides a signature for the coronal brain region in frames 140, 200, and 210 in **Figure 2**. Betti number 60 is an example of a brain activation subregion characteristic that persists over a sequence of video frames. The vector (frame number, Betti number, inner vortex cycle area) =

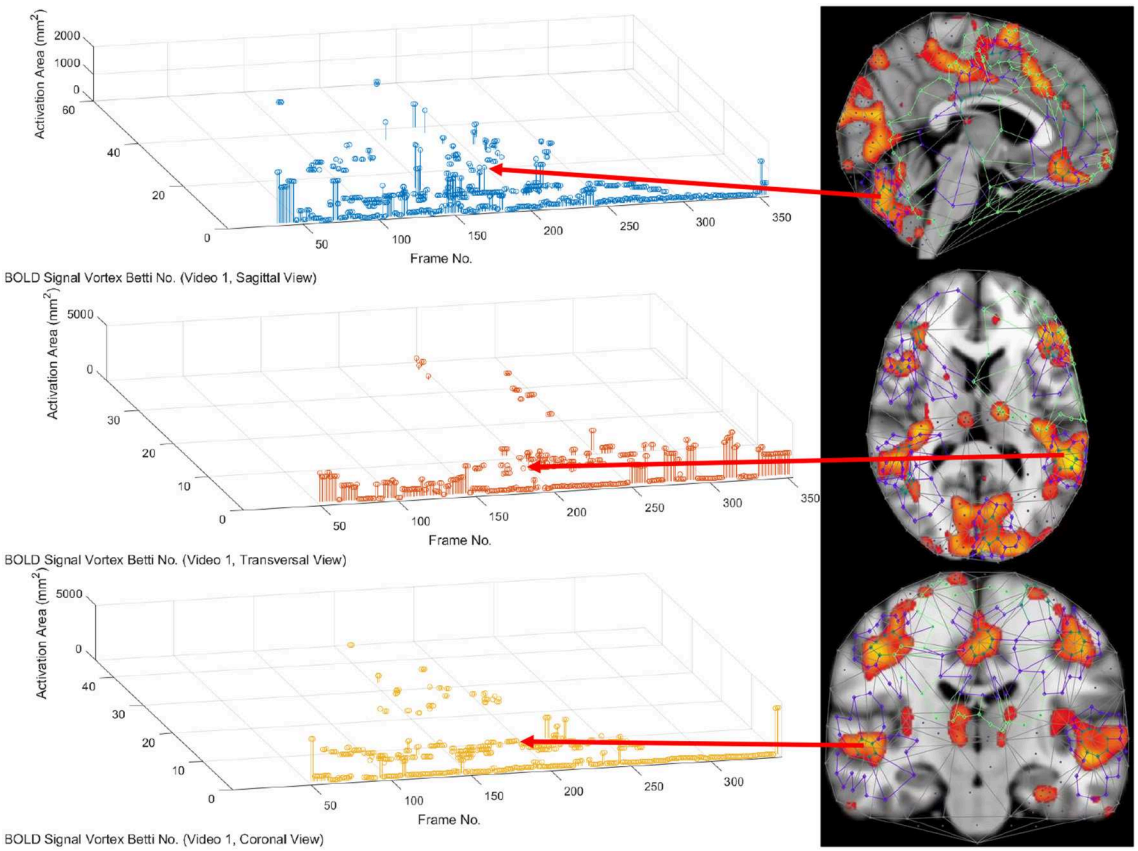

FIGURE 8 | 3D persistence barcode for 3 rs-fMRI BOLD signal propagated in brain regions.

(180, 40, 10 mm^2) depicts a brain activation subregion in the sagittal brain region in **Figure 7**. A repetition of the same Betti number for the same brain region across multiple video frames defines a lag thread pattern. For example, $\mathcal{B}_s = \mathcal{B}_t = \mathcal{B}_c = 100$ for the sagittal, transversal and coronal brain regions defines a lag thread pattern for multiple frames in **Figure 7**.

The gaps between the sequences of contiguous bars are important, since gaps in a particular row of pictograph bars indicates rs-fMRI video frames that do not have the same level of brain activity represented by bars in the row. The proximity of the bars (not necessarily contiguous) in a pictograph row call attention to BOLD signals that are close to each other, temporally. Repeated bars in a pictograph row indicate a repeated (persistent) level of brain activity recorded in rs-fMRI video frames. An example of a pictograph row containing multiple, contiguous bars can be seen in row 30 (Betti numbers = 30) in **Figure 2**. Contiguous bars in a pictograph row indicate the closeness in time of the corresponding brain activity.

Examples of pictograph rows containing multiple, non-contiguous bars can be seen in rows 10 and 60 (Betti numbers = 10 and 60) in **Figure 2**. A byproduct of the inspection of a sequence of contiguous bars (geometric β_1 Betti numbers) in a persistent barcode row can lead to the production of a reduced rs-fMRI video containing only video frames with either activation sub-regions with the same Betti number or a new video containing frames with activation sub-regions, each with a different β_1 (vortex cycle count) Betti number.

3.3. Part 3. Confirmation of Highly Reproducible Lag Thread Topography

From 3D barcodes in **Figure 7** as well as in **Figure 8**, Betti number-area patterns can be detected within frames in the same video. That is, one can find many examples of brain activation subregion Betti number (and corresponding subregion area in a lag thread in one video frame) that are reproduced in a lag thread in a different video frame. In other words, the Betti number-area combination persists across different frames. Many examples of this Betti number-area persistence phenomenon can be detected in the two sample 3D homology barcodes when compared with similar 3D homology barcodes derived from the frames in the videos available at the University of Manitoba Vortex Project at Don A. et al. (2019). These persistent repetitions in the topography detected in different lag threads confirm the observation that there are commonalities in signal propagation within each lag thread (Mitra et al., 2015).

3.4. Summary of Findings

This study of the Betti numbers and inner vortex areas of rs-fMRI BOLD signal propagation subregions of the brain confirms and supplements earlier findings given in Mitra et al. (2015). A main

result of this study of the persistence of brain activation subregion features confirms the contention that the topography of lag threads is highly reproducible. Starting with the Betti number of connected cycles derived from triangulated brain activation regions found in rs-fMRI videos, it is apparent that the vortexes on brain activation subregions appear over and over in the lag threads across different rs-fMRI video frames. In other words, we find that there are commonalities in BOLD signal propagation within each lag thread.

The question whether intrinsic brain activity contains reproducible temporal sequences is revisited. It is confirmed that a human resting state fMRI (rs-fMRI) contains persistent (repeatable) highly reproducible lag structure. The answer to this question is given twofold. This is done first in terms of Betti numbers that are counts of the number of connected cycles in vortexes on triangulated brain activation subregions. We introduce a 2D persistence pictograph (barcode) that exhibits the appearance, disappearance, and repeated reappearance of Betti numbers across lag threads in sequences of rs-fMRI video frames. In addition, the reproducibility question is also answered in terms of the introduction of a video frame-Betti number-vortex cycle area combination in 3D persistence barcodes that facilitates a check on how often these features of lag threads appear during a rs-fMRI session.

Concluding Remarks

This study considers Betti numbers that are counts of the number of connected barycentric cycles in vortexes on triangulated brain activation regions. In terms of the area occupied by a vortex on a brain activation subregion, we have only considered the area of the interior of the innermost barycentric cycle of each vortex. Also of interest and of considerable importance is the area in the interior of other cycles that includes the inner vortex cycle. Future work would expand the derivation of persistent barcodes to include the zeroth as well as the oneth Betti numbers. In working toward the approximation of the area of brain activation subregion shapes, the areas in the interior of the other cycles (summing on the innermost vortex cycle area) would be considered.

DATA AVAILABILITY STATEMENT

The dataset for this study can be found at the following URL: https://drive.google.com/drive/folders/1NxZ1Ydcuhzrdmgav-DxIGrbg-pq5GCI8?usp=sharing.

AUTHOR CONTRIBUTIONS

JP designed the materials and methods. AD developed the code and tools. AT and JP analyzed the data and performed the research. SR contributed to the research and writing of the paper.

FUNDING

This research has been supported by the Natural Sciences & Engineering Research Council of Canada (NSERC) discovery grants 185986 and 194376, Instituto Nazionale di Alta Matematica (INdAM) Francesco Severi, Gruppo Nazionale per le Strutture Algebriche, Geometriche e Loro Applicazioni grant 9 920160 000362, n.prot U 2016 / 000036.

ACKNOWLEDGMENTS

Authors acknowledge the NSERC Discovery Grants program for this research. This manuscript has been released as a Pre-print at https://www.biorxiv.org/content/10.1101/809293v2.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncom.2020.00034/full#supplementary-material

REFERENCES

Achard, S., and Bullmore, E. (2007). Efficiency and cost of economical brain functional networks. *PLoS Comput. Biol.* 3:e30017. doi: 10.1371/journal.pcbi.0030017

Alexandroff, P. (1965). *Elementary Concepts of Topology*. New York, NY: Dover Publications, Inc.

Buckner, R. L., Roffman, J. L., and Smoller, J. W. (2014). Brain Genomics Superstruct Project (GSP). *Harvard Dataverse*. doi: 10.7910/DVN/25833

Bullmore, E., and Sports, O. (2009). Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* 10, 186–198. doi: 10.1038/nrn2575

Carlsson, G. (2009). Topology and data. *Bull. Am. Math. Soc.* 46, 255–308. doi: 10.1090/S0273-0979-09-01249-X

Cassidy, B., Rae, C., and Solo, V. (2015). "Brain activity: conditional dissimilarity and persistent homology," in *ISBI, 12th International Symposium on Biomedical Imaging* (New York, NY: IEEE), 1356–1359.

Chazal, F., and Michel, B. (2017). An introduction to topological data analysis: fundamental and practical aspects for data scientists. *arXiv* 1710, 1–38. doi: 10.1142/9789813144446_0001

Chung, M., Hanson, J., Ye, J., Davidson, R., and Pollak, S. (2015). Persistent homology in sparse regression and its application to brain morphometry. *IEEE Trans. Med. Imaging* 34, 1928–1939. doi: 10.1109/TMI.2015.2416271

Chung, M., Lee, H., DiChristofano, A., Ombao, H., and Solo, V. (2019). Exact topological inference of the resting-state brain networks in twins. *Netw. Neurosci.* 3, 674–694. doi: 10.1162/netn_a_00091

Chung, M., Vilalta-Gil, V., Lee, H., Rathouz, P., Lahey, B., and Zald, D. (2017). "Exact topological inference for paired brain networks via persistent homology. Information processing in medical imaging," in *IPMI, Lecture Notes in Computer Science, 10265* (Springer), 299–310.

Delaunay, B. (1934). Sur la sphère vide. *Izvestia Akad. Nauk SSSR, Otdelenie Matematicheskii i Estestvennyka Nauk* 7, 793–800.

Di Concilio, A., Guadagni, C., Peters, J., and Ramanna, S. (2018). Descriptive proximities. Properties and interplay between classical proximities and overlap. *Springer Math. Comput. Sci.* 12, 91–106. doi: 10.1007/s11786-017-0328-y

Don, A., Peters, J., Ramanna, S., and Tozzi, A. (2019). *Vortex Signature Project.* Available online at: https://drive.google.com/drive/folders/1NxZ1Ydcuhzrdmgav-DxIGrbg-pq5GCI8?usp=sharing

Don, A. P., and Peters, J. F. (2019). Ghrist barcoded video frames. Application in detecting persistent visual scene surface shapes captured in videos. *Theory Appl. Math. Comput. Sci.* 9, 14–27. doi: 10.1007/978-3-030-22192-8

Don, A. P., Peters, J. F., Ramanna, S., and Tozzi, A. (2019). Topological inference from spontaneous activity structures in FMRI videos with persistence barcodes. *Cold Spring Harb. Lab.* 2019:809293. doi: 10.1101/809293

Edelsbrunner, H., and Harer, J. (2008). Persistent homology. A survey. *Contemp. Math.* 453, 257–282. doi: 10.1090/conm/453/08802

Edelsbrunner, H., and Harer, J. (2010). *Computational Topology. An Introduction*. Providence, RI: American Mathematical Society.

Edelsbrunner, H., Letscher, D., and Zomorodian, A. (2000). "Topological persistence and simplification," in *41st Annual Symposium on Foundations of Computer Science*, (Los Alamitos, CA: IEEE Comput. Soc. Press), 454–463.

Edelsbrunner, H., Letscher, D., and Zomorodian, A. (2001). Topological persistence and simplification. *Springer Discr. Comput. Geometry* 28, 511–533. doi: 10.1007/s00454-002-2885-2

Freeman, W. (2009). Vortices in brain activity: their mechanism and significance for perception. *Neural Netw.* 22, 491–501. doi: 10.1016/j.neunet.2009.06.050

Ghrist, R. (2008). Barcodes: the persistent topology of data. *Bull. Am. Math. Soc.* 45, 61–75. doi: 10.1090/S0273-0979-07-01191-3

Ghrist, R. (2014). *Elementary Applied Topology*. Philadelphia, PA: University of Pennsylvania.

Giusti, C., Pastalkova, E., Curto, C., and Itskov, V. (2015). Clique topology reveals intrinsic geometric structure in neural correlations. *Proc. Nat. Acad. Sci. U.S.A.* 112, 1355–1360. doi: 10.1073/pnas.1506407112

Lee, H., Chung, M., Kang, H., Kim, B.-N., and Lee, D. (2011a). "Computing the shape of brain networks using graph filtration and Gromov-Hausdor metric," in *MICCAI, Lecture Notes in Computer Science, 6892* (Toronto, ON: Springer), 302–309.

Lee, H., Chung, M., Kang, H., Kim, B.-N., and Lee, D. (2011b). "Discriminative persistent homology of brain networks," in *ISBI, IEEE International Symposium on Biomedical Imaging* (Chicago, IL), 841–844.

Matsui, T., Murakami, T., and Ohki, K. (2016). Transient neuronal coactivations embedded in globally propagating waves underlie resting-state functional connectivity. *Proc. Nat. Acad. Sci. U.S.A.* 113, 6556–6561. doi: 10.1073/pnas.1521299113

Meunier, D., Lambiotte, R., Fornito, A., Ersche, K., and Bullmore, E. (2009). Hierarchical modularity in human brain functional networks. *Front. Neuroinform.* 3:2009. doi: 10.3389/neuro.11.037.2009

Mitra, A., Snyder, A. Z., Blazey, T., and Raichle, M. (2015). Lag threads organize the brain's intrinsic activity. *Proc. Nat. Acad. Sci. U.S.A.* 112, E2235–E2244. doi: 10.1073/pnas.1503960112

Munkres, J. (2000). *Topology, 2nd Edn.* Englewood Cliffs, NJ: Prentice-Hall.

Park, Y. B., Shim, M. M., James, O., and Park, H. (2019). Possible links between the lag structure in visual cortex and visual streams using fMRI. *Sci. Rep.* 9, 4283–4293. doi: 10.1038/s41598-019-40728-x

Peters, J. (2020). *Computational Geometry, Topology and Physics of Digital Images with Applications. Shape Complexes, Optical Vortex Nerves and Proximities*. Cham: Springer Int. Pub.

Peters, J., Ramanna, S., Tozzi, A., and Inan, E. (2017). Bold-independent computational entropy assesses functional donut-like structures in brain fMRI images. *Front. Hum. Neurosci.* 11:38. doi: 10.3389/fnhum.2017.00038

Petri, G., Expert, P., Turkheimer, F., Carhart-Harris, R., Nutt, D., Hellyer, P. J., et al. (2014). Homological scaffolds of brain functional networks. *J. R. Soc. Interface* 11:20140873. doi: 10.1098/rsif.2014.0873

Sizemore, A. E., Giusti, C., Kahn, A., Vettel, J., Betzel, R., and Bassett, D. (2018). Cliques and cavities in the human connectome. *J. Comput. Neurosci.* 44, 115–145. doi: 10.1007/s10827-017-0672-6

van Wijk, B., Stam, C., and Daffertshofer, A. (2010). Comparing brain networks of different size and connectivity density using graph theory. *PLoS ONE* 5:e13701. doi: 10.1371/journal.pone.0013701

Yung, C., Choi, G.-T., Chen, K., and Lui, L. (2016). Trim: triangulating images for efficient registration. *arXiv cs.GR* 1605, 1–13. doi: 10.1016/j.jvcir.2018.07.005

Zomorodian, A., and Carlsson, G. (2005). Computing persistent homology. *Discrete Comput. Geomet.* 33, 249–274. doi: 10.1007/s00454-004-1146-y

Zomorodian, A. F. (2001). *Computing and comprehending topology: persistence and hierarchical morse complexes* (Ph.D. thesis). Department of Computer Science; University of Illinois at Urbana-Champaign, Chicago, IL., United States.

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.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 08 May 2020
doi: 10.3389/fncom.2020.00040

Exponential Time Differencing Algorithm for Pulse-Coupled Hodgkin-Huxley Neural Networks

*Zhong-qi Kyle Tian and Douglas Zhou**

School of Mathematical Sciences, MOE-LSC, Institute of Natural Sciences, Shanghai Jiao Tong University, Shanghai, China

The exponential time differencing (ETD) method allows using a large time step to efficiently evolve stiff systems such as Hodgkin-Huxley (HH) neural networks. For pulse-coupled HH networks, the synaptic spike times cannot be predetermined and are convoluted with neuron's trajectory itself. This presents a challenging issue for the design of an efficient numerical simulation algorithm. The stiffness in the HH equations are quite different, for example, between the spike and non-spike regions. Here, we design a second-order adaptive exponential time differencing algorithm (AETD2) for the numerical evolution of HH neural networks. Compared with the regular second-order Runge-Kutta method (RK2), our AETD2 method can use time steps one order of magnitude larger and improve computational efficiency more than ten times while excellently capturing accurate traces of membrane potentials of HH neurons. This high accuracy and efficiency can be robustly obtained and do not depend on the dynamical regimes, connectivity structure or the network size.

Keywords: Hodgkin-Huxley, exponential time differencing method, efficiency, pulse-coupled, second-order

OPEN ACCESS

Edited by:
Thomas Nowotny,
University of Sussex, United Kingdom

Reviewed by:
Lianchun Yu,
Lanzhou University, China
Marcel Stimberg,
INSERM U968 Institut de la Vision,
Université de la Sorbonne, France

****Correspondence:***
Douglas Zhou
zdz@sjtu.edu.cn

Received: *20 November 2019*
Accepted: *16 April 2020*
Published: *08 May 2020*

Citation:
Tian ZK and Zhou D (2020) Exponential Time Differencing Algorithm for Pulse-Coupled Hodgkin-Huxley Neural Networks. Front. Comput. Neurosci. 14:40. doi: 10.3389/fncom.2020.00040

1. INTRODUCTION

The Hodgkin-Huxley (HH) model (Hodgkin and Huxley, 1952; Hassard, 1978; Dayan and Abbott, 2003) is a classical neuron model, originally proposed to describe the behaviors of action potentials of the squid's giant axon. It provides a useful mechanism that accounts for the detailed generation of action potentials and the existence of the absolute refractory periods. It also serves as the foundation for other neuron models such as the one that can describe the behaviors of bursting and adaption (Pospischil et al., 2008). However, the HH equations are so complicated that it is difficult to study its properties analytically such as the Hopf bifurcation and chaotic dynamics (Aihara, 1986; Hansel and Sompolinsky, 1996; Guckenheimer and Oliva, 2002; Lin, 2006). Therefore, its investigation often relies on numerical simulations, for example, by the Runge-Kutta (RK) methods.

There are several difficulties to design an efficient and accurate numerical algorithm for the HH neural network, especially when the network size is large. First, when an HH neuron driven by external input generates an action potential (the interval of action potential is called spike period in this work), the HH neuron equations become stiff. Regular RK methods have to use very small time step to satisfy the requirement of numerical stability (Guckenheimer and Oliva, 2002; Börgers et al., 2005; Kassam and Trefethen, 2005; Börgers and Nectow, 2013). This small time step will significantly increase the computational cost when studying long time behavior of large-scale HH networks such as chaotic attractor dynamics or collecting reliable statistical information of HH neurons such as the distribution of inter-spike-intervals.

For more realistic situations, the neurons are generally driven by stochastic spike input and the interaction term is usually modeled by a Dirac delta function (pulse-coupled), while the spike-induced conductance dynamics are modeled by an *alpha* function (Somers et al., 1995; Hansel et al., 1998; Sun et al., 2009). These make the system become non-smooth and event-driven, while providing challenges for the design of efficient numerical simulation algorithms. For instance, it is impossible to predetermine the synaptic spike times since they are convoluted with neurons' trajectories themselves. As a result, one has to evolve the HH network by ignoring the spike interactions among neurons and then use spike-spike interaction to amend the neurons' trajectories at the end of the time step (Hansel et al., 1998; Brette et al., 2007). Without a careful recalibration for the neuronal spikes, the numerical algorithm often suffers from the issue of instability or relatively low numerical accuracy.

The exponential time differencing (ETD) method (Hochbruck et al., 1998; Cox and Matthews, 2002; Kassam and Trefethen, 2005; de la Hoz and Vadillo, 2008; Nie et al., 2008; Hochbruck and Ostermann, 2010) is proposed for efficient simulation of stiff ordinary differential equations (ODEs). The basic idea is to decompose the ODEs into a linear stiff part and a nonlinear non-stiff part. Then, the linear stiff part can be solved by using the integrating factor method, while the nonlinear non-stiff part can be approximated by numerical quadrature (Cox and Matthews, 2002). A second-order ETD (ETD2) method for HH neural networks has been proposed in a recent work (Börgers and Nectow, 2013), which allows using a large time step to raise computational efficiency. In Börgers and Nectow (2013), the HH equations are linearly approximated in each time step, and then solved analytically over the time step. The ETD2 method proposed in Börgers and Nectow (2013) is a reduced situation of that in Cox and Matthews (2002), but it is difficult to generalize to higher-order cases, e.g., the fourth-order ETD method. Besides, although the ETD2 method proposed in Börgers and Nectow (2013) is proven to be unconditionally stable for HH system, it will be inaccurate using a large time step (Börgers and Nectow, 2013).

In this work, we first provide an ETD2 method following the idea proposed in Cox and Matthews (2002) to evolve a pulse-coupled HH neural network driven by stochastic spike input. Note that the stiffness of HH equations are quite different, especially between the spike and non-spike periods, and we find that the ETD2 method may introduce a relatively large error in the membrane potentials in the non-stiff period if using the same time step as that in the stiff period. We then design an adaptive ETD2 method (AETD2) that using different decompositions of the linear and non-linear parts in stiff and non-stiff periods. In addition, for the situation where neurons generate spikes in the time step, the effects of the spikes are carefully recalibrated in our AETD2 method to achieve a second-order numerical accuracy. Our AETD2 method is capable of using a large time step, while achieving the same high accurate traces of membrane potential of each neuron as the second-order RK (RK2) method using a very small time step. It can improve computational efficiency more than one order of magnitude compared with the RK2 method. This high numerical accuracy and computational efficiency can be achieved over a wide range of dynamical regimes and does not depend on the network connectivity or size.

2. MATERIALS AND METHODS

2.1. The Model

The dynamics of the ith neuron of an HH neural network is governed by

$$C\frac{dV_i}{dt} = -(V_i - V_{Na})G_{Na}m_i^3h_i - (V_i - V_K)G_Kn_i^4 - (V_i - V_L)G_L + I_i^{\text{input}}, \quad (1)$$

$$\frac{dz_i}{dt} = (1 - z_i)\alpha_z(V_i) - z_i\beta_z(V_i), \quad \text{for } z = m, h, n, \quad (2)$$

where C is the cell membrane capacitance, V_i is the membrane potential, m_i, h_i, and n_i are gating variables for sodium and potassium currents, respectively (Dayan and Abbott, 2001). The parameters V_{Na}, V_K, and V_L are the reversal potentials for the sodium, potassium, and leak currents, respectively, G_{Na}, G_K, and G_L are the corresponding maximum conductances. The form of α_z and β_z are set as (Dayan and Abbott, 2001): $\alpha_m(V) = (0.1V+4)/(1-\exp(-0.1V-4))$, $\beta_m(V) = 4\exp(-(V+65)/18)$, $\alpha_h(V) = 0.07\exp(-(V+65)/20)$, $\beta_h(V) = 1/(1+\exp(-3.5-0.1V))$, $\alpha_n(V) = (0.01V+0.55)/(1-\exp(-0.1V-5.5))$, and $\beta_n(V) = 0.125\exp(-(V+65)/80)$.

The input current I_i^{input} is given by

$$I_i^{\text{input}} = -G_i^E(t)(V_i - V_G^E) - G_i^I(t)(V_i - V_G^I), \quad (3)$$

where G_i^E and G_i^I are excitatory and inhibitory conductances, respectively, V_G^E and V_G^I are the corresponding reversal potentials. The dynamics of conductance G_i^Q, $Q = E, I$, is governed by

$$\frac{dG_i^Q}{dt} = -\frac{G_i^Q}{\sigma_r^Q} + H_i^Q, \quad (4)$$

$$\frac{dH_i^Q}{dt} = -\frac{H_i^Q}{\sigma_d^Q} + F^Q\sum_l \delta(t - s_{il}) + \sum_j S_{ij}^Q\sum_l \delta(t - \tau_{jl}), \quad (5)$$

where H_i^Q is an auxiliary dynamical variable to make the conductance G_i^Q as a continuous function, $\delta(\cdot)$ is the Dirac delta function, s_{il} is the spike time of the feedforward Poisson input with strength F^Q and rate ν, τ_{jl} is the lth spike time of the jth neuron, and σ_d^Q and σ_r^Q are slow decay and fast rise time scale, respectively. Each neuron is either excitatory or inhibitory and its coupling strength is labeled by its type E or I, respectively. For example, S_{ij}^E (S_{ij}^I) is the coupling strength from the jth excitatory (inhibitory) neuron to its postsynaptic ith neuron. By analytically solving Equations (4) and (5), the spike-induced conductance change can be explicitly expressed as

$$\mathbb{G}(\sigma_d^Q, \sigma_r^Q, t) = \frac{\sigma_d^Q\sigma_r^Q}{\sigma_d^Q - \sigma_r^Q}(e^{-t/\sigma_d^Q} - e^{-t/\sigma_r^Q})\Theta(t), \quad (6)$$

where $\Theta(\cdot)$ is the Heaviside function. For all neurons, we take $F^E = f$ and $F^I = 0$. The model parameters are $C = 1\mu\text{F}\cdot\text{cm}^{-2}$, $V_{Na} = 50$ mV, $V_K = -77$ mV, $V_L = -54.387$ mV, $G_{Na} = 120\ \text{mS}\cdot\text{cm}^{-2}$, $G_K = 36\ \text{mS}\cdot\text{cm}^{-2}$, $G_L = 0.3\ \text{mS}\cdot\text{cm}^{-2}$, $V_G^E = 0$ mV, $V_G^I = -80$ mV, $\sigma_r^E = 0.5$ ms, $\sigma_d^E = 3.0$ ms, $\sigma_r^I = 0.5$ ms, and $\sigma_d^I = 7.0$ ms (Dayan and Abbott, 2001).

The voltage V_i evolves continuously according to Equations (1) and (2). When it reaches the firing threshold $V^{\text{th}} = -50$ mV (Sun et al., 2009), we say the ith neuron generates a spike at this time, say τ_{il}. Then it will trigger its postsynaptic jth neuron's conductance change in the form of $S_{ji}^Q \mathbb{G}(\sigma_d^Q, \sigma_r^Q, t-\tau_{il})$, $Q = E, I$. For the ease of discussion about our algorithm design, we consider an all-to-all connected network with $S_{ij}^Q = S/N$, where $Q = E, I$, S is the coupling strength, and N is the total number of neurons in the network. Note that our algorithm can be easily extended to networks with more complicated connectivity structure.

2.2. Runge-Kutta Method

Without loss of generality, we consider the RK2 method as the benchmark and compare it with the ETD methods. We first introduce the RK2 method to evolve the HH neural network with a fixed time step Δt, for example, to evolve the system from time $t = t_k = k\Delta t$ to $t = t_{k+1} = (k+1)\Delta t$. Since the synaptic spike times in $[t_k, t_{k+1}]$ can not be predetermined, one has to evolve the network without considering synaptic spike interactions and reconsider their effects by using spike-spike interactions at the end of time step (Hansel et al., 1998; Brette et al., 2007).

Due to the pulse-coupled dynamics in Equation (5), the numerical accuracy may be very low if the spike timing is not well estimated. For example, suppose that a presynaptic spike fired at $\tilde{t}$ between t_k and t_{k+1}. If one simply assigns it to be the end of time step t_{k+1}, then the error of the spike-induced conductance change is

$$\begin{aligned}\frac{S}{N}[\mathbb{G}(\sigma_d^Q, \sigma_r^Q, t-\tilde{t}) - \mathbb{G}(\sigma_d^Q, \sigma_r^Q, t-t_{k+1})] &= O(t_{k+1}-\tilde{t}) \quad (7)\\ &= O(\Delta t), Q = E, I.\end{aligned}$$

Therefore, the error with the magnitude of Δt will be introduced when the system evolves to $t = t_{k+1}$.

We now solve the above issue arising from the pulse-coupled dynamics to achieve a second-order numerical accuracy. First, we evolve the HH neural network without considering the feedforward and synaptic spikes during the time interval $[t_k, t_{k+1}]$. Then, at time $t = t_{k+1}$, some neuron's voltage may be above the threshold, i.e., generating a spike, say neuron i, if $V_{i,k} < Vth$ and $V_{i,k+1} \geq Vth$ where $V_{i,k}$ and $V_{i,k+1}$ represent $V_i(t_k)$ and $V_i(t_{k+1})$, respectively. The spike time, say τ_{il}, can be estimated following the idea proposed in Hansel et al. (1998) and Shelley and Tao (2001). The neuron's membrane potential during the time interval can be approximated by a linear interpolation:

$$V_i(t) \approx V_{i,k} + \frac{V_{i,k+1} - V_{i,k}}{\Delta t}(t - t_k), \quad (8)$$

and the spike time τ_{il} can be estimated by solving the equation:

$$V^{\text{th}} = V_{i,k} + \frac{V_{i,k+1} - V_{i,k}}{\Delta t}(\tau_{il} - t_k). \quad (9)$$

Since there may be some neurons firing and some feedforward spikes emitting during the time interval and they will induce the conductance change, the conductance should be then recalibrated. When the neuron firing and the feedforward spikes are not considered, the conductance variables in such cases, denoted by $\tilde{G}^Q$ and $\tilde{H}^Q$, $Q = E, I$, are

$$\tilde{H}_{j,k+1}^Q = H_{j,k}^Q e^{-\Delta t/\sigma_d^Q}, \quad (10)$$

$$\tilde{G}_{j,k+1}^Q = G_{j,k}^Q e^{-\Delta t/\sigma_r^Q} + H_{i,k}^Q \mathbb{G}(\sigma_d^Q, \sigma_r^Q, \Delta t), \quad (11)$$

and the conductance variables are then recalibrated by taking into account the neuron firing and the feedforward spikes as

$$\begin{aligned}H_{j,k+1}^Q = \tilde{H}_{j,k+1}^Q &+ F^Q \sum_{t_k < s_{jl} \leq t_{k+1}} e^{-(t_{k+1}-s_{jl})/\sigma_d^Q}\\ &+ \frac{S}{N}\sum_i \sum_{t_k < \tau_{il} \leq t_{k+1}} e^{-(t_{k+1}-\tau_{il})/\sigma_d^Q}, \quad (12)\end{aligned}$$

$$\begin{aligned}G_{j,k+1}^Q = \tilde{G}_{j,k+1}^Q &+ F^Q \sum_{t_k < s_{jl} \leq t_{k+1}} \mathbb{G}(\sigma_d^Q, \sigma_r^Q, t_{k+1} - s_{jl})\\ &+ \frac{S}{N}\sum_i \sum_{t_k < \tau_{il} \leq t_{k+1}} \mathbb{G}(\sigma_d^Q, \sigma_r^Q, t_{k+1} - \tau_{il}), \quad (13)\end{aligned}$$

for $j = 1, 2, ..., N$. A detailed algorithm of the RK2 method is given in Algorithm 1.

Algorithm 1: RK2 algorithm

Input: an initial time t_k and feedforward input times $\{s_{il}\}$
Output: Solutions at time t_{k+1}

```
for i = 1 to N do
    Solve the HH equations for the ith neuron without
    considering spike input using RK2 scheme.
    if V_i(t_k) < V^th and V_i(t_{k+1}) ≥ V^th then
        The ith neuron spiked in [t_k, t_{k+1}].
        Estimate the spike time τ_il by Equation (9).
    end
end
Recalibrate the conductance by Equations (12) and (13).
```

We show that the above algorithm can indeed achieve a second-order numerical accuracy as follows. If there are no feedforward or synaptic spikes, then all the dependent variables are infinitely differentiable and the RK2 method can achieve an error of order $O(\Delta t^2)$. For the time step that contains feedforward or synaptic spikes, an error of order $O(\Delta t)$ is introduced in the conductance with the form of $G^Q - \tilde{G}^Q, Q = E, I$. Nevertheless, the dependent variables of V, m, h, and n can have an error of order $O(\Delta t^2)$. The synaptic spike times are

estimated by a linear interpolation and also have an error of order $O(\Delta t^2)$. After recalibration shown in Equations (12) and (13), the conductance variables G^Q and H^Q, $Q = E, I$ can achieve numerical accuracy of second-order at the end of the time step. Therefore, all the dependent variables $V, m, h, n, G^E, G^I, H^E$, and H^I that are numerically solved in each time step have an error of order $O(\Delta t^2)$ (see below for verification of numerical results).

2.3. Exponential Time Differencing Method

Exponential time differencing method is proposed to solve the stiff problem in differential equations by decomposing the system into a linear stiff term and a nonlinear non-stiff term (Hochbruck et al., 1998; Cox and Matthews, 2002; Kassam and Trefethen, 2005; Nie et al., 2008). Following this idea, we propose the ETD schemes for HH Equations (1) and (2) below. As illustrated in Algorithm 1, each neuron in the HH network is evolved independently and their conductances are recalibrated at the end of time step. Thus, one can first derive an ETD scheme for a single HH neuron and then consider the spike interactions among neurons, and obtain an ETD scheme for the numerical evolution of an HH neural network.

Consider the evolution of a single HH neuron from t_k to t_{k+1}. The first step of the ETD method is to rewrite Equations (1) and (2) as

$$\frac{dz}{dt} = c_z z + F_z, \quad \text{for } z = V, m, h, n, \tag{14}$$

where

$$c_V = (-G_{Na} m_k^3 h_k - G_K n_k^4 - G_L)/C, \tag{15}$$

$$c_z = -\alpha_z(V_k) - \beta_z(V_k), \quad \text{for } z = m, h, n, \tag{16}$$

$$\begin{aligned} F_V(t, V, m, h, n) = \big[&-(V - V_{Na}) G_{Na} m^3 h - (V - V_K) G_K n^4 \\ &-(V - V_L) G_L + I^{\text{input}}\big]/C - c_V V \end{aligned} \tag{17}$$

and

$$F_z(t, V, m, h, n) = (1 - z)\alpha_z(V) - z\beta_z(V) - c_z z, \quad \text{for } z = m, h, n, \tag{18}$$

where z_k represents $z(t_k)$ for $z = V, m, h, n$ of this neuron. Here, $F_z(t, V, m, h, n)$ is actually a function of t, V, and z for $z = m, h, n$, but we write in this way for ease of illustration. Note that the linear coefficient c_z in Equation (14) is a constant value in the kth time step $[t_k, t_{k+1}]$ and is updated with respect to k. Multiplying Equation (14) by an integrating factor $e^{-c_z t}$ and taking integral from t_k to t_{k+1}, we obtain

$$\begin{aligned} z_{k+1} = \; & z_k e^{c_z \Delta t} + e^{c_z \Delta t} \int_0^{\Delta t} e^{-c_z \tau} F_z(t_k \\ & + \tau, V(t_k + \tau), m(t_k + \tau), h(t_k + \tau), n(t_k + \tau)) d\tau \end{aligned} \tag{19}$$

for $z = V, m, h$, and n.

The next step of the ETD method is to derive proper approximations to the above integration. We take a second-order ETD formula with RK time stepping which was proposed as ETD2RK method in Cox and Matthews (2002). Let,

$$a_{z,k} = z_k e^{c_z \Delta t} + F_{z,k}(e^{c_z \Delta t} - 1)/c_z, \tag{20}$$

and approximate F_z during the time interval $[t_k, t_{k+1}]$ by

$$\begin{aligned} &F_z(t_k + \tau, V(t_k + \tau), m(t_k + \tau), h(t_k + \tau), n(t_k + \tau) \\ &= F_{z,k} + \tau(F_z(t_{k+1}, a_{V,k}, a_{m,k}, a_{h,k}, a_{n,k}) - F_{z,k})/\Delta t + O(\Delta t^2), \end{aligned} \tag{21}$$

for $z = V, m, h$, and n, where $F_{z,k}$ represents $F_z(t_k, V_k, m_k, h_k, n_k)$. Substituting the above approximation into Equation (19) yields the ETD2 scheme (the ETD method which has second-order numerical accuracy) which is given by

$$\begin{aligned} z_{k+1} \; = \; & a_{z,k} + [F_z(t_{k+1}, a_{V,k}, a_{m,k}, a_{h,k}, a_{n,k}) \\ & - F_{z,k}](e^{c_z \Delta t} - 1 - c_z \Delta t)/c_z^2 \Delta t, \end{aligned} \tag{22}$$

for $z = V, m, h$, and n. The procedure of the ETD2 algorithm for an HH neural network is similar to that of the RK2 algorithm given in Algorithm 1, but the RK2 scheme is replaced by the ETD2 scheme in Equation (22).

2.4. Adaptive Exponential Time Differencing Method

The ETD2 method can indeed use a large time step to improve computational efficiency, but we find that it will introduce relatively large error in the trajectories of neurons' membrane potentials and even lead to the missing of action potentials (see below for numerical results). In addition, the number of the missed action potentials in the ETD2 method can grow with the increase of time steps compared with the RK2 method using a small time step. Thus, it is important to design an efficient but also reliable ETD method to solve this issue.

As shown in **Figure 1A**, the slope of voltage has a very large value when the neuron generates an action potential (spike period) and quickly reduces to a value around zero in the non-spike period until the next spike time. Therefore, the stiffness of HH equations is quite different between spike and non-spike periods. In the non-spike period, the slope of voltage is almost zero, while the linear and nonlinear parts in Equation (14) have a much larger absolute value and nearly cancel each other out as shown in **Figure 1B**. Therefore, the decomposition in Equation (14) may not be appropriate in the non-spike period since both the linear and nonlinear parts become stiff while the summation of them is indeed non-stiff. Based on this, we propose a different decomposition in the non-stiff period from that in the stiff period: taking $c_z = 0$ and the whole right hand side of Equation (14) as the nonlinear part. For such a decomposition, the ETD2 scheme reduces to the RK2 scheme in the non-stiff period.

The stiff period of HH equations can be clearly identified as shown in **Figure 1A** and is defined as follows. For each spike event, the starting point of the stiff period is determined as the spike time when the voltage reaches the firing threshold $Vth = -50$ mV and the interval of stiff period is chosen as 3.5 ms which is sufficient long to cover the highly stiff region of

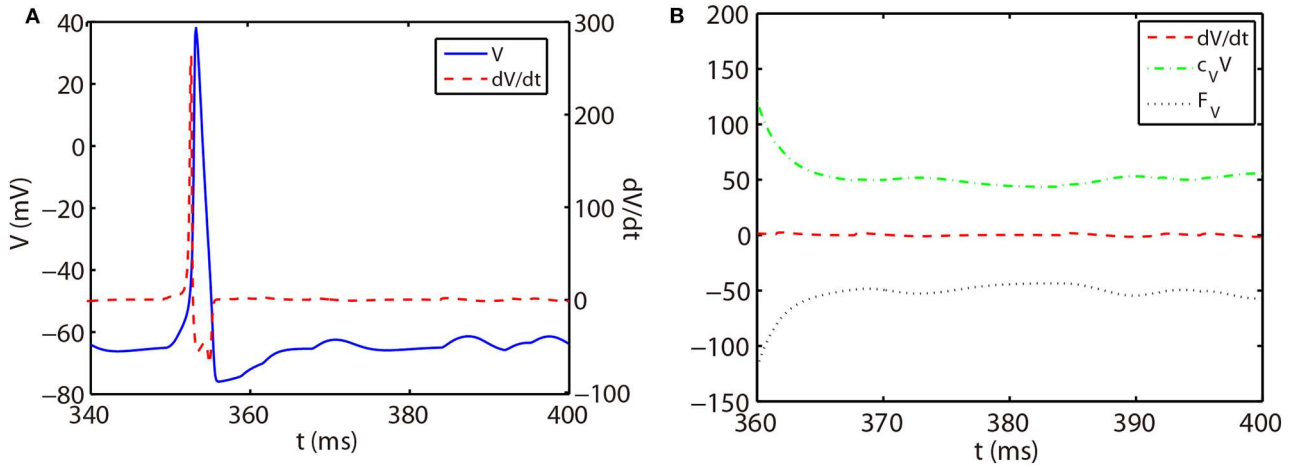

FIGURE 1 | (A) Trajectory of voltage (blue solid curve) and the slope of voltage (red dashed curve) for a single HH neuron. **(B)** Trajectory of the slope of voltage (red dashed curve), linear part $c_V V$ (green dash-dotted curve), and nonlinear part F_V (black dotted curve) in Equation (14) for the non-spike period. The time interval of **(B)** zooms into the later part of **(A)**.

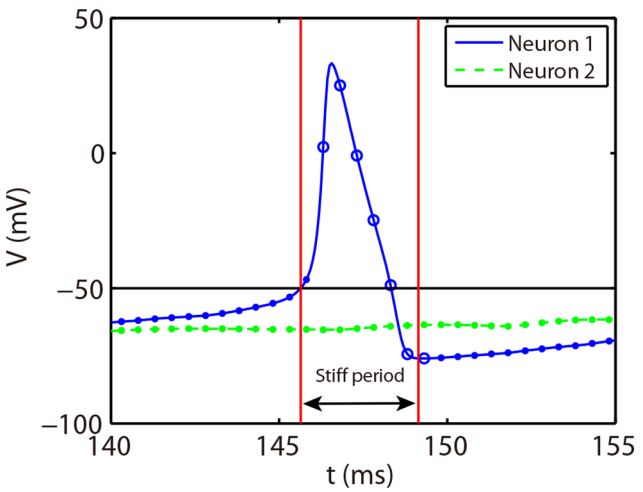

FIGURE 2 | Illustration of the AETD2 method. After neuron 1 fires a spike, we use the ETD2 scheme to evolve the HH equations for neuron 1 during the stiff period indicated by the red vertical lines, while the HH equations for neuron 2 is evolved using RK2 scheme since neuron 2 is in the non-stiff period. The starting point of the stiff period is determined as the spike time and it lasts for the following about 3.5 ms. The circles and dots indicate the time nodes where we use the ETD2 and RK2 schemes, respectively.

Algorithm 2: AETD2 algorithm

Input: an initial time t_k, feedforward input times $\{s_{il}\}$
Output: Solutions at time t_{k+1}
for $i = 1$ *to* N **do**
  Solve the HH equations for the ith neuron without considering spike input:
  if *The ith neuron is inside the stiff period* **then**
    use ETD2 scheme
  else
    use RK2 scheme
  end
  if $V_i(t_k) < V^{th}$ *and* $V_i(t_{k+1}) \geq V^{th}$ **then**
    The ith neuron spiked in $[t_k, t_{k+1}]$.
    Estimate the spike time τ_{il} by Equation (9).
  end
end
Recalibrate the conductance by Equations (12) and (13).

the spike. Based on the above observation, we give our AETD2 method for HH neural network as following: each neuron is evolved using ETD2 scheme if it is in the stiff period and use the reduced ETD2 scheme, the RK2 scheme, otherwise, as shown in **Figure 2**. Detailed AETD2 algorithm is given in Algorithm 2.

3. RESULTS

We consider an all-to-all connected network of 80 excitatory and 20 inhibitory neurons driven by Poisson feedforward input. For the ease of illustration, we choose the Poisson input strength $f = 0.06$ mS·cm^{-2} and input rate $\nu = 300$ Hz, and the coupling strength between neurons are chosen as $S = 0.2$ mS·cm^{-2} throughout this work, unless indicated otherwise. However, our algorithm can be applied to HH neural networks under a variety of dynamical regimes.

First, we verify the second-order numerical accuracy by performing convergence tests. A high precision solution is obtained by using RK2 method with a sufficiently small time step $\Delta t = 1 \times 10^{-6}$ ms and is denoted by a superscript "high." It is compared with the solutions computed by the RK2, ETD2, and AETD2 methods with various values of larger time steps $\Delta t = 2^{-4}, 2^{-5}, ..., 2^{-12}$ ms which is denoted by a superscript

"Δt." Errors of membrane potentials at final run time $T = 2,000$ ms and the last spike time of each neuron are computed:

$$\text{Error}_V = \sqrt{\sum_i (V_i^{(\Delta t)}(T) - V_i^{(\text{high})}(T))^2}, \tag{23}$$

$$\text{Error}_\tau = \sqrt{\sum_i (\tau_{il^*}^{(\Delta t)} - \tau_{il^*}^{(\text{high})})^2}, \tag{24}$$

where τ_{il^*} indicates the last spike time of the ith neuron during the run time interval. As shown in **Figure 3**, if one naively assigns the end of time step as the spike times in the RK2 method (naive RK2), the numerical accuracy of the membrane potentials and spike times can only be of the first-order. In contrast, if one determines the spike times by linear interpolation and recalibrate the conductances accordingly, all the RK2, ETD2, and AETD2 methods can achieve a second-order numerical accuracy. In addition, we find that the ETD2 method has much larger error compared with the RK2 and AETD2 methods using the same time step as shown in **Figure 3**. When using a time step larger than $\Delta t = 2^{-6} = 0.0156$ ms, the ETD2 method performs even worse than the naive RK2 method. The underlying reason is that the HH equations are almost non-stiff in the non-spike period, but the decomposition in Equation (14) induces a relatively large stiffness for the nonlinear term as discussed previously.

We next discuss the numerical performance of our AETD2 method and compare it with other different numerical methods. As shown in the top panel of **Figure 4**, the AETD2 method with large time steps (maximum time step $\Delta t = 0.277$ ms) can obtain the same high accuracy in membrane potentials as the RK2 method using a very small time step $\Delta t = 0.01$ ms. The bottom panel of **Figure 4** shows the raster plots (neuron index vs. its spike time) of the spike events in the network. It can be seen that the spike times are well-captured by the AETD2 method with large time steps. In contrast, as shown in **Figures 5A,B**, the ETD2 method is highly inaccurate in terms of voltage traces and raster plots when the time step $\Delta t = 0.277$ ms is used (the maximum time step in AETD2 method). **Figure 5C** shows the relative error in the mean firing rate (the average number of synaptic spikes per unit time) between the ETD2 and RK2 methods, and that between the AETD2 and RK2 methods over different values of coupling strength. It can be seen that the ETD2 method can achieve only one digit of numerical accuracy while the AETD2 method can robustly achieve more than two digits of numerical accuracy when the time step $\Delta t = 0.277$ ms is used in both methods. Therefore, the ETD2 method has worse numerical performance compared with the AETD2 method.

To demonstrate the efficiency of our AETD2 method, we compare the simulation time that RK2, ETD2, and AETD2 methods take for a common total run time. We simulate the all-to-all connected network by the RK2, ETD2, and AETD2 methods on a Windows platform using an Intel i7 2.6 GHz processor (the weblink of the source codes is given in section 4), and the simulation time and numerical accuracy of mean firing rate are given in **Table 1**. The AETD2 method can achieve over an order of magnitude of speedup compared with the RK2 method while achieving the same high accuracy in terms of the mean firing rate.

In addition, we define the efficiency ratio of the AETD2 method over the RK2 method as

$$E = \frac{T_{\text{RK2}}}{T_{\text{AETD2}}} \tag{25}$$

where T_{RK2} and T_{AETD2} indicate the simulation times of the RK2 and AETD2 methods, respectively, for the HH neural network to evolve the run time T. Note that the RK2 and ETD2 methods take almost the same simulation time when using the same small time step as shown in **Table 1**. Thus, the above efficiency ratio can be approximated by the ratio of the total number of time steps each method requires as

$$E \approx \frac{T/\Delta t_{\text{RK2}}}{T/\Delta t_{\text{AETD2}}} = \frac{\Delta t_{\text{AETD2}}}{\Delta t_{\text{RK2}}}, \tag{26}$$

where Δt_{RK2} and Δt_{AETD2} indicate the time steps used in the RK2 and AETD2 methods, respectively. To demonstrate that the above efficiency ratio is independent of the network connectivity, size, and dynamical regimes, we evolve the all-to-all connected network of 80 excitatory and 20 inhibitory neurons and a randomly connected network of 800 excitatory and 200 inhibitory neurons with a variety choice of coupling strength. Not surprisingly, the efficiency ratio approximated by Equation (26) agrees well with the one measured by the ratio of simulation time between the RK2 and AETD2 methods in both two networks as shown in **Figure 6**. Hence, the efficiency ratio of the AETD2 method relies on only the size of evolved time steps.

4. DISCUSSION

We have presented an adaptive second-order ETD method to evolve the pulse-coupled HH neural network. Our AETD2 method can solve the stiff problem in the HH equations when an HH neuron generates an action potential (spike period). It can use a large time step to raise computational efficiency while accurately capturing dynamical properties of HH neurons such as the trace of membrane potentials, spike times of each neuron, and the mean firing rate. We point out that our AETD2 method can robustly enlarge time steps and raise computational efficiency over one order of magnitude compared with the RK2 method. This high efficiency seems to be independent of parameter choice of connectivity structure, dynamical regimes, or network size.

Our adaptive ideas of ETD methods can be applied to dynamical systems with stiff and non-stiff periods. In addition, we point out that the ETD scheme in our AETD2 algorithm can be chosen in a variety of forms according to the properties of dynamical systems. Here, we use the ETD2 scheme derived by approximating the integration in Equation (19) with RK time stepping. Other forms of numerical schemes can also be used to approximate the integration. For example, one can use a liner interpolation to approximate the nonlinear part in Equation (14) to obtain another form of ETD2 scheme. Besides, one can derive an ETD scheme following the idea proposed in Börgers and

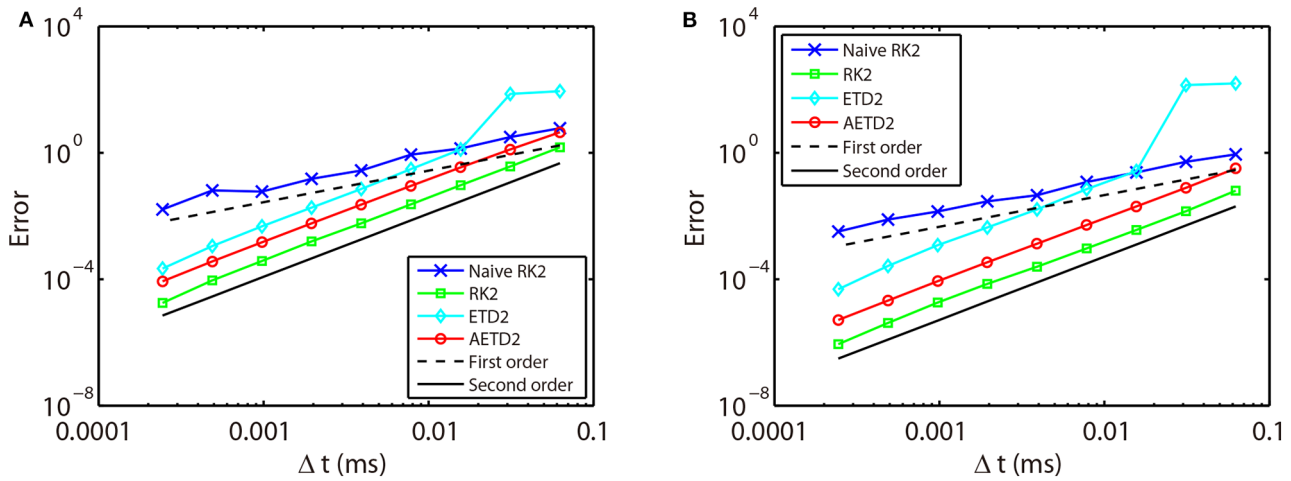

FIGURE 3 | Errors of membrane potentials **(A)** and the last spike time of each neuron **(B)** in the all-to-all connected network when it is evolved using various time steps. Blue crosses are naive RK2 method without performing the linear interpolation for the estimate of the spike times. Green squares are RK2 method, cyan diamonds are ETD2 method, and red circles are AETD2 method. The last three methods all perform the linear interpolation to estimate the spike times. The dashed line and the solid line indicate the numerical convergence of the first-order and the second-order, respectively. The total run time $T = 2,000$ ms.

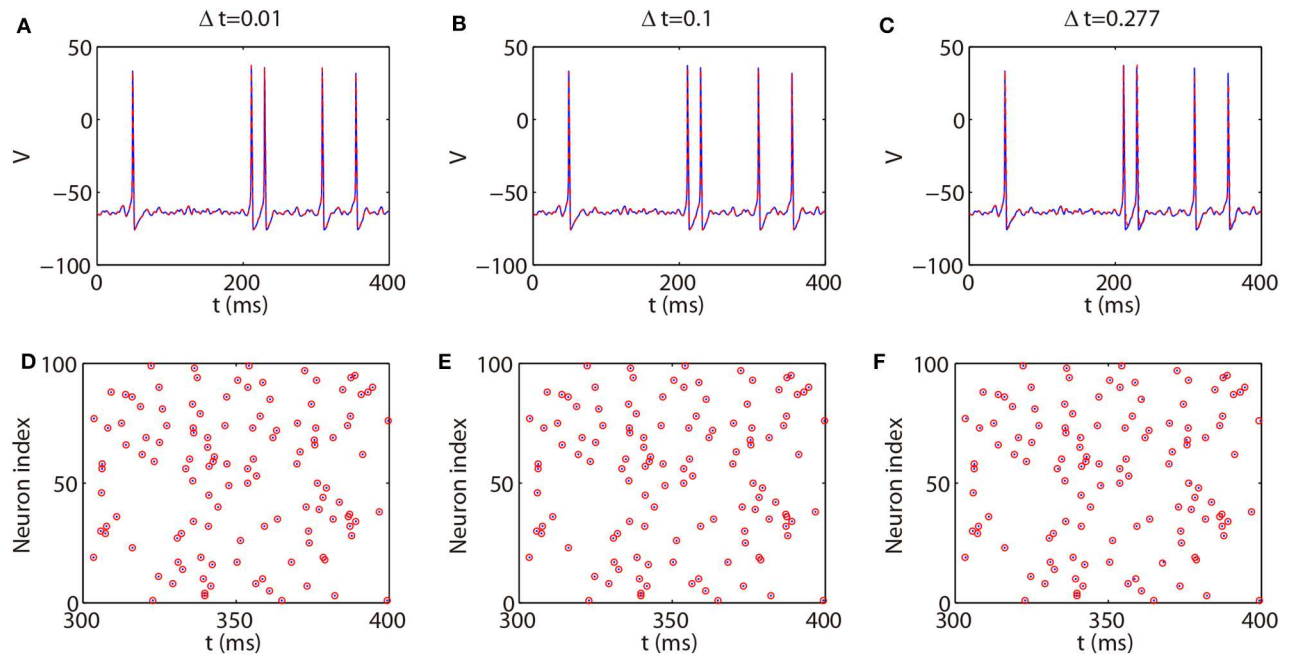

FIGURE 4 | Comparing the AETD2 method with the RK2 method. (Top) Traces of membrane potential of an HH neuron in the all-to-all connected network. (Bottom): Raster plots of the network spikes. The blue solid curves and dots indicate the results by the RK2 method with time step $\Delta t = 0.01$ ms, while the red dashed curves and circles indicate the results by the AETD2 method. The time steps for the AETD2 method are $\Delta t = 0.01, 0.1, 0.277$ ms for **(A,D)**, **(B,E)**, and **(C,F)**, respectively.

Nectow (2013) by linearly approximating the HH equations. The derived ETD scheme is proven to be unconditionally stable for HH system in Börgers and Nectow (2013). All these different ETD schemes can be easily embedded into our AETD2 method in the same way as given in Algorithm 2. For example, we can embed the ETD formula proposed in Börgers and Nectow (2013) into the AETD2 method to evolve the reduced Traub Miles (RTM) neural networks (Ermentrout and Kopell, 1998; Olufsen et al., 2003; Börgers and Nectow, 2013). The dynamical equations for an RTM neuron is almost the same as that for an HH neuron except that the gating variable m is described by $m_i = \alpha_m(V_i)/(\alpha_m(V_i) + \beta_m(V_i))$. The forms of α and β for the RTM neurons are set as: $\alpha_m(V_i) = 0.32(V_i+54)/(1-\exp(-(V_i+54)/4))$, $\beta_m(V_i) = 0.28(V_i+27)/(\exp((V_i+27)/5)-1)$, $\alpha_h(V_i) =$

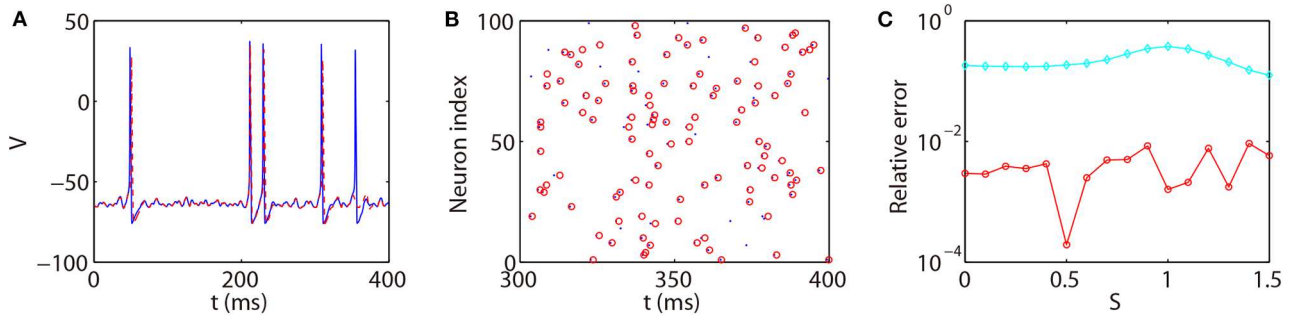

FIGURE 5 | Comparing the ETD2 and AETD2 methods with the RK2 method. **(A)** Voltage trace of the same HH neuron used in **Figure 4**. **(B)** Raster plot of the network spikes. The blue solid curve and dots indicate the results by the RK2 method with time step $\Delta t = 0.01$ ms while the red dashed curve and circles indicate the results by the ETD2 method with time step $\Delta t = 0.277$ ms. The coupling strength is $S = 0.2$ mS·cm^{-2}. **(C)** Relative error in the mean firing rates between the ETD2 and RK2 methods (cyan diamonds), and that between the AETD2 and RK2 methods (red circles) for different choice of the coupling strength. Both the ETD2 and AETD2 methods use time step $\Delta t = 0.277$ ms. The benchmark mean firing rate is computed by the RK2 method with a very small time step $\Delta t = 1 \times 10^{-6}$ ms.

TABLE 1 | Simulation of the all-to-all connected network with a total run time $T = 10$ s.

	RK2			ETD2			AETD2		
Δt (ms)	CPU	Relative error		CPU	Relative error		CPU	Relative error	
0.005	60.56 s	0	(13.61 Hz)	60.07 s	0	(13.61 Hz)	60.82 s	0	(13.61 Hz)
0.01	30.22 s	0	(13.61 Hz)	30.05 s	0	(13.61 Hz)	30.55 s	0	(13.61 Hz)
0.02	14.99 s	0	(13.61 Hz)	15.03 s	0.074%	(13.60 Hz)	15.30 s	0	(13.61 Hz)
0.05	***	***	***	5.95 s	0.66%	(13.52 Hz)	6.16 s	0.074 %	(13.62 Hz)
0.1	***	***	***	2.94 s	2.65%	(13.25 Hz)	3.11 s	0.074 %	(13.62 Hz)
0.2	***	***	***	1.48 s	9.99%	(12.25 Hz)	1.57 s	0.15 %	(13.63 Hz)
0.277	***	***	***	1.09 s	12.56%	(11.29 Hz)	1.15 s	0.59 %	(13.69 Hz)
0.5	***	***	***	0.57 s	41.59%	(7.95 Hz)	***	***	***
1	***	***	***	0.29 s	87.07%	(1.76 Hz)	***	***	***

The simulation time is measured in seconds. The relative error in the mean firing rate between each method using different time steps and the RK2 method using a very small time step $\Delta t = 1 \times 10^{-6}$ ms is measured in percentage and the mean firing rate is measured in Hz given inside the parentheses. Asterisks indicate overflow errors.

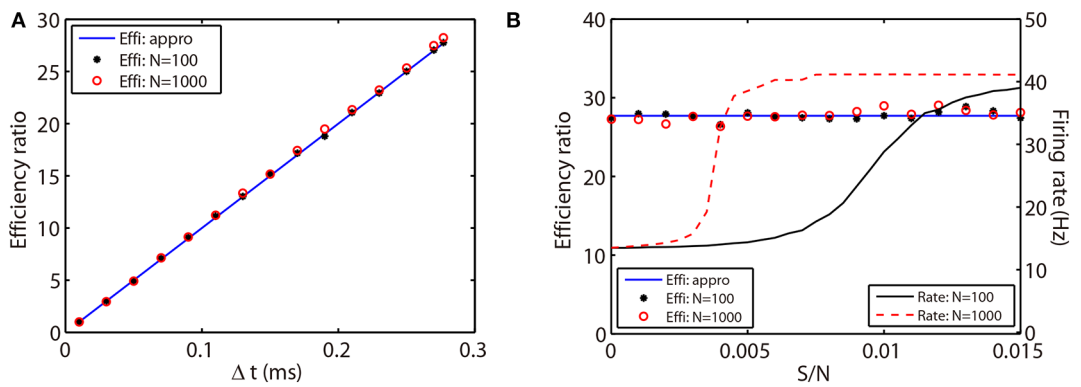

FIGURE 6 | Efficiency ratio of the AETD2 method when evolving the HH neural network using various time steps **(A)** and coupling strength **(B)**. In **(A,B)**, the blue lines are efficiency ratio measured by the approximation in Equation (26), while the black stars and red circles are the efficiency ratio measured by the ratio of simulation times between the RK2 and AETD2 methods. The black stars represent the results for the all-to-all connected HH neural network of 80 excitatory and 20 inhibitory neurons, while the red circles represent the results for an HH neural network of 800 excitatory and 200 inhibitory neurons randomly connected with probability 25%. The black solid and red dashed curves in **(B)** are the mean firing rates in the smaller network of 100 neurons and larger network of 1,000 neurons, respectively. The coupling strength in **(A)** is $S/N = 0.002$ mS·cm^{-2} and the time step for AETD2 method in **(B)** is $\Delta t = 0.277$ ms. The time step for RK2 method is 0.01 ms and total run time is $T = 50$ seconds in both **(A,B)**.

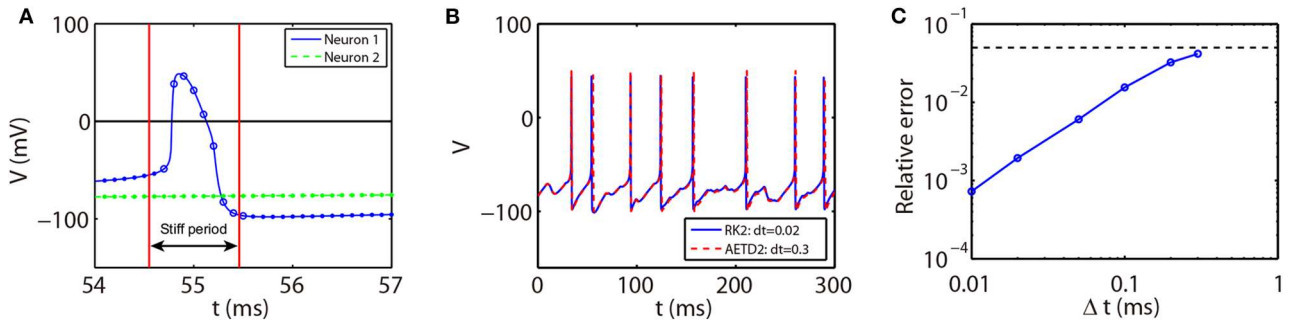

FIGURE 7 | The AETD2 method for the RTM neural network. **(A)** Illustration of the AETD2 method. We take a relatively high firing threshold $V^{th} = 0$ mV for the RTM neurons indicated by the solid horizontal line. The stiff period is defined as the region where $|dV/dt| \geq 20$. The circles and dots indicate the time nodes where we use the ETD2 and RK2 schemes, respectively. **(B)** Traces of membrane potential of an RTM neuron. The blue solid and red dashed curves indicate the results by the RK2 method with time step $\Delta t = 0.02$ ms and the AETD2 method with time step $\Delta t = 0.3$ ms, respectively. **(C)** Relative error in the mean firing rates between the AETD2 and RK2 methods for different choice of time steps used in the AETD2 method for an all-to-all connected RTM network of 80 excitatory and 20 inhibitory neurons with Poisson input. The dashed horizontal line indicates 5% error, relative error = 0.05. The benchmark mean firing rate is computed by the RK2 method with a very small time step $\Delta t = 1 \times 10^{-6}$. The parameters for RTM model is $C = 1\mu$F·cm^{-2}, $V_{Na} = 50$ mV, $V_K = -100$ mV, $V_L = -67$ mV, $G_{Na} = 100$ mS·cm^{-2}, $G_K = 80$ mS·cm^{-2}, and $G_L = 0.1$ mS·cm^{-2}.

$0.128\exp(-(V_i+50)/18)$, $\beta_h(V_i) = 4/(1+\exp(-(V_i+27)/5))$, $\alpha_n(V_i) = 0.032(V_i+52)/(1-\exp(-(V_i+52)/5))$, and $\beta_n(V_i) = 0.5\exp(-(V_i+57)/40)$.

Note that the rising phase of action potentials for the RTM neurons is extremely short, around 0.03 ms as shown in **Figure 7A**. In such a situation, it may not be appropriate to choose the spike time (when the voltage reaches the firing threshold) as the starting point of the stiff period in the AETD2 method as shown in **Figure 7A** since large numerical error will be introduced, especially when a large time step is used, e.g., time step $\Delta t = 0.3$ ms. This is because the system is evolved by the RK2 scheme during the time step that contains the rapid rising region of the neuron's action potential. Therefore, to achieve high numerical accuracy, the interval of the stiff period should cover the rapid rising region. To achieve this, we then define the stiff period of the RTM neurons as the region where the magnitude of the slope of voltage is over a proper threshold as shown in **Figure 7A**. We point out that our AETD2 method using a large time step can still accurately capture the membrane potential traces and the mean firing rates compared with the RK2 method using a small time step as shown in **Figures 7B,C**. Therefore, the definition of the stiff period in our AETD2 method can be flexibly determined based on dynamical properties of studied systems.

In this work, the numerical accuracy of our AETD2 method is second-order. In some situations, high accurate traces of membrane potentials may be required, especially the accurate shapes of action potentials (Traub et al., 2001; Kopell and Ermentrout, 2004). Therefore, one future work may be the design of the fourth-order ETD method. As illustrated above, due to the discontinuity arising from the pulse-coupled dynamics, an even more careful recalibration needs to be designed to achieve fourth-order numerical accuracy.

Finally, we point out that our AETD2 method can be easily extended to networks of other HH type neurons (Pospischil et al., 2008). And our AETD2 method can also robustly achieve high numerical accuracy and efficiency. In addition, our method is naturally a parallel algorithm which can be applied to simulations of large-scale neural network dynamics. For reproducibility of our results by other researchers, all the source codes written in C++ can be accessed at http://github.com/KyleZhongqi/ETD2_HH.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

ZT and DZ designed the research, wrote and code, ran the simulations, analyzed the results, and wrote the paper.

FUNDING

This work was supported by National Science Foundation of China with Grant Nos. 11671259, 11722107, and SJTU-UM Collaborative Research Program, and by the Student Innovation Center at Shanghai Jiao Tong University (DZ).

ACKNOWLEDGMENTS

This manuscript has been released as a Pre-Print at arXiv repository (Tian and Zhou, 2019). We dedicate this paper to our late professor David Cai.

REFERENCES

Aihara, K. (1986). Chaotic oscillations and bifurcations in squid giant axons. *Chaos* 257–269. doi: 10.1515/9781400858156.257

Börgers, C., Epstein, S., and Kopell, N. J. (2005). Background gamma rhythmicity and attention in cortical local circuits: a computational study. *Proc. Natl. Acad. Sci. U.S.A.* 102, 7002–7007. doi: 10.1073/pnas.0502366102

Börgers, C., and Nectow, A. R. (2013). Exponential time differencing for Hodgkin-Huxley-like odes. *SIAM J. Sci. Comput.* 35, B623–B643. doi: 10.1137/120883657

Brette, R., Rudolph, M., Carnevale, T., Hines, M., Beeman, D., Bower, J. M., et al. (2007). Simulation of networks of spiking neurons: a review of tools and strategies. *J. Comput. Neurosci.* 23, 349–398. doi: 10.1007/s10827-007-0038-6

Cox, S. M., and Matthews, P. C. (2002). Exponential time differencing for stiff systems. *J. Comput. Phys.* 176, 430–455. doi: 10.1006/jcph.2002.6995

Dayan, P., and Abbott, L. (2003). Theoretical neuroscience: computational and mathematical modeling of neural systems. *J. Cogn. Neurosci.* 15, 154–155. doi: 10.1162/089892903321107891

Dayan, P., and Abbott, L. F. (2001). *Theoretical Neuroscience, Vol. 806*. Cambridge, MA: MIT Press.

de la Hoz, F., and Vadillo, F. (2008). An exponential time differencing method for the nonlinear schrödinger equation. *Comput. Phys. Commun.* 179, 449–456. doi: 10.1016/j.cpc.2008.04.013

Ermentrout, G. B., and Kopell, N. (1998). Fine structure of neural spiking and synchronization in the presence of conduction delays. *Proc. Natl. Acad. Sci. U.S.A.* 95, 1259–1264. doi: 10.1073/pnas.95.3.1259

Guckenheimer, J., and Oliva, R. A. (2002). Chaos in the Hodgkin-Huxley model. *SIAM J. Appl. Dyn. Syst.* 1, 105–114. doi: 10.1137/S1111111101394040

Hansel, D., Mato, G., Meunier, C., and Neltner, L. (1998). On numerical simulations of integrate-and-fire neural networks. *Neural Comput.* 10, 467–483. doi: 10.1162/089976698300017845

Hansel, D., and Sompolinsky, H. (1996). Chaos and synchrony in a model of a hypercolumn in visual cortex. *J. Comput. Neurosci.* 3, 7–34. doi: 10.1007/BF00158335

Hassard, B. (1978). Bifurcation of periodic solutions of the Hodgkin-Huxley model for the squid giant axon. *J. Theor. Biol.* 71, 401–420. doi: 10.1016/0022-5193(78)90168-6

Hochbruck, M., Lubich, C., and Selhofer, H. (1998). Exponential integrators for large systems of differential equations. *SIAM J. Sci. Comput.* 19, 1552–1574. doi: 10.1137/S1064827595295337

Hochbruck, M., and Ostermann, A. (2010). Exponential integrators. *Acta Numer.* 19, 209–286. doi: 10.1017/S0962492910000048

Hodgkin, A. L., and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol.* 117:500. doi: 10.1113/jphysiol.1952.sp004764

Kassam, A.-K., and Trefethen, L. N. (2005). Fourth-order time-stepping for stiff pdes. *SIAM J. Sci. Comput.* 26, 1214–1233. doi: 10.1137/S1064827502410633

Kopell, N., and Ermentrout, B. (2004). Chemical and electrical synapses perform complementary roles in the synchronization of interneuronal networks. *Proc. Natl. Acad. Sci. U.S.A.* 101, 15482–15487. doi: 10.1073/pnas.0406343101

Lin, K. K. (2006). Entrainment and chaos in a pulse-driven Hodgkin-Huxley oscillator. *SIAM J. Appl. Dyn. Syst.* 5, 179–204. doi: 10.1137/050631677

Nie, Q., Wan, F. Y., Zhang, Y.-T., and Liu, X.-F. (2008). Compact integration factor methods in high spatial dimensions. *J. Comput. Phys.* 227, 5238–5255. doi: 10.1016/j.jcp.2008.01.050

Olufsen, M. S., Whittington, M. A., Camperi, M., and Kopell, N. (2003). New roles for the gamma rhythm: population tuning and preprocessing for the beta rhythm. *J. Comput. Neurosci.* 14, 33–54. doi: 10.1023/A:1021124317706

Pospischil, M., Toledo-Rodriguez, M., Monier, C., Piwkowska, Z., Bal, T., Frégnac, Y., et al. (2008). Minimal Hodgkin-Huxley type models for different classes of cortical and thalamic neurons. *Biol. Cybern.* 99, 427–441. doi: 10.1007/s00422-008-0263-8

Shelley, M. J., and Tao, L. (2001). Efficient and accurate time-stepping schemes for integrate-and-fire neuronal networks. *J. Comput. Neurosci.* 11, 111–119. doi: 10.1023/A:1012885314187

Somers, D. C., Nelson, S. B., and Sur, M. (1995). An emergent model of orientation selectivity in cat visual cortical simple cells. *J. Neurosci.* 15, 5448–5465. doi: 10.1523/JNEUROSCI.15-08-05448.1995

Sun, Y., Zhou, D., Rangan, A. V., and Cai, D. (2009). Library-based numerical reduction of the Hodgkin-Huxley neuron for network simulation. *J. Comput. Neurosci.* 27, 369–390. doi: 10.1007/s10827-009-0151-9

Tian, Z.-Q. K., and Zhou, D. (2019). Exponential time differencing algorithm for pulse-coupled Hodgkin-Huxley neuronal networks. *arXiv [preprint] arXiv:1910.08724*.

Traub, R. D., Kopell, N., Bibbig, A., Buhl, E. H., LeBeau, F. E., and Whittington, M. A. (2001). Gap junctions between interneuron dendrites can enhance synchrony of gamma oscillations in distributed networks. *J. Neurosci.* 21, 9478–9486. doi: 10.1523/JNEUROSCI.21-23-09478.2001

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.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 12 June 2020
doi: 10.3389/fncom.2020.00052

Spike-Timing Dependent Plasticity Effect on the Temporal Patterning of Neural Synchronization

Joel Zirkle[1] *and Leonid L. Rubchinsky*[1,2]*

[1] Department of Mathematical Sciences, Indiana University Purdue University Indianapolis, Indianapolis, IN, United States,
[2] Stark Neurosciences Research Institute, Indiana University School of Medicine, Indianapolis, IN, United States

Neural synchrony in the brain at rest is usually variable and intermittent, thus intervals of predominantly synchronized activity are interrupted by intervals of desynchronized activity. Prior studies suggested that this temporal structure of the weakly synchronous activity might be functionally significant: many short desynchronizations may be functionally different from few long desynchronizations even if the average synchrony level is the same. In this study, we used computational neuroscience methods to investigate the effects of spike-timing dependent plasticity (STDP) on the temporal patterns of synchronization in a simple model. We employed a small network of conductance-based model neurons that were connected via excitatory plastic synapses. The dynamics of this network was subjected to the time-series analysis methods used in prior experimental studies. We found that STDP could alter the synchronized dynamics in the network in several ways, depending on the time scale that plasticity acts on. However, in general, the action of STDP in the simple network considered here is to promote dynamics with short desynchronizations (i.e., dynamics reminiscent of that observed in experimental studies). Complex interplay of the cellular and synaptic dynamics may lead to the activity-dependent adjustment of synaptic strength in such a way as to facilitate experimentally observed short desynchronizations in the intermittently synchronized neural activity.

Keywords: STDP, synaptic plasticity, intermittency, synchronization, phase-locking, neural oscillations

OPEN ACCESS

Edited by:
Julie S. Haas,
Lehigh University, United States

Reviewed by:
Wilten Nicola,
University of Waterloo, Canada
Xiaojuan Sun,
Beijing University of Posts and Telecommunications (BUPT), China

***Correspondence:**
Leonid L. Rubchinsky
lrubchin@iupui.edu

Received: *26 November 2019*
Accepted: *12 May 2020*
Published: *12 June 2020*

Citation:
Zirkle J and Rubchinsky LL (2020) Spike-Timing Dependent Plasticity Effect on the Temporal Patterning of Neural Synchronization. Front. Comput. Neurosci. 14:52. doi: 10.3389/fncom.2020.00052

INTRODUCTION

Synchronization of neural activity in the brain is involved in multiple neural functions (e.g., Buzsáki and Draguhn, 2004; Fell and Axmacher, 2011; Fries, 2015; Harris and Gordon, 2015). Neural synchronization that is either too strong or too weak may be one of the neurophysiological factors behind symptoms of several disorders such as Parkinson's disease and schizophrenia (Schnitzler and Gross, 2005; Uhlhaas and Singer, 2006; Oswal et al., 2013; Pittman-Polletta et al., 2015). Thus, the synchronization of neural activity is a ubiquitous phenomenon. In the rest state, the strength of this synchronization is usually moderate. This means that the intervals of stronger synchrony are interspersed with desynchronized intervals. This is probably not surprising given the plausibility of the very general nature of the transient character of neural activity (Rabinovich et al., 2008).

Recent developments in time-series analysis allowed for the exploration of the temporal patterning of synchronized activity in brain dynamics on very short time-scales. Studies of different brain signals in different conditions and species suggest an apparently universal feature:

synchronous activity is interrupted by very short (although potentially numerous) intervals of desynchronized dynamics (as opposed to few longer desynchronized episodes). This phenomenon was observed in the synchrony between local field potentials (LFPs) and spikes in different parts of the basal ganglia and EEG in Parkinson's disease (Park et al., 2010; Ratnadurai-Giridharan et al., 2016; Ahn et al., 2018), in synchronization between LFPs recorded in the prefrontal cortex and hippocampus of normal and amphetamine-sensitized mice (Ahn et al., 2014), in EEG of healthy human subjects (Ahn and Rubchinsky, 2013), and in EEG in autism spectrum disorders (Malaia et al., 2020). The differences in the temporal patterning are correlated with certain behavioral features but the prevalence of short desynchronizations persisted nevertheless (Ahn et al., 2014, 2018; Malaia et al., 2020). Therefore, short desynchronizations may be functionally important and the properties and mechanisms of desynchronization durations merit exploration.

These observations of the persistence of short desynchronizations naturally suggest the question about the biological mechanisms behind this phenomenon. The modeling study (Ahn and Rubchinsky, 2017) suggested one possible mechanism: the short desynchronization dynamics was promoted by the substantial difference in the timescales of spike-producing sodium and potassium currents. The relative slowness of the potassium delayed-rectifier current may be one of the reasons for why short desynchronizations are observed in different neural systems. However, there may also be other mechanisms. This paper is aimed at the exploration of one potential mechanism related to synaptic plasticity. We use computational modeling to explore how spike-timing dependent plasticity (STDP) can affect the temporal patterning of neural synchrony on short timescales.

STDP is a very common neural phenomenon with potentially multiple effects on neural synchronization. In particular, a synapse whose conductance is modulated by STDP can enhance neural synchrony (Nowotny et al., 2003; Cassenaer and Laurent, 2007; Knoblauch et al., 2012; Ratnadurai-Giridharan et al., 2015). We use a simple neural network of two conductance-based model neurons coupled via excitatory synapses with STDP and apply the same time-series analysis techniques as were used in the prior experimental studies. While this model network can hardly adequately model field potentials recorded in some of the experimental studies mentioned above, it serves as a simple model system exhibiting rich synchronization dynamics, which is substantially modulated by synaptic plasticity. Numerical analysis of this model shows that STDP may affect not only the strength of synchronization, but also the temporal patterning of synchronization, with an ability to facilitate the short desynchronizations dynamics observed in experiments.

METHODS

Neuronal and Synaptic Modeling

We utilize the network model from Ahn and Rubchinsky (2017) except that the synapses are plastic in this study. The model is described below.

The neurons are modeled using a two-dimensional conductance-based model of a Hodgkin-Huxley type that is mathematically equivalent to the Morris-Lecar model (Izhikevich, 2007; Ermentrout and Terman, 2010). The sodium conductance is assumed to activate instantaneously and to have no inactivation, while the potassium conductance is controlled by its gating variable and so varies dynamically.

$$\frac{dv}{dt} = -I_{Na} - I_K - I_L - I_{syn} + I_{app}$$

$$\frac{dw}{dt} = \frac{w_\infty(v) - w}{\tau(v)}$$

Here v is the neuron's transmembrane potential and w is the gating variable for the potassium current. The synaptic current between neurons, I_{syn}, is given below and I_{app} is a constant input current to each neuron to control the frequency of spiking. The sodium, potassium, and leak currents are:

$$I_{Na} = g_{Na} m_\infty(v)(v - v_{Na})$$

$$I_K = g_K w(v - v_K)$$

$$I_L = g_L(v - v_L)$$

g_{Na}, g_K, and g_L are the maximal conductances for the sodium, potassium, and leak currents, respectively. The steady-state values for the gating variables of the sodium and potassium currents are:

$$m_\infty(v) = \frac{1}{1 + exp\left(-2\frac{v - v_{m1}}{v_{m2}}\right)}$$

$$w_\infty(v) = \frac{1}{1 + exp\left(-2\frac{v - v_{w1}}{\beta}\right)}$$

The voltage-dependent activation time constant of the potassium current is:

$$\tau(v) = \frac{1}{\epsilon} * \frac{2}{exp\left(\frac{v - v_{w1}}{2\beta}\right) + exp\left(\frac{v_{w1} - v}{2\beta}\right)}$$

All synapses are excitatory, and the synaptic current to neuron i is given by:

$$I_{syn,i} = g_{syn}(v_i - v_{syn}) \sum_{j \neq i} s_j$$

Where g_{syn} is the maximal conductance of the synapse (i.e., the synaptic strength), and s_j is the synaptic variable for neuron j and the summation is taken over all neurons that are connected to the i-th neuron. The synaptic variable s is governed by:

$$\frac{ds}{dt} = \alpha_s (1 - s) H_\infty (v - \theta_v) - \beta_s s$$

H_∞ is a sigmoidal function whose input is the presynaptic neuronal voltage:

$$H_\infty (v) = \frac{1}{1 + exp\left(-\frac{v}{\sigma_s}\right)}$$

The values of cellular and synaptic parameters are the same as used in Ahn and Rubchinsky (2017): $g_{Na} = 1, g_K = 3.1, g_L = 0.5$, $v_{Na} = 1$, $v_K = -0.7$, $v_L = -0.4$, $v_{m1} = -0.01$, $v_{m2} = 0.15$, $v_{w1} = 0.08$, $\beta = 0.145$, $I_{app} = 0.045$, $\varepsilon_1 = 0.02$, $\varepsilon_2 = 1.2\varepsilon_1$, $v_{syn} = 0.5$, $\propto_s = 5$, $\beta_s = 0.2$, $\theta_v = 0.0$, $\sigma_s = 0.2$.

STDP modeling follows (Zhigulin et al., 2003). If neuron i spikes at time t_i and neuron j spikes at time t_j, then the strength of the synapse from neuron i to neuron j is additively updated by the amount

$$\Delta g_{syn} = sgn(\Delta t) A exp\left(-k|\Delta t|\right)$$

where $\Delta t = t_j - t_i$. The synaptic conductance from neuron j to neuron i is simultaneously updated by an equal, but opposite, amount. While the additive update rule does not necessary need to be symmetric (as it is here), there is experimental evidence supporting the nature of the update, see for example (Zhang et al., 1998; Feldman, 2012). We varied the values of our plastic parameters, in particular $A \in [0.0001, 0.01]$, $k \in [0.01, 50]$. The synaptic conductance is bounded below by zero.

Numerical Implementation

The system of differential equations was solved numerically in Python using the built-in odeint function from the SciPy module (v.1.4.1). This function implements either the Adams method or a backward differentiation formula (BDF) method depending on the stiffness of the problem. The solution was reported at multiples of the time step $dt = 0.1$ (assuming the time units are milliseconds), however the function uses an adaptive step size and there was no lower bound on the length of the intermediate time steps that may be used (similarly, there was no upper bound restriction on the number of intermediate steps that were taken). The absolute and relative tolerances for the method were kept at the default value of 1.49×10^{-8}. While the solution depends on the initial conditions, its statistical properties (such as the firing rate, synchrony pattern characteristics etc.) do not. The system was solved on the time interval $[0, 25000]$, the first 20% of the time-series was removed from analysis. To implement plasticity, the integration was paused after each time step and, if necessary, the synaptic strength was updated. Specifically, the voltage threshold to define an action potential was set at 0.2.

Synchronization Analysis

The time-series analysis of synchronized dynamics in the network follows that of Ahn et al. (2011) and Ahn and Rubchinsky (2017) and is similar to the analysis of the temporal patterns of neural synchrony in the experimental studies mentioned in the Introduction. We will briefly describe this analysis here.

The phase, $\varphi(t)$, of a neuron is defined as

$$\varphi(t) = \tan^{-1}\left(\frac{v(t) - \hat{v}}{w(t) - \hat{w}}\right)$$

where $(\hat{w}, \hat{v})$ is a point selected inside the neuron's limit cycle in the (w, v) – plane. The synchronization strength is computed as

$$\gamma = \left| \frac{1}{N} \sum_{j=1}^{N} exp\left(i \Delta\varphi\left(t_j\right)\right) \right|^2$$

where $\Delta\varphi\left(t_j\right) = \varphi_1\left(t_j\right) - \varphi_2(t_j)$ is the difference of the phases of neurons 1 and 2 at time t_j. N is the number of data points. The value of γ ranges from 0 to 1, which represent a complete lack of synchrony and perfect phase synchrony, respectively.

If there is some degree of phase locking present, then there is a synchronized state, i.e., a preferred value of the phase difference $\Delta\varphi$. For each cycle of oscillation one can check if the actual phase difference is close to this preferred value or not. Note that the index γ only represents an average value of phase-locking over the interval $[t_1, t_N]$, however to describe the patterning of synchrony one needs to look at the transitions to and from a synchronized state on much shorter timescales. This is done as follows.

When φ_1 increases past zero, say at time $t_{j,i}$, then $\varphi_2\left(t_{j,i}\right)$ is recorded. This generates a sequence of numbers $\left\{\varphi_2\left(t_{j,i}\right)\right\}_{i=1}^{M}$. Due to the presence of some synchrony, there is a clustering about some phase value, say φ_0. This is taken as the preferred phase value, and if $\varphi_2\left(t_{j,i}\right) = \varphi_i$, for $1 \leq i \leq M$, differs from it by more than $\frac{\pi}{2}$ then the neurons are desynchronized, otherwise they are synchronized. The choice of $\frac{\pi}{2}$ is not only convenient (it partitions the $(\varphi_i, \varphi_{i+1})$ space into quadrants) but was also used in the experimental studies described in the section Introduction.

The length of a desynchronization event is defined as the number of consecutive times the system spends in the desynchronized states. In other words, the length of desynchronization is the length of the time interval the system is away from the synchronized state (as defined above); this length is measured not in the absolute time units, but in the number of cycles of oscillations (in line with the experimental studies mentioned in the Introduction). The lengths of all desynchronization events are recorded and the distribution of durations is reconstructed. The mode of this distribution is used as a characteristic of the temporal patterning of synchronized dynamics. For later reference, a "mode n" system means that the mode of all lengths of desynchronization events for that particular system is n. Thus, a mode 1 system ($n = 1$

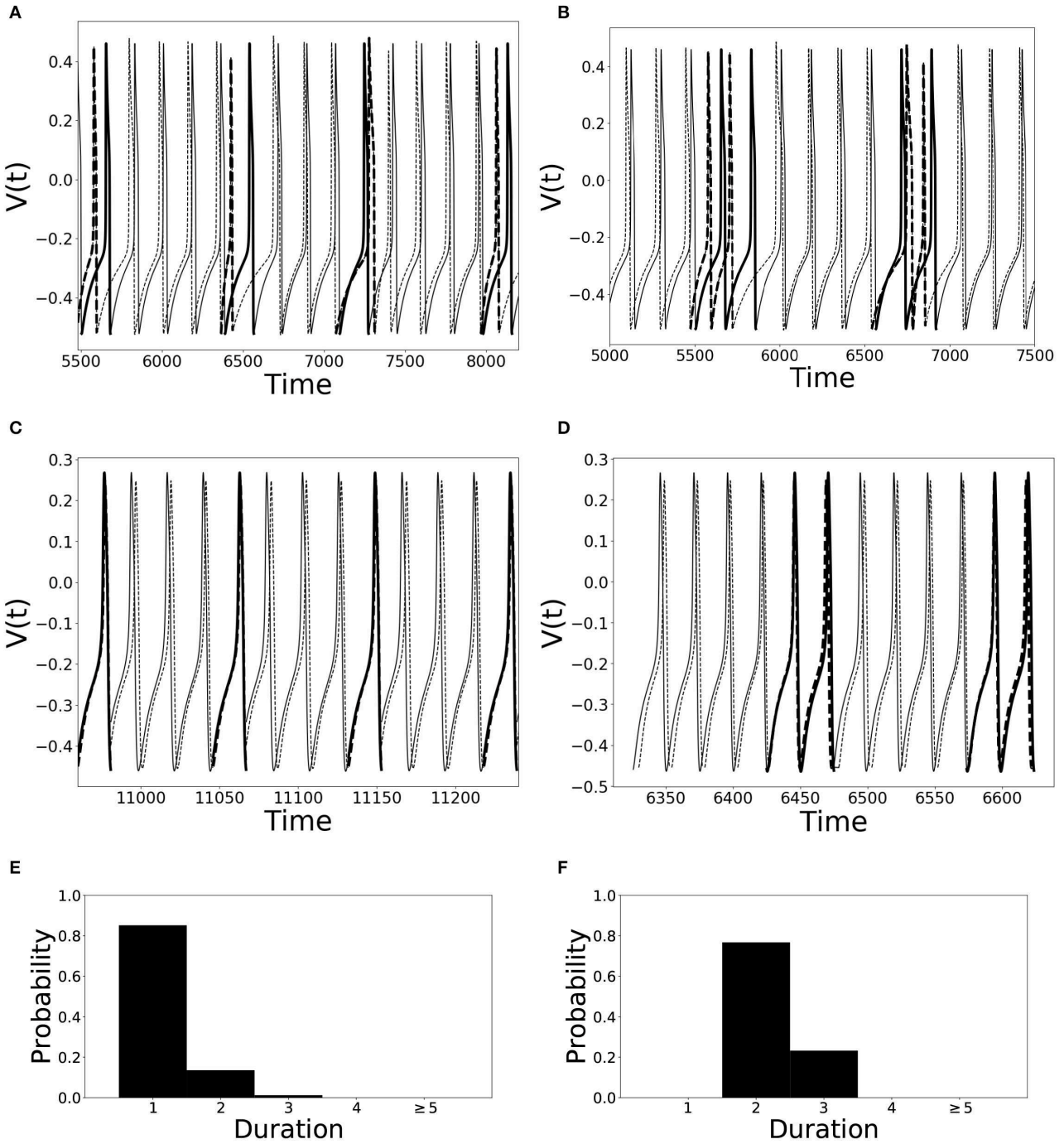

FIGURE 1 | Illustration of dynamics with different desynchronization durations (mode 1 and mode 2 dynamics). **(A–D)** depict voltage traces of two partially synchronized neurons (solid and dashed lines). When the neurons exhibit the preferred time difference the voltage traces are thin lines, indicating proximity to a synchronized state. However, when the phase difference is not close to the preferred one, the lines are thick to indicate the desynchronizations (as defined above). **(A,C)** illustrate short desynchronizations (lasting one cycle of oscillations), **(B,D)** show longer desynchronizations (lasting two cycles of oscillations). **(A,B)** are artificially generated examples, while **(C,D)** present examples generated by the network considered in the section below. In a longer time-series, the desynchronizations of different durations may coexist, however, usually one duration will prevail. The distributions showing relative frequency of different desynchronizations for the dynamics with predominantly short desynchronizations (like **A,C**) and with longer desynchronizations (like **B,D**) are presented in **(E,F)**, respectively. The mode of the distribution in **(E)** is 1, thus this is mode 1 dynamics; the mode of the distribution in **(F)** is 2, thus this is mode 2 dynamics.

case) is the system with synchronized dynamics interrupted by predominantly short desynchronization intervals. The larger n is, the more prominent the tendency for long desynchronizations is. This does not necessarily affect the overall synchrony strength, because it depends not only on the duration of desynchronizations, but also on their number. The mode is used to characterize the durations because experimental studies used the mode for this purpose.

An illustration of different desynchronization durations and dynamics with different modes of desynchronizations is provided in **Figure 1**. Voltages and distributions of desynchronization durations for mode 1 dynamics are in the left column, the ones for mode 2 dynamics are in the right column. The synchronization is not perfect and synchronized dynamics (phase difference is close to the preferred one) are interspersed with desynchronized intervals. Note that the preferred phase difference is not necessarily zero so that the zero lag state is not necessarily a synchronized state.

Finally, we would like to reiterate that in this approach the time is measured in terms of cycles of oscillations of the neural activity, not in absolute time units. This allows one to compare the properties of variability of synchrony of brain rhythms with different frequencies.

The phase-locking strength index γ was observed to be usually about 0.2–0.3 in this study (even after STDP adjustments). These are moderate values, comparable with experimental results (in particular with the results reported in the studies references in the section Introduction). With this moderate synchrony strength, synchronization effects are hard to see by the naked eye, however, the quantitative time-series analysis techniques are able to quantify the synchronized dynamics and its properties including the temporal patterning of weakly synchronous dynamics.

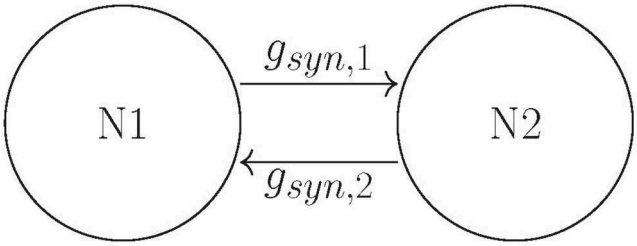

FIGURE 2 | The schematics of the network: two neurons coupled with mutually excitatory synapses.

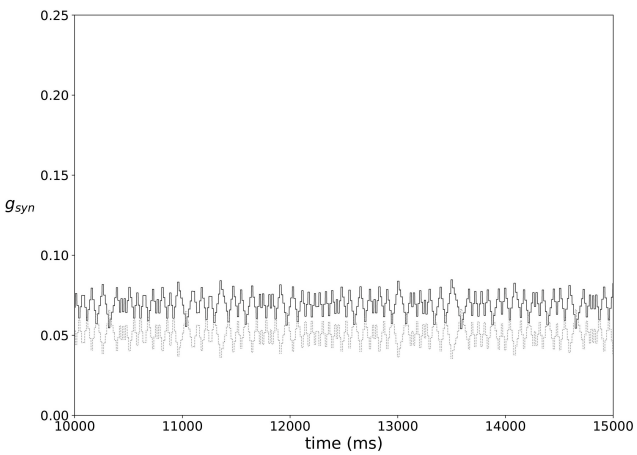

FIGURE 3 | An example of typical temporal evolution of synaptic weights in a network with plasticity ($\varepsilon = 0.15$, $A = 0.009$, $k = 0.3$).

RESULTS

Building on Ahn and Rubchinsky (2017), we used a simple network consisting of two neurons connected via excitatory synapses (see **Figure 2**); however the synapses are now plastic. The two neurons have a slightly different firing rate, i.e., their respective ε values differ slightly (see the list of parameter values in section Methods). The initial value of the maximal synaptic conductance is $g_{syn} = 0.005$, so that the coupling is weak. This heterogeneity and weak synaptic coupling ensure that the synchrony between the two neurons is relatively weak.

The dynamics of the non-plastic variant of this system was studied in Ahn and Rubchinsky (2017). Based on that study, we vary values of three parameters of the potassium current in such a way as to change the dynamics of the non-plastic network from exhibiting predominantly short desynchronizations (i.e., those observed in experiments) to one with a large mode of desynchronization durations. These parameters are ε [the reciprocal of the peak value of the activation time-constant $\tau(v)$], β [which characterizes the widths of the activation time-constant $\tau(v)$ and the steady-state function $w_\infty(v)$], and v_{w1} [a horizontal translation in $w_\infty(v)$ and $\tau(v)$ which changes their values over the specific voltage range]. Changes in all these parameters effectively change the activation time-constant $\tau(v)$ to either large or small, which delays or accelerates the activation of potassium current, respectively. Consequently, the lengths of the desynchronization events shift to predominantly short or long. Next, we explore how the introduction of plasticity affects the durations of desynchronization events. Hence our parameter space is two-dimensional for each case considered, and consists only of the plasticity parameters A and k.

In most of the simulations the synaptic weights do not reach a steady state, but rather exhibit fairly stationary variations, as illustrated in **Figure 3**.

Variation of ε

Let us mention here that $\varepsilon \propto \frac{1}{\tau}$ and the maximum value of $\tau(v)$ is $\frac{1}{\varepsilon}$. Hence as ε is increased, the value of $\tau(v)$ is decreased across its entire domain as it is a unimodal function. This in turn accelerates the activation of potassium current because $\frac{dw}{dt} \propto \frac{1}{\tau(v)}$. From Ahn and Rubchinsky (2017) we know that smaller values of ε promote shorter desynchronization events.

For $\varepsilon = 0.05$, the non-plastic system is mode 1. This means the synchronized dynamics has the following property. As the system is exhibiting partially synchronized dynamics, it will be either close in the synchronized state or away from

synchronized state, the latter is termed desynchronization. The desynchronized interval length (measured in the number of cycles of oscillations) varies in time. We obtain the distribution of the desynchronization durations from numerical simulation and find the mode of this distribution. If this mode equals one cycle of oscillation, then the system is mode 1 (see section Methods for a more detailed explanation). Mode 1 means the desynchronizations are predominantly short.

Now the non-plastic system is changed to include STDP. The changes in the temporal patterning of synchronization dynamics are illustrated in **Figure 4**. **Figure 4A** is a diagram of the mode of the desynchronization durations in the space of plasticity parameters, A and k. The plasticity effects are negligible across the top (very large k implies a quick decay of the change in synaptic strength), and especially in the upper left corner (large k and a small amplitude A). In these areas the values of the plasticity parameters are such that the magnitude of the update, Δg_{syn}, is negligible (the average update is usually in the interval $[0.0, 10^{-5}]$, on the larger end this corresponds to about 0.2% of the initial value of g_{syn}). Hence, the plastic system continues to be mode 1 in these areas.

The rest of the parameter space, in particular the central region, displays a high proportion of mode 1 dynamics as well. In these areas plasticity is not negligible, as the synaptic strength can vary to a substantial degree. However, even in the presence of STDP, mode 1 dynamics persist. For the diagram in **Figure 4A**, about 85% of the parameter space points correspond to mode 1 systems.

To illustrate the effect of plasticity on a distribution of desynchronization durations, refer to **Figures 4B–D**. Plasticity effects increase from left to right. The distribution of durations changes: at a weak level of plasticity the durations are exclusively length one, while at a stronger level of plasticity some longer durations are observed. Yet the preponderance of length one desynchronization durations is preserved.

Now let us look at the effect of plasticity on the dynamics in systems with a mode larger than one. We consider $\varepsilon = 0.15$. The non-plastic system is mode 2 (the synchronization index γ is virtually unchanged from that of $\varepsilon = 0.05$, although the frequency of oscillations increases by several times, Ahn and Rubchinsky, 2017). Mode 2 means the desynchronizations tend to be longer than those of the mode 1 case.

Figure 5 shows the effect of STDP on the system that is mode 2 in the non-plastic case. As explained earlier, the plasticity effects are negligible across the top of **Figure 5A**, and especially in the upper left corner. We note that this region of the parameter space exhibits mode 2 dynamics (as expected). However, throughout the entire parameter space it is seen that a majority of parameter values correspond to mode 1 systems (the large central region in **Figure 5A**). Overall, about 20% of the parameter space points stay mode 2, while over 65% exhibit mode 1 dynamics (and less than 15% correspond to larger than mode 2 systems).

To illustrate the effect of plasticity on a distribution of desynchronization durations, refer to **Figures 5B–D**. Plasticity effects increase from left to right. Here we see that the introduction of weak plasticity can be sufficient to shift the system from mode 2 to mode 1 (**Figure 5C**). This means desynchronizations tend to become shorter in the plastic case. At stronger levels of plasticity (**Figure 5D**), the distribution widens, however the vast majority of desynchronization events remain length one.

Overall, we have seen that mode 1 dynamics are generally preserved when STDP is introduced to a non-plastic mode 1 system. When STDP is introduced to a non-plastic mode 2 system, the dynamics largely shifts from mode 2 to mode 1. The same was found with other non-plastic systems exhibiting higher modes: the introduction of STDP generally shifts the mode of the system down to one. Finally, we would like to note that there are several points in the parameter space (see **Figures 4A**, **5A**) that have very large modes. For example, in **Figure 4A** when $A = 0.0006, k = 0.01$, the resulting system is mode 38 (i.e., most common desynchronizations are very long). Generally, these cases have a wide distribution of desynchronization durations. Therefore, while these systems have a large mode, the mode does not present a strong tendency in the distribution. Nevertheless, these situations are relatively rarely found.

Variation of β

The parameter β changes the widths of the voltage-dependent time-constant of activation $\tau\ (v)$ and the width of the steady-state activation function $w_\infty(v)$ for potassium current. In particular, as β is decreased, the slope at the half-height of $w_\infty(v)$ is increased, and this decreases the width of the step ($w_\infty(v)$ is a sigmoidal function). Similarly, for $\tau(v)$, a decrease in β decreases the width of the function around the peak. This causes an advancement in the activation of the potassium current.

A larger value of β promotes shorter desynchronization durations (Ahn and Rubchinsky, 2017). For $\beta = 0.124$, the non-plastic system is mode 1. The effect of STDP on this system is presented in **Figure 6**. Across the top and in the upper left corner of **Figure 6A** we see that virtually every point corresponds to a mode 1 system, as expected. Indeed, a substantial portion of the entire parameter space displays mode 1 dynamics; about 80% of the parameter space studied.

To illustrate the effect of plasticity on a distribution of desynchronization durations, refer to **Figures 6B–D**. Plasticity effects increase from left to right. The introduction of plasticity has a minimal effect on the distribution; there is very little change visibly. Indeed, the proportion of desynchronization durations of length one increases with plasticity.

Decreasing β increases the mode of a system. If $\beta = 0.091$, the non-plastic system is mode 2. With the introduction of very weak plasticity (across the top and the upper left corner of **Figure 7A**) we see that the dynamics are relatively unchanged, i.e., the mode of most systems remains two. However, if plasticity is not very weak, the dynamics shift to mode 1 in a significant portion of the parameter space. The effect is not as substantial as in the previous section, but about 35% of parameter space becomes mode 1 (about 45% remains mode 2, i.e., the mode is unchanged).

To illustrate the effect of plasticity on a distribution of desynchronization durations, refer to **Figures 7B–D**. Plasticity effects increase from left to right. We see that the vast majority

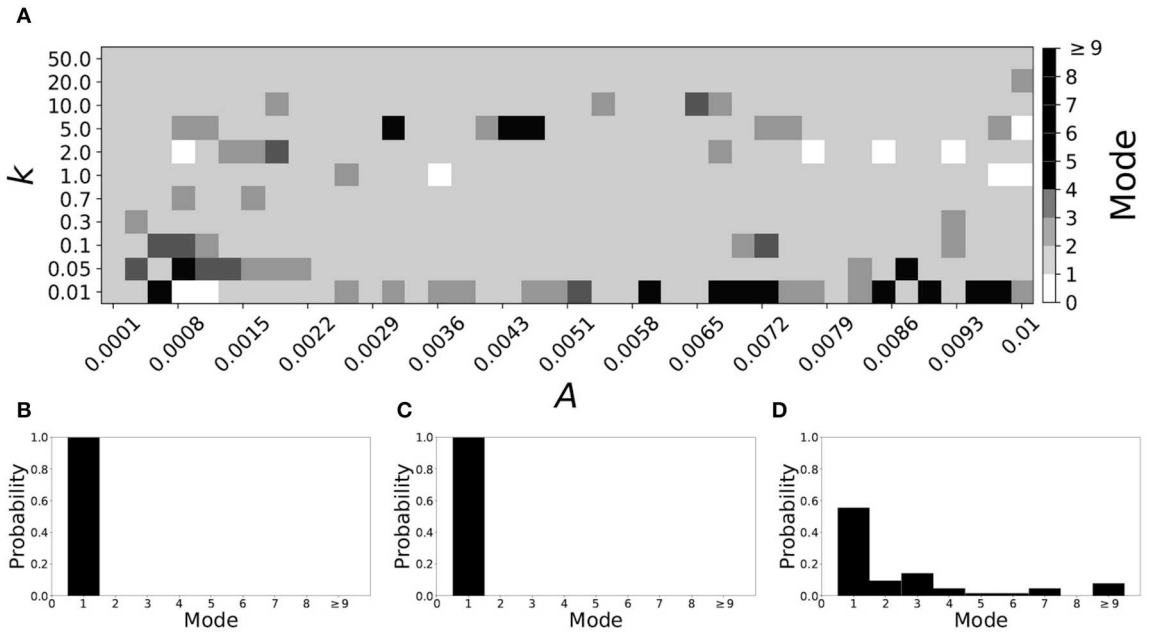

FIGURE 4 | A system exhibiting mode 1 dynamics in the non-plastic case is subjected to plasticity ($\varepsilon = 0.05$). **(A)** Mode is colored via gray scale, see legend on the right of the diagram. The amplitude of the synaptic update, A, is varied along the horizontal axis. The reciprocal of the time-scale of the synaptic update, k, is varied along the vertical axis. **(B–D)** show the changes in the histogram of desynchronization durations as plasticity becomes stronger. **(B)** The system without plasticity. **(C)** The system with very weak plasticity: $A = 0.0047,\ k = 20.0$. **(D)** The system with moderate plasticity: $A = 0.0047,\ k = 0.05$.

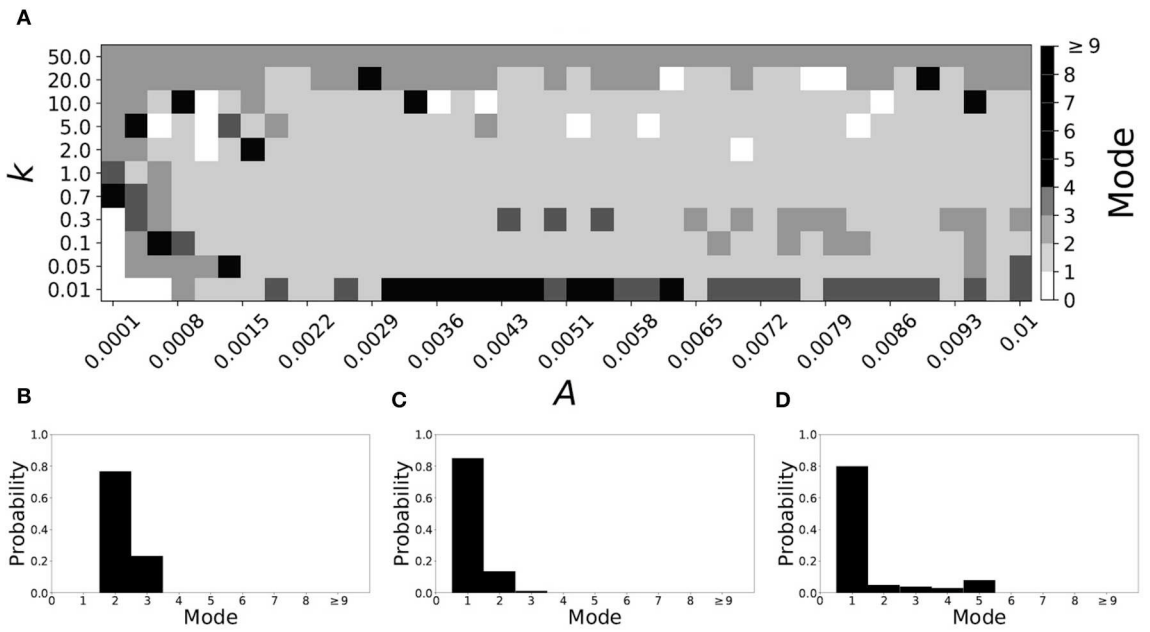

FIGURE 5 | A system exhibiting mode 2 dynamics in the non-plastic case is subjected to plasticity ($\varepsilon = 0.15$). **(A)** Mode is colored via gray scale, see legend on the right of the diagram. The amplitude of the synaptic update, A, is varied along the horizontal axis. The reciprocal of the time-scale of the synaptic update, k, is varied along the vertical axis. **(B–D)** show the changes in the histogram of desynchronization durations as plasticity becomes stronger. **(B)** The system without plasticity. **(C)** The system with very weak plasticity: $A = 0.0047,\ k = 20.0$. **(D)** The system with moderate plasticity: $A = 0.0047,\ k = 0.7$.

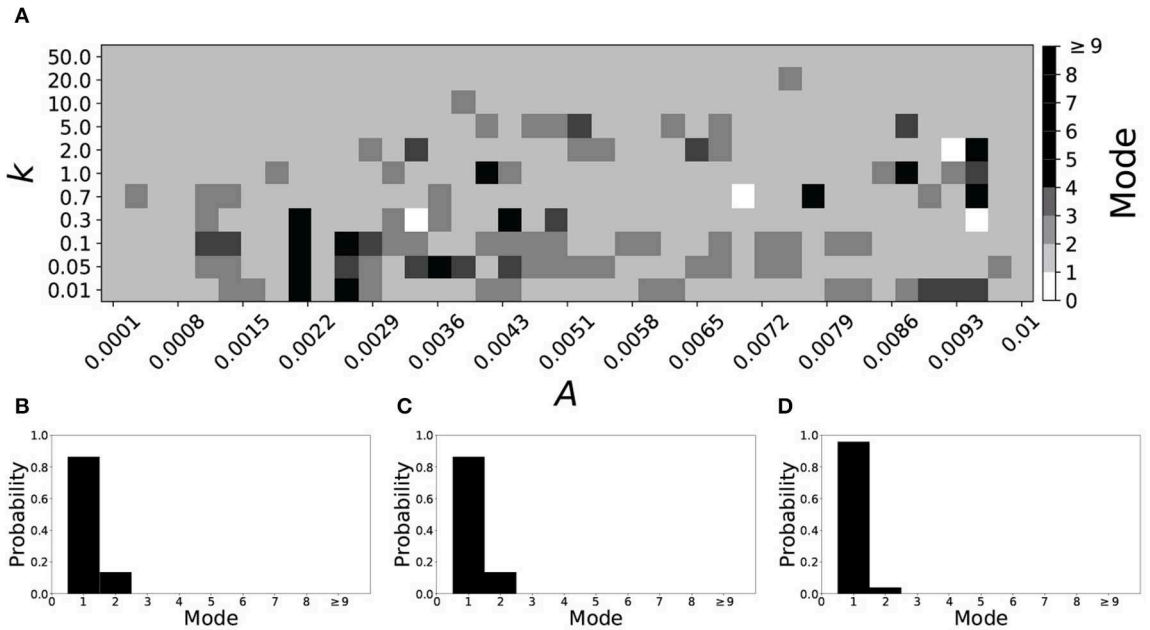

FIGURE 6 | A system exhibiting mode 1 dynamics in the non-plastic case is subjected to plasticity ($\beta = 0.124$). **(A)** Mode is colored via gray scale, see legend on the right of the diagram. The amplitude of the synaptic update, A, is varied along the horizontal axis. The reciprocal of the time-scale of the synaptic update, k, is varied along the vertical axis. **(B–D)** show the changes in the histogram of desynchronization durations as plasticity becomes stronger. **(B)** The system without plasticity. **(C)** The system with very weak plasticity: $A = 0.0052,\ k = 20.0$. **(D)** The system with moderate plasticity: $A = 0.0052,\ k = 0.7$.

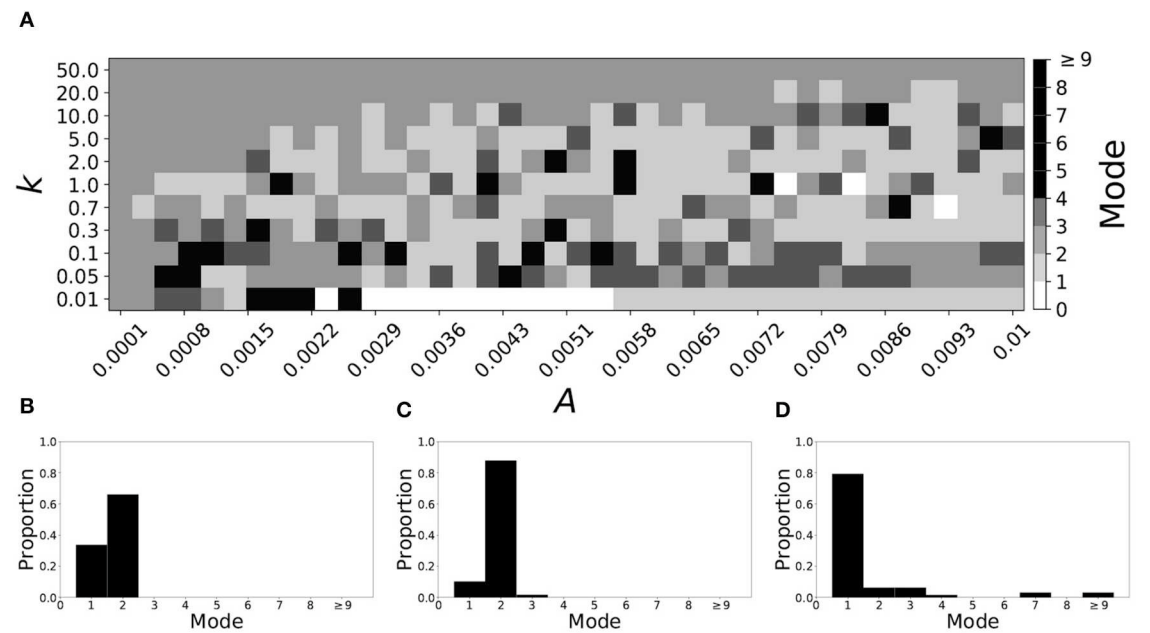

FIGURE 7 | A system exhibiting mode 2 dynamics in the non-plastic case is subjected to plasticity ($\beta = 0.091$). **(A)** Mode is colored via gray scale, see legend on the right of the diagram. The amplitude of the synaptic update, A, is varied along the horizontal axis. The reciprocal of the time-scale of the synaptic update, k, is varied along the vertical axis. **(B–D)** show the changes in the histogram of desynchronization durations as plasticity becomes stronger. **(B)** The system without plasticity. **(C)** The system with very weak plasticity: $A = 0.0047,\ k = 20.0$. **(D)** The system with moderate plasticity: $A = 0.0047,\ k = 0.7$.

of desynchronization durations become length one as plasticity becomes stronger.

Variation of ν_{w1}

The parameter ν_{w1} affects a horizontal translation in $w_\infty(\nu)$ and $\tau(\nu)$. Increasing ν_{w1} shifts both curves to the right, i.e., toward higher voltages; this results in a potassium current that activates faster.

Smaller values of ν_{w1} result in short desynchronization durations (Ahn and Rubchinsky, 2017). For $\nu_{w1} = 0.102$, the non-plastic system is mode 1. The effect of STDP on this system is presented in **Figure 8**. We see that mode 1 dynamics is observed not only for the weak plasticity region (top and upper left corner of **Figure 8A**), but for most of the parameter space (about 85% of the parameter space studied).

To illustrate the effect of plasticity on a distribution of desynchronization durations, refer to **Figures 8B–D**. Plasticity effects increase from left to right. We see that as plasticity increases to a higher level, the prevalence of mode 1 is unchanged.

Varying ν_{w1} to larger values leads to shorter desynchronization durations becoming less prevalent. For $\nu_{w1} = 0.161$, the non-plastic system is mode 2. The effect of STDP is presented in **Figure 9**. When plasticity is added we see that the dynamics are similar to the non-plastic case when plasticity is weak enough (top and upper left corner of **Figure 9A**). However, when the plasticity effects are moderate, the system exhibits mode 1 dynamics frequently (central region of **Figure 9A**). For the domain of parameter space studied, the majority of points (about 45%) correspond to mode 1 systems, the rest are either mode 2 (about 40%) or higher.

To illustrate the effect of plasticity on a distribution of desynchronization durations, refer to **Figures 9B–D**. Plasticity effects increase from left to right. We see that the mode of the system shifts down from two to one as plasticity becomes stronger.

Variation of β_w and β_τ

Varying either ε, β, or ν_{w1} may affect not only the durations of the desynchronizations, but also synchronization strength and the frequency of activity in the system. To control desynchronization durations while keeping both spiking frequency and synchronization strength near constant in a non-plastic system, one can consider the parameter β and separate it into two independent parameters, β_τ and β_w. As a result, the lengths of desynchronization events are almost independent of the frequency and synchrony strength (Ahn and Rubchinsky, 2017).

Smaller β_w and larger β_τ result in shorter desynchronization durations (Ahn and Rubchinsky, 2017). For $\beta_w = 0.098$, $\beta_\tau = 0.079$, the non-plastic system is mode 1. **Figure 10** illustrates the impact of STDP on this system. Mode 1 dynamics is observed not only for the weak plasticity region (top and upper left corner of **Figure 10A**), but for the majority of the parameter space (about 60% of the parameter space studied).

To illustrate the effect of plasticity on a distribution of desynchronization durations, refer to **Figures 10B–D**. Plasticity effects increase from left to right. We see that as plasticity progresses to a moderate level, the proportion of short desynchronizations stays largely unchanged. In particular, the system is still mode 1.

If $\beta_w = 0.115$, $\beta_\tau = 0.071$, the non-plastic system is mode 2. **Figure 11** illustrates the impact of STDP on this system. With the addition of plasticity, we see that the system is largely mode 2 if the plasticity is weak (top and upper left corner of **Figure 11A**). However, stronger plasticity shifts the dynamics to mode 1 for a substantial portion of the parameter space (about 55% of points considered).

To illustrate the effect of plasticity on a distribution of desynchronization durations, refer to **Figures 11B–D**. Plasticity effects increase from left to right. We see that the distribution is largely unchanged for very weak plasticity, but as plasticity increases, the system becomes mode 1.

DISCUSSION

This study considered intermittent synchronous dynamics in a small network of simple conductance-based model neurons. While strong synaptic strength in general can promote synchronization between neurons, moderate values of synaptic coupling lead to dynamics with relatively weak synchronization, and where the episodes of synchronization are interspersed with episodes of desynchronized dynamics. Intermittent synchronization in the presence of moderate (and fixed in time) coupling is quite typical for coupled oscillatory systems (Pikovsky et al., 2001). In other words, temporal variability of correlations is observed due to the relative weakness of a fixed coupling strength. The temporal signatures of this variability have been previously modeled in Ahn and Rubchinsky (2017) and were in good agreement with the analysis of the temporal variability observed in experimental data (see section Introduction and references therein).

However, many actual synapses are plastic and thus the synaptic coupling between neurons experiences temporal variations. This variation may contribute to the temporal variability of intermittent synchrony as well. This study considered how one common type of neural plasticity—spike-timing dependent plasticity—might affect this temporal variability. Experimental data ubiquitously points to the prevalence of short desynchronization dynamics in neural synchrony. This kind of dynamics is naturally generated in synaptically coupled conductance-based model neurons. We showed here that the introduction of STDP under quite general conditions preserves this realistic fine temporal structure of intermittent neural synchrony. Moreover, when the non-plastic system parameters are selected in such a way as to predominantly express longer desynchronizations, STDP changes the intermittently synchronous dynamics back to one with short desynchronizations. This was observed while varying several different parameters, so that STDP may reverse dynamics from long to short desynchronizations regardless of how the desynchronizations were obtained in the non-plastic system.

The overall dependence of the dynamics on the characteristics of plasticity is quite complicated. Numerical simulations

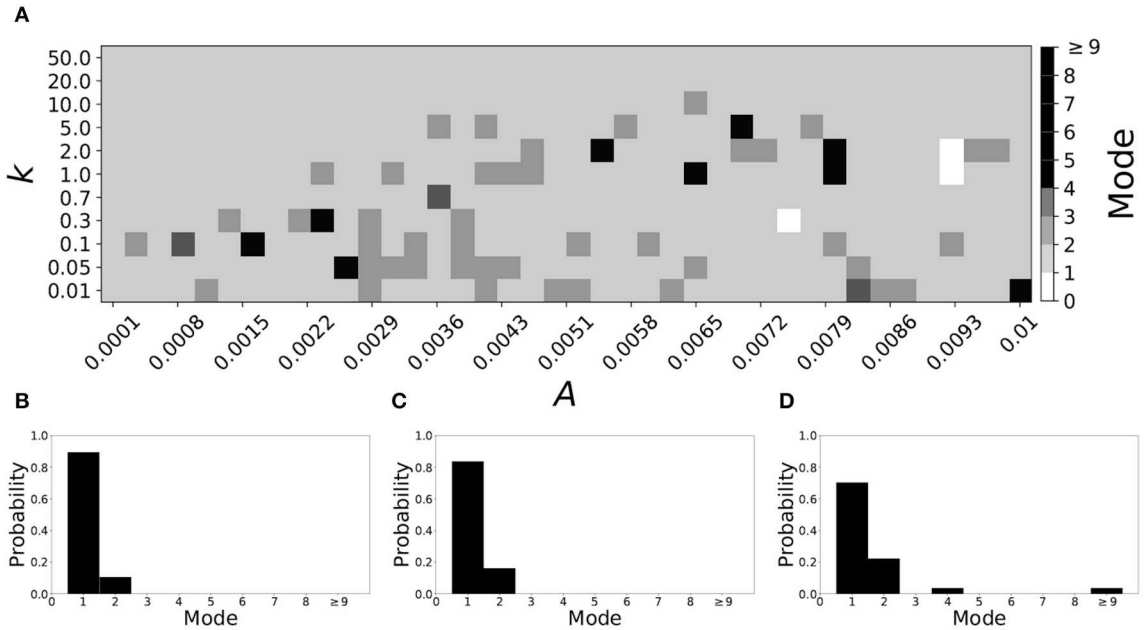

FIGURE 8 | A system exhibiting mode 1 dynamics in the non-plastic case is subjected to plasticity ($v_{w1} = 0.102$). **(A)** Mode is colored via gray scale, see legend on the right of the diagram. The amplitude of the synaptic update, A, is varied along the horizontal axis. The reciprocal of the time-scale of the synaptic update, k, is varied along the vertical axis. **(B–D)** show the changes in the histogram of desynchronization durations as plasticity becomes stronger. **(B)** The system without plasticity. **(C)** The system with very weak plasticity: $A = 0.0047, k = 20.0$. **(D)** The system with moderate plasticity: $A = 0.0047,\ k = 0.7$.

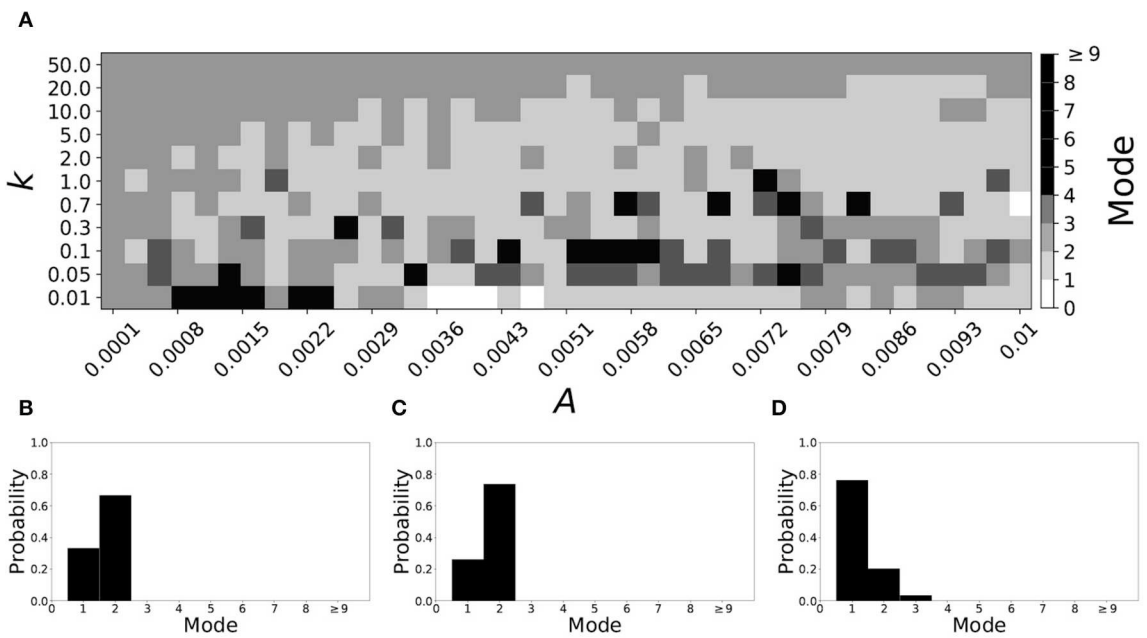

FIGURE 9 | A system exhibiting mode 2 dynamics in the non-plastic case is subjected to plasticity ($v_{w1} = 0.161$). **(A)** Mode is colored via gray scale, see legend on the right of the diagram. The amplitude of the synaptic update, A, is varied along the horizontal axis. The reciprocal of the time-scale of the synaptic update, k, is varied along the vertical axis. **(B–D)** show the changes in the histogram of desynchronization durations as plasticity becomes stronger. **(B)** The system without plasticity. **(C)** The system with very weak plasticity: $A = 0.0047,\ k = 20.0$. **(D)** The system with moderate plasticity: $A = 0.0054,\ k = 1.0$.

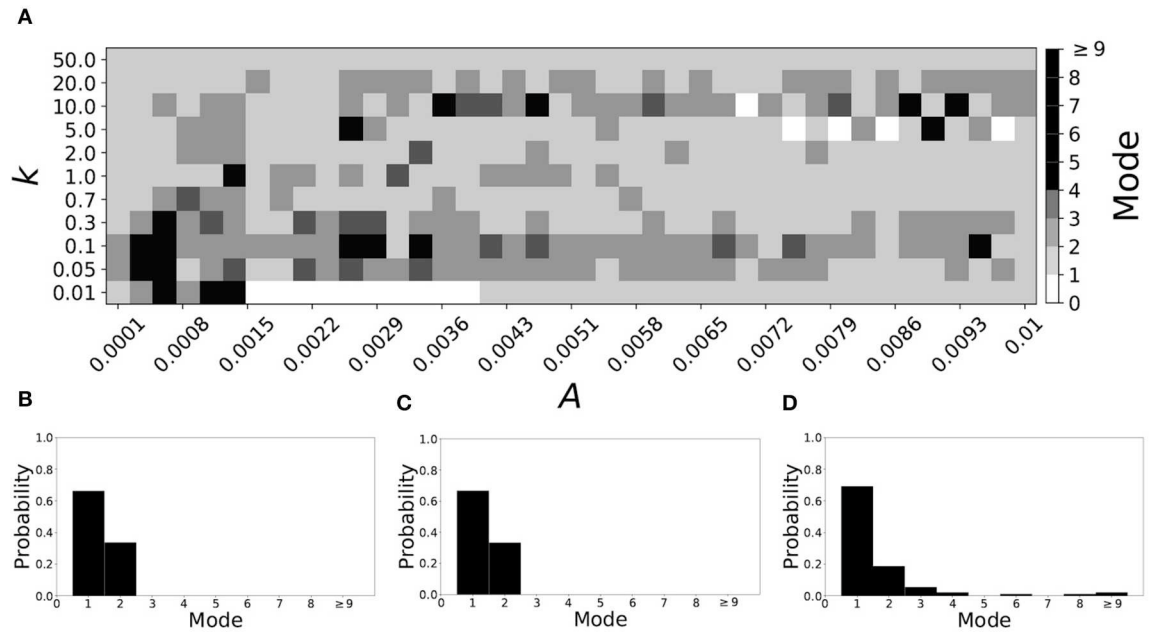

FIGURE 10 | A system exhibiting mode 1 dynamics in the non-plastic case is subjected to plasticity ($\beta_w = 0.098,\ \beta_\tau = 0.079$). **(A)** Mode is colored via gray scale, see legend on the right of the diagram. The amplitude of the synaptic update, A, is varied along the horizontal axis. The reciprocal of the time-scale of the synaptic update, k, is varied along the vertical axis. **(B–D)** show the changes in the histogram of desynchronization durations as plasticity becomes stronger. **(B)** The system without plasticity. **(C)** The system with very weak plasticity: $A = 0.0049,\ k = 50.0$. **(D)** The system with moderate plasticity: $A = 0.0052,\ k = 0.7$.

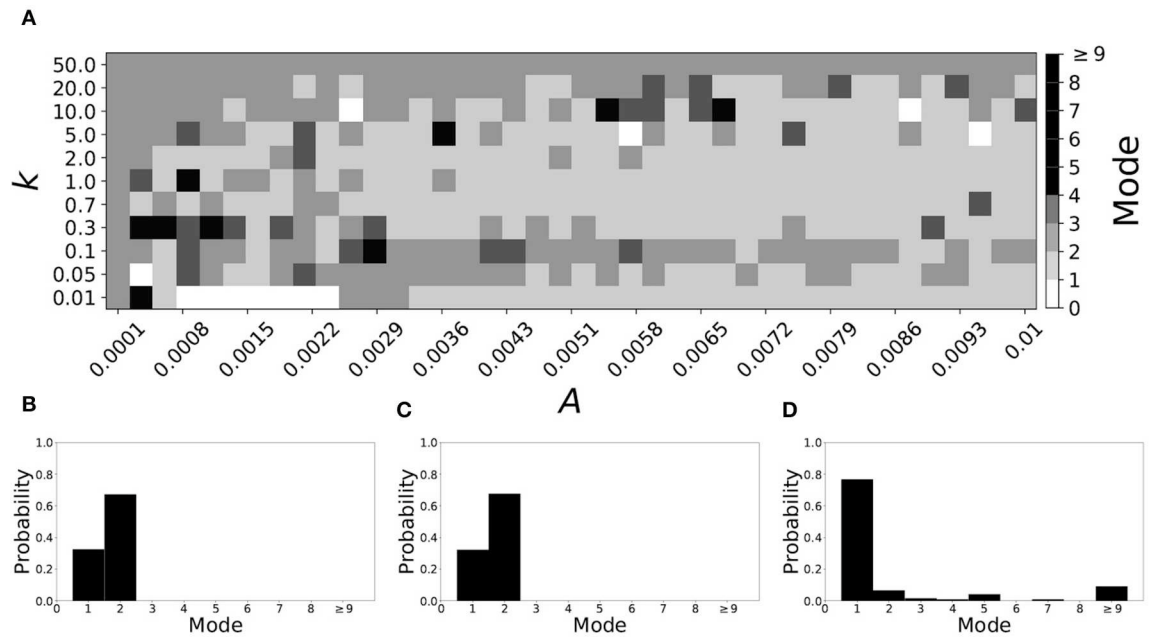

FIGURE 11 | A system exhibiting mode 2 dynamics in the non-plastic case is subjected to plasticity ($\beta_w = 0.115,\ \beta_\tau = 0.071$). **(A)** Mode is colored via gray scale, see legend on the right of the diagram. The amplitude of the synaptic update, A, is varied along the horizontal axis. The reciprocal of the time-scale of the synaptic update, k, is varied along the vertical axis. **(B–D)** show the changes in the histogram of desynchronization durations as plasticity becomes stronger. **(B)** The system without plasticity. **(C)** The system with very weak plasticity: $A = 0.0049,\ k = 50.0$. **(D)** The system with moderate plasticity: $A = 0.0054,\ k = 0.7$.

indicate that some plasticity parameter values may promote very unrealistic synchronized dynamics. However, under the conditions considered, the short desynchronization dynamics were obtained in large regions of the parameter space. This was regardless of whether the corresponding non-plastic system was mode 1, or had a higher mode.

The results of these numerical simulations suggest that STDP may be one of the contributing factors behind experimentally observed short desynchronization dynamics. Moreover, STDP and cellular mechanisms proposed in Ahn and Rubchinsky (2017) may act cooperatively in promoting short desynchronizations.

The results discussed here were obtained in the framework of relatively simple modeling. The actual neuronal synchrony is, of course, a much more complicated phenomenon than the model considered here, and there were multiple factors not included in the model. For example, inhibitory synapses (e.g., see Nowotny et al., 2008). The experimental observations of short desynchronizations were mostly done with LFP and EEG signals, and the simple network considered here is too simple to adequately model these signals. However, the similarity between experimentally observed intermittent neural synchrony and the temporal patterning of synchrony observed in our study with a relatively simple model with STDP may speak to the very general nature of this phenomenon.

The variability of the dynamics on short time-scales may be a functionally beneficial phenomenon. Short desynchronization dynamics (which is essentially a high degree of variability of synchrony on very short time-scales) have been conjectured to be conducive for quick and efficient formation and break-up of neural assemblies (Ahn and Rubchinsky, 2013, 2017). As was noted in these studies, the ease of formation and disappearance of synchronized states at rest may suggest that a transient synchronized assembly may be easily formed whenever needed to facilitate a particular function. The results of this study suggest that the temporal variability of synaptic strength due to STDP may potentially further facilitate this phenomenon.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

LR conceived research. LR and JZ designed research, analyzed and interpreted the results, and wrote the manuscript. JZ performed numerical simulations. Both authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by NSF DMS 1813819.

REFERENCES

Ahn, S., Park, C., and Rubchinsky, L. L. (2011). Detecting the temporal structure of intermittent phase locking. *Phys. Rev. E Stat. Nonlinear Soft Matter Phys.* 84:016201. doi: 10.1103/PhysRevE.84.016201

Ahn, S., and Rubchinsky, L. L. (2013). Short desynchronization episodes prevail in synchronous dynamics of human brain rhythms. *Chaos* 23:013138. doi: 10.1063/1.4794793

Ahn, S., and Rubchinsky, L. L. (2017). Potential mechanisms and functions of intermittent neural synchronization. *Front. Comput. Neurosci.* 11:44. doi: 10.3389/fncom.2017.00044

Ahn, S., Rubchinsky, L. L., and Lapish, C. C. (2014). Dynamical reorganization of synchronous activity patterns in prefrontal cortex - hippocampus networks during behavioral sensitization. *Cereb. Cortex* 24, 2553–2561. doi: 10.1093/cercor/bht110

Ahn, S., Zauber, S. E., Witt, T., Worth, R. M., and Rubchinsky, L. L. (2018). Neural synchronization: average strength vs. temporal patterning. *Clin. Neurophysiol.* 129, 842–844. doi: 10.1016/j.clinph.2018.01.063

Buzsáki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929. doi: 10.1126/science.1099745

Cassenaer, S., and Laurent, G. (2007). Hebbian STDP in mushroom bodies facilitates the synchronous flow of olfactory information in locusts. *Nature* 448, 709–713. doi: 10.1038/nature05973

Ermentrout, G. B., and Terman, D. H. (2010). *Mathematical Foundations of Neuroscience*. New York, NY: Springer. doi: 10.1007/978-0-387-87708-2

Feldman, D. (2012). The spike timing dependence of plasticity. *Neuron* 75, 556–571. doi: 10.1016/j.neuron.2012.08.001

Fell, J., and Axmacher, N. (2011). The role of phase synchronization in memory processes. *Nat. Rev. Neurosci.* 12, 105–118. doi: 10.1038/nrn2979

Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235. doi: 10.1016/j.neuron.2015.09.034

Harris, A. Z., and Gordon, J. A. (2015). Long-range neural synchrony in behavior. *Annu. Rev. Neurosci.* 38, 171–194. doi: 10.1146/annurev-neuro-071714-034111

Izhikevich, E. M. (2007). *Dynamical Systems in Neuroscience: The Geometry of Excitability and Bursting*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/2526.001.0001

Knoblauch, A., Hauser, F., Gewaltig, M. O., Körner, E., and Palm, G. (2012). Does spike-timing-dependent synaptic plasticity couple or decouple neurons firing in synchrony? *Front. Comput. Neurosci.* 6:55. doi: 10.3389/fncom.2012.00055

Malaia, E., Ahn, S., and Rubchinsky, L. L. (2020). Dysregulation of temporal dynamics of synchronous neural activity in adolescents on autism spectrum. *Autism Res.* 13, 24–31. doi: 10.1002/aur.2219

Nowotny, T., Huerta, R., and Rabinovich, M. I. (2008). Neural synchrony: peculiarity and generality. *Chaos* 18:037119. doi: 10.1063/1.2949925

Nowotny, T., Zhigulin, V. P., Selverston, A. I., Abarbanel, H. D., and Rabinovich, M. I. (2003). Enhancement of synchronization in a hybrid neural circuit by spike-timing dependent plasticity. *J. Neurosci.* 23, 9776–9785. doi: 10.1523/JNEUROSCI.23-30-09776.2003

Oswal, A., Brown, P., and Litvak, V. (2013). Synchronized neural oscillations and the pathophysiology of Parkinson's disease. *Curr. Opin. Neurol.* 26, 662–670. doi: 10.1097/WCO.0000000000000034

Park, C., Worth, R. M., and Rubchinsky, L. L. (2010). Fine temporal structure of beta oscillations synchronization in subthalamic nucleus in Parkinson's disease. *J. Neurophysiol.* 103, 2707–2716. doi: 10.1152/jn.00724.2009

Pikovsky, A., Rosenblum, M., and Kurths, J. (2001). *Synchronization: A Universal Concept in Nonlinear Sciences*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511755743

Pittman-Polletta, B. R., Kocsis, B., Vijayan, S., Whittington, M. A., and Kopell, N. J. (2015). Brain rhythms connect impaired inhibition to

altered cognition in schizophrenia. *Biol. Psychiatry* 77, 1020–1030. doi: 10.1016/j.biopsych.2015.02.005

Rabinovich, M., Huerta, R., and Laurent, G. (2008). Neuroscience. Transient dynamics for neural processing. *Science* 321, 48–50. doi: 10.1126/science.1155564

Ratnadurai-Giridharan, S., Khargonekar, P. P., and Talathi, S. S. (2015). Emergent gamma synchrony in all-to-all interneuronal networks. *Front. Comput. Neurosci.* 9:127. doi: 10.3389/fncom.2015.00127

Ratnadurai-Giridharan, S., Zauber, S. E., Worth, R. M., Witt, T., Ahn, S., and Rubchinsky, L. L. (2016). Temporal patterning of neural synchrony in the basal ganglia in Parkinson's disease. *Clin. Neurophysiol.* 127, 1743–1745. doi: 10.1016/j.clinph.2015.09.005

Schnitzler, A., and Gross, J. (2005). Normal and pathological oscillatory communication in the brain. *Nat. Rev. Neurosci.* 6, 285–296. doi: 10.1038/nrn1650

Uhlhaas, P. J., and Singer, W. (2006). Neural synchrony in brain disorders: relevance for cognitive dysfunctions and pathophysiology. *Neuron* 52, 155–168. doi: 10.1016/j.neuron.2006.09.020

Zhang, L., Tao, H. W., Holt, C. E., Harris, W. A., and Poo, M. (1998). A critical window for cooperation and competition among developing retinotectal synapses. *Nature* 395, 37–44. doi: 10.1038/25665

Zhigulin, V. P., Rabinovich, M. I., Huerta, R., and Abarbanel, H. D. (2003). Robustness and enhancement of neural synchronization by activity-dependent coupling. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 67:021901. doi: 10.1103/PhysRevE.67.021901

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.

frontiers in Computational Neuroscience

HYPOTHESIS AND THEORY
published: 23 June 2020
doi: 10.3389/fncom.2020.00046

Object Recognition at Higher Regions of the Ventral Visual Stream via Dynamic Inference

Siamak K. Sorooshyari[1]*, *Huanjie Sheng*[1†] and *H. Vincent Poor*[2]

[1] Department of Integrative Biology, University of California, Berkeley, Berkeley, CA, United States, [2] Department of Electrical Engineering, Princeton University, Princeton, NJ, United States

The ventral visual stream (VVS) is a fundamental pathway involved in visual object identification and recognition. In this work, we present a hypothesis of a sequence of computations performed by the VVS during object recognition. The operations performed by the inferior temporal (IT) cortex are represented as not being akin to a neural-network, but rather in-line with a dynamic inference instantiation of the untangling notion. The presentation draws upon a technique for dynamic maximum a posteriori probability (MAP) sequence estimation based on the Viterbi algorithm. Simulation results are presented to show that the decoding portion of the architecture that is associated with the IT can effectively untangle object identity when presented with synthetic data. More importantly, we take a step forward in visual neuroscience by presenting a framework for an inference-based approach that is biologically inspired via attributes implicated in primate object recognition. The analysis will provide insight in explaining the exceptional proficiency of the VVS.

Keywords: object recognition, sequence estimation, decoding, IT cortex, dynamic inference, Viterbi algorithm

OPEN ACCESS

Edited by:
Tatyana Sharpee,
Salk Institute for Biological Studies, United States

Reviewed by:
Guy Elston,
University of the Sunshine Coast, Australia
Jan Lauwereyns,
Kyushu University, Japan

***Correspondence:**
Siamak K. Sorooshyari
siamak_sorooshyari@berkeley.edu

†Present address:
Huanjie Sheng,
Roche Sequencing Solutions, Inc., Santa Clara, CA, United States

Received: *06 January 2020*
Accepted: *30 April 2020*
Published: *23 June 2020*

Citation:
Sorooshyari SK, Sheng H and Poor HV (2020) Object Recognition at Higher Regions of the Ventral Visual Stream via Dynamic Inference. Front. Comput. Neurosci. 14:46. doi: 10.3389/fncom.2020.00046

1. INTRODUCTION

A prevalent hypothesis is that the identities of viewed objects are represented as patterns of activity across populations of neurons with increasingly complicated computations occurring further along the ventral visual stream (VVS). Presenting a biologically inspired algorithm where the stimulus information is processed and exchanged among different populations of neurons is a challenge. Since the higher visual areas such as inferior temporal (IT) cortex are selective to the more complex stimuli characteristics than populations in lower levels such as V1 and V2, it has been postulated that more complicated processing techniques are used by the IT (Riesenhuber and Poggio, 1999). The term "encoding" has been applied extensively to the manner by which neurons in the early visual stages respond to and represent stimuli. The presented analysis will treat the object recognition process performed by the higher regions of the VVS as a decoding operation and present a model that can commence to unify an understanding of the computations involved during such a cognitive process. Topical overviews such as DiCarlo et al. (2012) advocate the first step of unequivocally defining the question of how the brain solves the problem. It is sensible to presume that as large amounts of data become available the object recognition question will be asked in different ways. Computer vision algorithms have been lauded for efficacy in categorizing objects after being trained on large sets of sample data. However, they are also known to suffer from the invariance problem that has been studied by visual neuroscientists. This is especially true when

a large number of object categories are considered, and imparts one to question whether computer vision models are the optimal means for studying the computational operations performed by the brain during real-time object recognition. The encoding-decoding methodology discussed in this work will provide a model that is a closer, biologically-plausible explanation of the VVS operation.

Neurons at progressive stages of neuroanatomy receiving weighted excitatory and inhibitory inputs prior to their state being subject to a thresholding operation is not a new concept to vision neuroscience. We consider somewhat more sophisticated operations that will occur over several populations of neurons. While the algorithmic operations may be deemed sophisticated, it is noteworthy that such operations are being performed by millions of neurons. Furthermore, the fact that primates are extremely efficient in conducting object recognition vindicates the use of algorithms to explain the seemingly effortless manner by which the recognition is performed. The input to the model will be the representation that the viewed object should evoke at the IT. This representation is obviously associated with the visual stimulus, and is immediately encoded by the retina and lateral geniculate nucleus (LGN) circuitry in order for its meaning to be communicated along the VVS in a reliable manner. The model presented in this work provides an alternative to neural network techniques employing max-pooling, and an alternative to machine learning approaches that consider object categorization rather than classification of object attributes during the recognition process. The analysis additionally brings forth the question of what metrics to consider in assessing how well a model performs object recognition. Within the encoding-decoding framework it is possible to distinguish between different gradations of recognition. Specifically, one would be able to quantify the error rate in recognizing objects, the attributes of an object, and the object category.

Algorithmic operations will be suggested herein for various stages of the VVS to mirror the functional operations implicated by prior works in visual neuroscience. The algorithmic structure in **Figure 1** is novel within the context of visual neuroscience. The biologically-inspired system will be referred to as the communication-theoretic object recognition (CTOR) model and will encompass high-level visual function processing low-level sensory signals. A natural impetus for the derivation of CTOR is the brain consisting of communication channels with a task such as object recognition invoking the interchange of signals between neural circuits as part of the interplay between top-down and bottom-up processing. There are several themes that subsist when considering statistical inference on the output of a non-ideal channel in engineering or biology: the time-sensitive nature of the information, the presence of stochastic perturbations, and the possible compression of the recovered information. Refinements of CTOR that may spawn from this presentation will need to include a decoding algorithm for inference. Indeed, alternate decoding algorithms may be proposed and different definitions for the elements that comprise the decoded sequence may emerge.

2. ATTRIBUTES OF BIO-INSPIRED OBJECT RECOGNITION MODELS

It has been reputed that neural connectivity dictates a hierarchical organization at the VVS with visual information traversing the retina to the LGN, and then through cortical area V1, V2, and V4 before reaching the IT. Neurons in V1 have small receptive fields and respond to simple features such as edge orientation (Hubel and Wiesel, 1962). The receptive fields of V4 neurons are on average four to seven times greater than those in V1, but are smaller than the receptive fields of IT neurons. Many V4 neurons are sensitive to stimulus features of moderate complexity (Cadieu et al., 2007), whereas the IT neurons are selective to much more complex stimuli such as faces. The tuning properties of IT cells seem to be shaped by task learning with their dendritic arbors being more expansive than those of V1, V2, or V4 neurons (Elston, 2002; Luebke, 2017). The untangling notion advocated in DiCarlo et al. (2012) serves as motivation for the decoding module in CTOR. As the viewed object is processed beyond the retina and along the successive stages of the VVS, it is believed that increasingly sophisticated processing power is applied to untangle the object's identity. When considering the statistics of the input to the lowest stage in the model, works such as Simoncelli and Olshausen (2001) have provided a litany of studies that contain empirical evidence for the non-Gaussianity of natural images. The authors proceed to describe the neural coding/representation that occurs in portions of the visual cortex. An array of works have discussed attributes of the visual cortex that enable the system to be exceptionally proficient at performing object recognition in a rapid and effortless manner. The following are what we consider the most crucial attributes that a biologically-inspired model for object recognition should address.

- *Selectivity*: The ability to accurately discriminate between different objects. Object recognition models typically do not quantitatively distinguish between object identification and categorization. The model herein will distinguish between the two domains and focus on the identification of an object rather than a rapid categorization.
- *Invariance*: The ability to recognize an object under transformations such as scale or position alterations in the field of view. Furthermore, inconsiderable alternations in the object's features should not preclude recognition.
- *Robustness*: Aspects of the viewed stimulus such as illumination and clutter may decrease the signal-to-noise ratio (SNR) of the neural signals communicated along the lower visual stages. The VVS is frequently able to distinguish among objects in light of perturbations to the viewed object that reduce the SNR of the neural information progressing along the pathway.
- *Processing Speed*: The recognition of an object within either a strict or lax temporal constraint imposed by the task. From a psychophysics perspective, the processing speed corresponds to how rapidly the object recognition is performed by the brain.

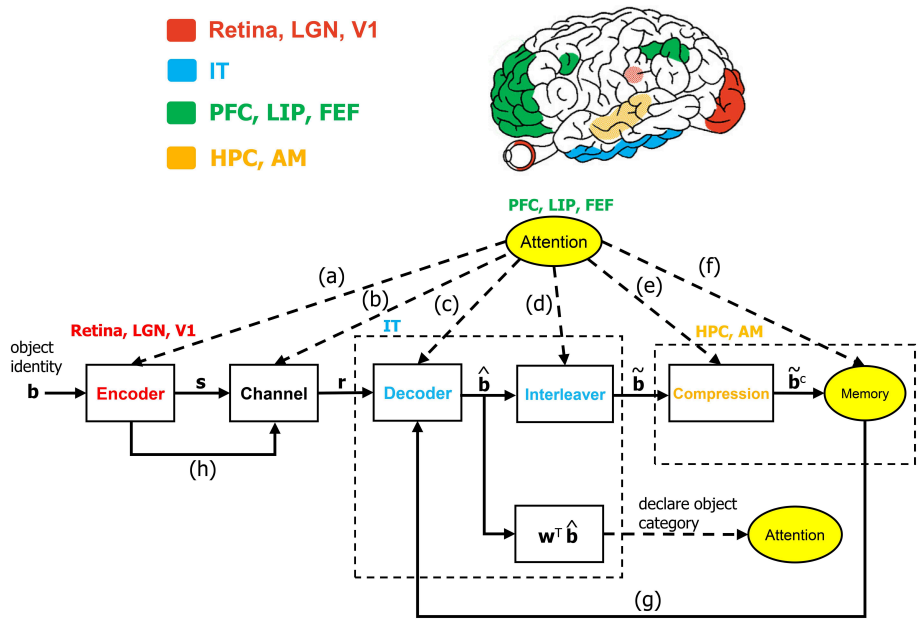

FIGURE 1 | The communication-theoretic object recognition (CTOR) model. The structure of CTOR consists of feedforward processing with memory and attention providing feedback to the decoder portion of the IT. **(a)** The communication of T, $\{\alpha_i\}$, and $\{G_i(t)\}$ of equations (1) and (2) from the PFC, LIP, and FEF regions to V1, V2, and V4. **(b)** Attention modulating the channel properties. **(c,g)** The communication of the transition probabilities $\{P[\mathbf{r}_k|\mathbf{b}_{k+1}, \mathbf{b}_k]\}$ and the priors $\{P[\mathbf{b}_{k+1}|\mathbf{b}_k]\}$ to the IT from the attention and memory circuitry, respectively. **(d)** The conveying of the number of features F and the number of bits allocated to each feature $\{M_i\}$ by attention to the interleaver. **(e)** Conveyance of the degree of compression from attention. **(f)** Attention gating memory as far as the object features that are retained following recognition. **(h)** The contribution of the retinal, LGN, and V1 stages to the neural noise process constituting the channel. The corresponding brain regions are marked in the brain above the diagram. The units drawn with solid lines are modeled by algorithmic operations. Filled areas represent brain regions on the surface, while shaded areas represent those embedded inside the brain. LGN, lateral geniculate nucleus; V1, primary visual cortex; IT, inferior temporal cortex; LIP, lateral intraparietal area; FEF, frontal eye fields; HPC, hippocampus; AM, amygdala.

- *Attentional Gating*: The degree and implications of attention allocated to recognizing an object. The dynamics of the allocated attention will govern how the brain parses object features and what is retained following recognition.
- *Dynamic Recurrence*: The consideration of feedback as a necessary complement to the feedforward processing. The recurrence should be dynamic and involve interaction between multiple brain areas.

The first two attributes have been discussed in works such as Serre and Riesenhuber (2004), whereas robustness has been considered in a multitude of studies (e.g., Cadieu et al., 2007). The processing speed was elegantly discussed in Thorpe and Van Rullen (2001), while attentional modulation during object recognition has also been extensively investigated in the literature. The processing speed and attentional gating attributes will have analogues in CTOR. Dynamic recurrence in the VVS during object recognition has been experimentally instantiated by works such as Wyatte et al. (2012), O'Reilly et al. (2013), and Poggio and Kreiman (2013). It seems natural for the brain to take advantage of feedback pathways to coordinate between top-down and bottom-up signals during more challenging recognition tasks such as object completion or identification in the presence of clutter. In fact, studies on neural circuit specialization and connectivity have discussed areas V1 and V2 receiving connections from IT and parahippocampal regions (Rockland, 1997). Consideration of the above attributes presents an avenue to discuss how CTOR is a bio-inspired model for object recognition at the VVS. Primate circuits such as the cerebral cortex, hippocampus, and amygdala are associated with advanced cognitive functions and have been shown to contain pyramidal neurons whose architecture seem to be specialized for the posited task of such neural circuits (Jacobs and Scheibel, 2002; Elston, 2003). Interestingly, substantial differences are noted in the number of spines on the basal dendritic fields of neurons in V1, V2, and IT with the quantity and density multiplicatively increasing when progressing from V1 to IT. This is believed to lead to the increased capability of pyramidal neurons in the latter stages of information processing such as the IT and PFC to integrate a broader range of synaptic inputs than neurons at the lower cortical areas such as V1 and V2 (Elston et al., 1999). Thus, the anatomy and connectivity of the cortical circuitry are crucial in determining any prospective computation (Elston, 2003; Spruston, 2008; Luebke, 2017). The intriguing discussion of Biederman (1987) brought forth the recognition-by-components (RBC) view of vision where it was suggested that the brain parses viewed objects into parts. Partial matches among the segments are then possible, and the proportion of the similarity

in the components between the viewed object and a stored representation is used to assess the fidelity of the match. Elements of CTOR have been motivated by the valuable discussion in (Biederman, 1987) and the presented model aims to further concretize RBC.

3. MODELING OBJECT RECOGNITION AS DYNAMIC INFERENCE

It is reputed that a study of how the neural populations of the visual system process scenes so that the brain is capable of object recognition leads to an overcomplete problem. In a nominal example an information-rich scene is presented to a subject with an object of interest embedded in the scene. Regardless of the object's salience, the subject has been provided with a plethora of visual information for the prospective task. The hierarchical and non-linear nature of the layers that govern the computations among simple and cortical complex cells implicate the difficulty of formulating optimization functions that the visual system may be attempting to minimize/maximize during such a nominal task. It has also been argued that the difficulty in attempting to mimic functions of the visual cortex is further complicated by its columnar organization and the heterogeneity among the columns Roe (2019). In light of this, works such as Serre et al. (2007) have motivated the approach of studying each layer in the system separately. We believe that a graceful unison should exist between the two disparate avenues of viewing the system as a whole and dividing it into disjoint units. **Figure 1** depicts the architecture that will be motivated as a sensible hypothesis for high-level computational processing occurring in the VVS during object recognition. The conjecture is unique since it is biologically inspired to reflect the VVS's operation while concomitantly being an ideology borrowed from communication theory. From a communication-theoretic perspective, the seminal work of Shannon (1948) has led to countless developments in the design of structured redundancy applied to information that is conveyed over a noisy channel to a receiver with processing capability. The transmission of such structured redundancy is often perturbed in a stochastic manner by a channel prior to it being decoded, or more appropriately for this presentation, "untangled" by the destination. The necessary background on the encoder-channel-decoder structure within a communication-theoretic setting has been provided in Fano (1963) as well as classical texts such as Wozencraft and Jacobs (1965).

It is evident that psychological processes such as attention and memory are prerequisites for visual perception. There is a wealth of literature on the computational capacity of cortical circuitry and the quantitative differences among the population of neurons associated with vision—see Elston, 2002; Jacobs and Scheibel, 2002; Spruston, 2008; Elston and Fujita, 2014; Luebke, 2017 for reviews. The work of Mishkin has provided clear evidence for the inclusion of the hippocampus and amygdala in the so-called recognition memory circuitry. In fact, Mishkin (1982) concludes that a model of object recognition would be incomplete without considering recognition memory and the corresponding feedback and feedforward projections to the hippocampus and amygdala. Furthermore, the pyramidal neurons present in the visual cortex are also seen in the hippocampus and the amygdala (Feldman, 1984; DeFelipe and Farinas, 1992). The notion of re-integration is also advocated by Mishkin; lending credence to the presence of concatenated operations such as the decoder in **Figure 1** being followed by an interleaving operation. The hierarchical nature of the visual system consists of bi-directional information flow between the various levels (Van Essen and Gallant, 1994). Studies such as O'Reilly et al. (2013) and Lamme and Roelfsema (2000) have advocated the interaction of feedforward and feedback processing in delineating between the quick and detailed categorization of an object. The architecture of **Figure 2** considers feedforward connections as well as feedback projections that are guided by memory and the neural circuitry associated with attention. It is noteworthy that neuroanatomical evidence for cell structure influencing function in the visual system is provided in studies such as Elston et al. (2005) and Jacobs et al. (2001), and there are abundant discussions on the specialization of feedforward and feedback connections along the VVS (Rockland, 1997).

The CTOR formulation is fundamentally different from prior computational vision works such as Salinas and Abbott (1997) by considering the operation of the IT neural circuitry along with the functionality posited to be performed by the lower layers of the visual system. The notion that visual objects are represented by patterns of activity across populations of neurons has been advocated in discussions such as Zhang et al. (2011) and Lee and Mumford (2003). In accordance, the processing considered in CTOR can incorporate the representation of the neural activity via vectors that have dimensionality corresponding to the considered neural population. The encoding and decoding operations hypothesize that the neural activity has structure and is affected by the external environment and a subject's memory. Attention is suggested as having an impact on all operations of CTOR including the encoder, channel, decoder, interleaver, compression, and memory (**Figure 1**). The work of Lee and Mumford (2003) presented Bayesian inference as part of a graphical model for the viewing of an object by the early visual cortex. Their analysis makes mention of neural populations from the IT and V4, but is primarily focused on V1 and V2. The CTOR formulation will focus on the higher visual regions by presenting maximum a posteriori probability (MAP) inference within the context of the IT's role in object recognition. We shall use a binary alphabet to present the signals at the various stages of CTOR, however, the components should not be automatically associated with spikes. It is logical to inquire if the elements of the encoded and decoded CTOR signals are outputs of individual neurons, the result of a principal component analysis applied to output of populations of neurons, or perhaps the binary-thresholded outputs of neural circuits. The dimensionality of the signals in **Figure 1** can be specified to encompass all of the aforementioned scenarios. While such level of abstraction may be deemed unnecessary, it is productive for a new model to allow flexibility so that it can be fit to various data sets. As advocated in works such as DiCarlo and Cox (2007), CTOR encourages a

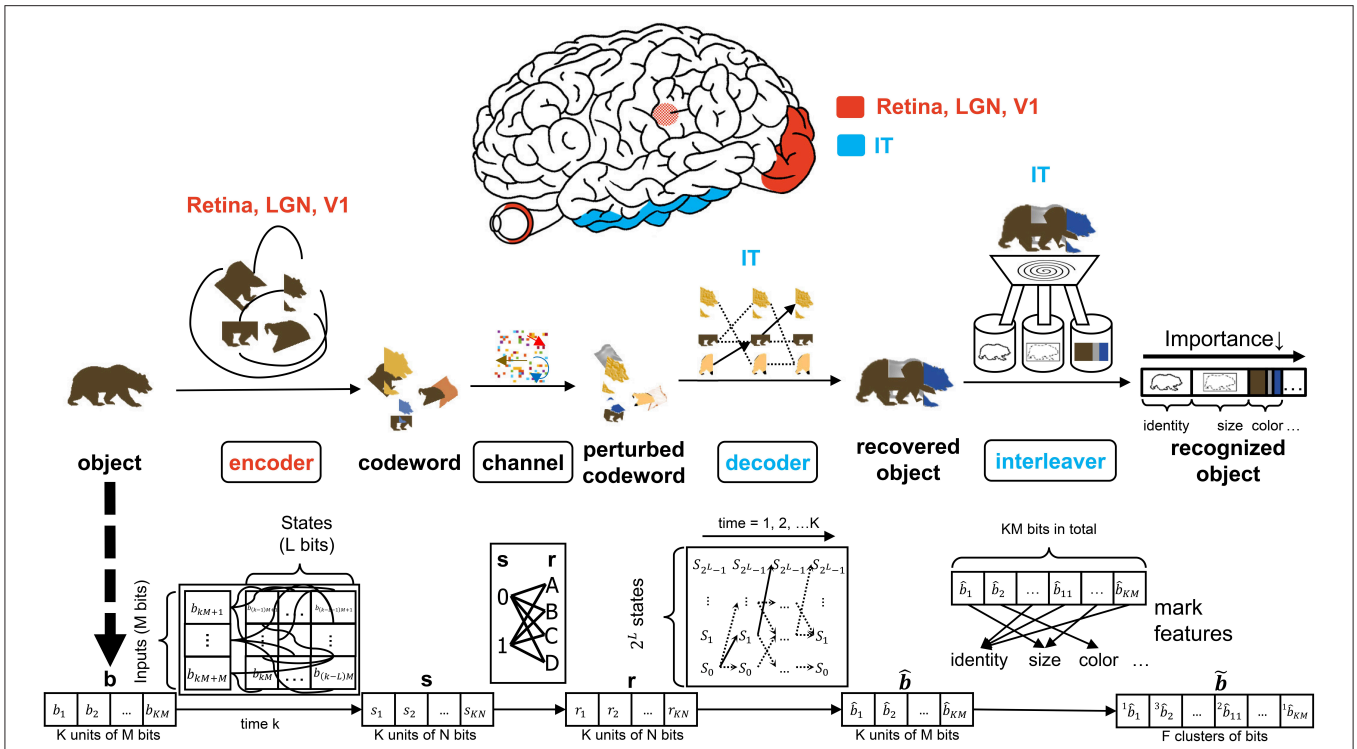

FIGURE 2 | The interaction of the input, encoder, channel, decoder, and interleaver in the CTOR model. The top row depicts the brain regions implicated for the encoding, decoding, and interleaving operations. The middle row provides an illustration of CTOR operations with the viewed object being a brown bear. The bottom row illustrates the progression of the dimensions of the neural signals that are subject to the CTOR operations.

shift in emphasis from single-unit spiking activity in favor of the processing performed by neural circuits.

3.1. Model Input

A seemingly fundamental facet of a model is the input. Object i will be denoted by a binary representation $\mathbf{b}^i$ that encompasses the object's attributes. The stimulus index $i = 1, 2, 3, \ldots$ will serve as the identity of the viewed object and the representation that the object should evoke at the IT for correct recognition. The representation **b** will be tangled by the retina, LGN, and V1 prior to being untangled by the IT. In a nominal object recognition trial the stimulus representation of a viewed object such as a brown bear (**Figure 2**) will be faithfully recovered by the IT and then compressed prior to being stored in memory. The tangling of the object identity prior to its progression along the VVS has been elegantly discussed in DiCarlo and Cox (2007) via the notion of an intertwining of object manifolds. CTOR provides a concrete means of representing such a tangling, namely the mapping of **b** to a codeword as will be discussed below.

3.2. Object Tangling via the Encoder

The early stages of the visual system will tangle the representation **b** that the viewed stimulus should evoke at the IT. The CTOR example illustrated in **Figure 2** considers the encoder as being stimulus-driven. A rate coding operation has been advocated as taking place in various visual areas (Van Essen and Gallant, 1994). The viewed object manifolds conveyed to area V1 by the retinal and LGN processing are nearly as tangled as the pixel representation (DiCarlo et al., 2012). This is largely attributed to the receptive fields in the aforementioned two populations being functionally akin to point-wise spatial filters (Olshausen and Field, 2005). Interestingly, as the retinal- and LGN-processed signals are processed by V1, the total dimensionality of the representation is increased approximately 30-fold (Stevens, 2001). However, the V1-processed signal is still considered highly tangled since its response is significantly inferior to human performance for real-world recognition problems (DiCarlo et al., 2012). Such biological characteristics are motivation for CTOR to postulate the encoder as being comprised of the retina, LGN, and V1 circuitry. Since the object representation is tangled by the encoder, it is debatable whether LGN—rather than V1—should be considered as the last stage of the encoder. This judgment is based on the V1 output still being highly tangled, and that the dimensionality increase that occurs following V1's processing of the LGN output is a trademark of the encoding operation. An example is shown in **Figure 2** where the representation **b** of a viewed object, i.e., a brown bear, is encoded into the stream **s** as the tangled version of the representation which should be evoked at the IT when viewing this object. Two parameters are crucial to the discussion. Assuming a binary alphabet, the integer M will denote the number of input bits processed by the encoding stage at a time. The integer K will denote the number of M-bit units allocated for representing the viewed object. Thus, the IT representation of a viewed object will consist of KM bits, and the

IT may have 2^{KM} distinct representations for a temporal window of duration K. The CTOR model considers a continuous stream of input bits being processed by the VVS. The continuous stream of information has been segmented into KM bits at different object boundaries. In effect, a larger K will correspond to an increase in the complexity of the viewed object.

Anatomically, the output of the encoder circuitry will be a length KN codeword $\mathbf{s}$ that comprises the neuronal response that the IT must decode (**Figure 2**). Although exceptionally large, we shall consider the number of possible representations as being finite. From a communication-theoretic perspective, encoding is an operation where a M-component input is mapped to a message consisting of $N \geq M$ components. From a reliability perspective it is advantageous to have $N \gg M$ because it behooves the decoder to have access to as many information-bearing signals as possible in its decision of which message to declare as the untangled representation. The ratio $M/N \leq 1$ is dubbed the code rate and the $N = M$ scenario is the somewhat anomalous case referred to as rateless coding because it provides no redundancy. An important parameter stems from the non-restrictive assumption that the encoder generates the codewords via a shift-register structure (Lin and Costello, 1983). The maximal memory order of the shift register will be designated by L. For ease in presentation, we shall assume a simple shift-register structure where the total memory is equal to the maximal memory (L). In communication theory, this quantity is referred to as the encoder constraint length and the same name will be used henceforth. It shall be assumed that only one bit is fed into the encoder at each time instant (i.e., $M = 1$)—this is also a non-restrictive assumption that is made for ease of presentation. At each time instant there will be 2^L possible states $\{S_0, S_1, \ldots, S_{2^L-1}\}$, that the encoder can take, and we shall denote the encoder state at time k by $S_{k,i} : i = 0, 1, \ldots, 2^L - 1$. The time index $k = 1, 2, 3, \ldots$ will be suppressed unless when necessary. When in state $S_{k,i}$ an encoder can produce only one of two possible codewords at time $k+1$. Similarly, a generated codeword could have only been preceded by two possible codewords at time $k - 1$. The length-N codeword $s_1, s_2, \ldots, s_N$ at time k will be denoted by the vector $\mathbf{s}_k$. The N components of the output codeword $\mathbf{s}_k$ will be dependent on $\mathbf{b}_k^i$ as well as the L prior inputs to the encoder: $\mathbf{b}_{k-1}^i, \mathbf{b}_{k-2}^i, \ldots, \mathbf{b}_{k-L}^i$. The codewords $\mathbf{s}_1, \mathbf{s}_2, \ldots, \mathbf{s}_K$ for duration K are concatenated as the encoder output vector $\mathbf{s}$ (**Figure 2**). In evaluating the CTOR operation and performance in the ensuing sections we shall consider $\mathbf{b}_k^i$ as being comprised of a small number of bits of synthetic data. The use of such synthetic data is a logical first step for introducing and motivating the model. In subsequent works an image stimuli can be considered by devising the sequences $\mathbf{b}_{k-1}^i, \mathbf{b}_{k-2}^i, \ldots, \mathbf{b}_{k-L}^i$ to be binary representations of the pixels in the object that is viewed by a subject. Expanding CTOR functionality to operate on input consisting of pixel intensities is a future consideration. As a summary, the biological implication of the encoder is relatively simple - a viewed object should elicit a representation at the IT; the representation is tangled via the encoding operation performed by the lower layers of the visual system. Assuming a binary alphabet, the neural signal corresponding to the encoded object will be represented by KN bits.

It is believed that the spiking of visual neurons is greater when attention is allocated to a stimulus than when attention has not been allocated to the same stimulus. The spiking rate of the retinal and V1 populations of neurons will be represented via the relation

$$s_i(t) = G_i(t) s_{i,\text{rest}}(t) \quad \text{for } i = 1, 2, \ldots N. \tag{1}$$

The above reflects attention, modeled by a positive and time-varying quantity $G_i(t)$ that has a multiplicative effect on the firing rate of the neurons. The process $s_{i,\text{rest}}(t)$ denotes the unmodulated firing rate of the ith V1 neuron. Works such as McAdams and Maunsell (1999) and Salinas and Abbott (1997) have provided evidence for $G_i(t)$ being a Gaussian function with parameters dependent on the attended location and the preferred attentional locus of the ith neuron. A codeword of length N denoted by a stream $s_1, s_2, \ldots, s_N$ will designate the activity of the V1 population of neurons, with the ith codeword component being "1" if the ith neuron has fired more than $\alpha_i > 0$ times during an interval (e.g., 50 ms as noted in DiCarlo et al., 2012), and "0" otherwise. In other words, it is conceivable to consider an assignment

$$s_i = \begin{cases} 1 & \text{if } \int_0^T s_i(t)dt > \alpha_i \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

as the rate coding rule for each of the N units over a time epoch of T seconds. More elaborate scenarios can be concocted where sub-populations of the lower-visual level neurons each form codewords that are multiplexed to form a larger codeword that is signaled to the IT. The CTOR model will specify the codeword $s_1, s_2, \ldots, s_N$ constructed by an encoder with a state-machine structure such as that shown in **Figure 3A**. The state diagram in **Figure 3B** illustrates the input-output dynamics of this encoder, where it is evident from **Figure 3A** that the encoder output will depend on the prior inputs to the encoder. Biologically, this implies that the output of the retina, LGN, and V1 stage is not a memoryless sequence, but rather follows a pattern that is modulated by various processes.

3.3. The Channel

Visual recognition is affected by dynamic perturbations that can have impeding effects such as obfuscating the object identity, delaying the recognition, and possibly leading to an erroneous identification of the object or its characteristics. The hindrances might stem from the properties of the viewed object (e.g., the novelty, or the object being obscured in the scene), or a subject's attentional state. Since neurons are inherently noisy, it is also possible for the encoder to be imperfect during its encoding of the stimulus. The CTOR model will subsume such impediments within a channel that separates the encoding and decoding operations (**Figure 2**). The output of the channel will be denoted by the vector $\mathbf{r}$ and shall constitute the input to the IT. For clarity in presentation, and in cadence with the communication and information theory literature, we separate the encoder and channel in **Figure 2** despite the fact that they are compound entities within a VVS. Although the channel separates the early

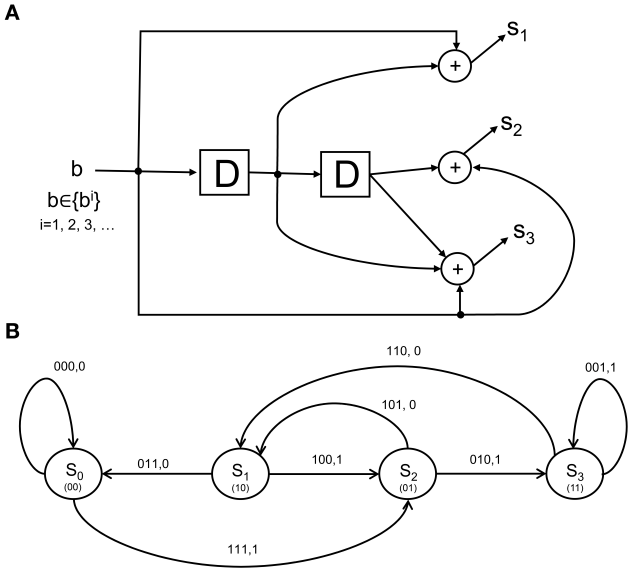

FIGURE 3 | The structure and state diagram of a simple encoder. **(A)** The so-called simple-encoder considered as performing the tangling operation at the retina and LGN in **Figure 1**. The "D" elements denote a unit delay inherent in the encoder's processing. The above is a definite simplification of encoding performed by the neural circuitry, but will serve to illustrate an instance of CTOR's encoding operation when the early visual stages are presented with an object. **(B)** A state diagram representation depicting the dynamics of the encoder in **(A)**. The grouping 100, 1 above the state transition $S_1 \rightarrow S_2$ indicates that when the neural circuits representing the retinal and V1 stages (i.e., the encoder) are in a state of S_1, a stimulus value of 1 would result in a transition to state S_2 as well as the encoder output of 100. The contents of the "D" elements in **(A)** representing the corresponding state are given between parentheses in **(B)**.

visual stages from the IT, the early stages' operations will resonate in shaping the stochastic perturbations that are modeled via the channel.

The instantiation of a channel plays a role in studying the robustness attribute that we have discussed for object recognition. A channel provides a source of dispersion (**Figure 2**) by distorting the codeword and will be represented via a conditional distribution $P[\mathbf{r}|\mathbf{s}]$ where $\mathbf{r}$ is a perturbed version of the signal and $\mathbf{s}$ is the encoder output. The channel may perturb the encoder output in either a continuous or discrete (i.e., quantized) fashion, accordingly, $P[\mathbf{r}|\mathbf{s}]$ will be represented either by a probability density function (pdf) or a probability mass function (pmf), respectively. The simplest linear, continuous channel consists of a noise process $\mathbf{n}$ being added to the encoder output via

$$\mathbf{r} = \mathbf{s} + \mathbf{n}. \tag{3}$$

A prevalent channel quality metric (CQM) for a continuous channel is the SNR. For (3) the SNR of neural signals conveyed to the IT will be expressed as

$$\mathrm{SNR}_i = \frac{\max(s_i) - \min(s_i)}{E[n_i^2]} = \frac{1}{E[n_i^2]} \quad \text{for } i = 1, 2, \ldots, N \tag{4}$$

with the denominator representing the neuronal noise power. The SNR of single neurons has been considered in numerous studies. In the spirit of works such as Mar et al. (1999), we consider an aggregate, population-wide CQM for the collective effect of the units comprising the retina, LGN, V1, V2, and V4. An insightful CQM for a discrete channel will quantify the uncertainty in the probabilistic mapping of the channel inputs to the channel outputs. The conditional entropy

$$H(r_i|s_i = n) = -\sum_{m=1}^{|r|} P[r_i = m|s_i = n] \log\left(P[r_i = m|s_i = n]\right) \tag{5}$$
$$\text{for } n = 0, 1 \quad \text{and} \quad i = 1, 2, \ldots, N$$

is viewed as the equivocation between a discrete channel's input and output, with $|r|$ denoting the cardinality of the set of possible channel outputs. From a biological perspective it is sensible to assume that over a short time-scale associated with a task, a continuous channel will maintain a probability distribution, but the parameters that characterize the distribution (e.g., mean and variance) will vary. Similarly, for a discrete channel it would be expected that during the viewing of an object the components of $P[\mathbf{r}|\mathbf{s}]$ change but the values $\{H(r_i|s_i)\}$ do not drastically vary. Over longer time-scales that span the viewing of different scenes it is expected that the channel's distribution will vary due to different stimuli and changes in attention.

3.4. The Decoder

There is evidence that in the visual cortex, neurons such as pyramidal cells become increasingly large, more branched, and more spinous as one progresses along the VVS (Elston, 2002). From the perspective of information transmission, the identity of a viewed object propagates along the VVS until reaching the IT. Works such as Karklin and Lewicki (2009) have suggested that sensory signals from early visual areas convey information that allows the higher visual areas to construct more complex representations of the sensory input. With CTOR, it is the objective of the decoder to determine the object identity and classify its characteristics. In effect, the decoded message will represent the object that the IT has identified from the representation propagated to the IT by the lower visual stages. After K time instances the sequence of vectors $\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_K$ will be available to the decoder with $\mathbf{r}_k$ representing a length-N perturbed codeword that is to be untangled into a length-M message. Accordingly, the decoder will continuously process the channel output at every time instant, with its output being a length-KM binary vector denoted by $\widehat{\mathbf{b}}$. The selectivity attribute discussed in section 2 is accounted for by the fact that the objective of decoding is discriminating between different patterns. A good decoder operating over a channel that is not inordinately dispersive will be capable of discriminating among various object representations with high likelihood. In effect, at each discrete time instant the decoder transforms a N-length sequence that may take on a number of possibilities to a M-length binary sequence. A decoding operation is conceived by considering various metrics, for example, a MAP decoder

would select $\widehat{\mathbf{b}} = \mathbf{b}^i$ for the $\mathbf{b}^i$ that maximizes the probability $P[\mathbf{b}^i|\mathbf{r}]$ where $i \in \{1, 2, \ldots, 2^{KM}\}$. Dynamic programming is often used to solve large-scale inference problems when it is desired to recover a sequence that has the highest possibility of having occurred. The Viterbi algorithm provides the most probable sequence of states when the environment is described by a hidden Markov model (HMM) (Eddy, 2004) with the similarities between the principle and dynamic programming discussed in the seminal work of Forney (1973). A description of the Viterbi algorithm is provided in **Appendix A** and the terminology there will be incorporated henceforth. The CTOR proposal for the VVS proficiency at object recognition lies in the IT implementing the untangling notion via a MAP decoding algorithm in order to infer the object identity and attributes. The untangling notion can be equated to seeking the most likely path in a state transition diagram with 2^L states at time k. The length of prospective transitions between two states $\mathbf{b}_{k+1}$ and $\mathbf{b}_k$ at time k is quantified via

$$\lambda(\mathbf{b}_{k+1}, \mathbf{b}_k) = -\ln(P[\mathbf{b}_{k+1}|\mathbf{b}_k]) - \ln(P[\mathbf{r}_k|\mathbf{b}_{k+1}, \mathbf{b}_k]) \quad (6)$$

where $P[\mathbf{b}_{k+1}|\mathbf{b}_k]$ is the a priori probability of state $\mathbf{b}_{k+1}$ given the observance of state $\mathbf{b}_k$, and the transition probability $P[\mathbf{r}_k|\mathbf{b}_{k+1}, \mathbf{b}_k]$ denotes the probability between a given pair of successive states and the sequence $\mathbf{r}_k$. The process is illustrated in the decoder portion of the example in **Figure 2** with the IT performing dynamic sequence estimation of the tangled representation.

An appeasing feature of the CTOR proposal is that the invariance, robustness, and selectivity attributes discussed in section 2 may be considered in unison. This is because when decoding $\mathbf{r}$ the MAP sequence estimation technique attempts to recover the correct message, or one that is as "close" as possible to the correct message despite disparity in certain attributes. The disparity is noted by the bit streams disagreeing at various positions, and the degree of closeness is quantified by the Hamming distance between the sequence decoded by the IT and the representation that the viewed object should have evoked at the IT. We define the deviation by

$$d(\widehat{\mathbf{b}}, \mathbf{b}^i) \triangleq ||\widehat{\mathbf{b}} \oplus \mathbf{b}^i||_0 \quad (7)$$

where $\oplus$ denotes the component-wise XOR operation and $||\mathbf{x}||_0$ denotes the number of non-zero elements in the vector $\mathbf{x}$. Invariance has been considered since correct decoding and object recognition are possible despite transformations induced to the sequence $\mathbf{b}^i$ (via the channel) prior to its entering the decoder. Works such as (Usher and Niebur, 1996) have advocated the IT exhibiting a larger overlap in its representations of similar objects than in its representation of dissimilar objects. The overlap of the similar objects is conveniently modeled in CTOR by such objects having decoded sequences that are relatively close in Hamming distance. Conversely, the decoding of dissimilar objects will result in sequences that have a larger discrepancy in Hamming distance. For instance, the representation of an object such as $\mathbf{b}^1$ =brown bear is expected to be closer in Hamming distance to $\mathbf{b}^2$ =baby elephant than to $\mathbf{b}^3$ =green hat. Inspection of a simple, synthetic example such as

$$\begin{aligned} \text{brown bear}: \mathbf{b}^1 &= 110010111001001 \\ \text{baby elephant}: \mathbf{b}^2 &= 110010111001100 \\ \text{green hat}: \mathbf{b}^3 &= 101100101100101 \end{aligned} \quad (8)$$

indicates that $d(\mathbf{b}^1, \mathbf{b}^2) < d(\mathbf{b}^1, \mathbf{b}^3)$, and $d(\mathbf{b}^1, \mathbf{b}^2) < d(\mathbf{b}^2, \mathbf{b}^3)$. In other words, the first two decoded sequences are closer to each other than either sequence is to the third. We note that the decoding accuracy is dependent not only on the decoder, but also the encoder and the channel properties. For instance, the robustness attribute can not be realized by the encoder and decoder alone because a channel with a very poor CQM would perturb the encoded representation to a degree that the decoder would be incapable of correctly untangling the object's identity.

Visual processing works such as Reynolds and Chelazzi (2004) and Usher and Niebur (1996) have discussed the so-called competition among the neural representation of objects along the VVS. The competition occurs between a target object and distractors that are concomitantly present during the viewing. We posit that there is also competition among the objects stored in memory that are vying to be declared the viewed object. Such competition is incorporated in CTOR as the closeness among the decoded codewords. For example, in (8) there will be more competition among the representations $\mathbf{b}^1$ and $\mathbf{b}^2$ than among $\mathbf{b}^2$ and $\mathbf{b}^3$. **Figure 4** depicts this notion with the closest messages competing within a decision space to be the representation associated with the viewed object. Models such as Usher and Niebur (1996) consider a suppression of the neural activity for a competing stimuli following a decision as to which object is present. With CTOR, the suppression of competitive stimuli occurs by the decoding operation discarding all prospective messages except for the selected $\widehat{\mathbf{b}}$. A comparative mechanism is inherent during the decoding operation since the codeword that is closest to the represented formulation is selected by the decoder as the decoded message. It was reported in Rust and DiCarlo (2010) that performance on visual discrimination tasks depend considerably on the number of neurons included in the analysis and the number of images included in the stimulus set. The decoding framework incorporates analogues for these two dimensions via the codeword length N, and the cardinality of the set of possible representations (i.e., $|\{\mathbf{b}^i\}| = 2^{KM}$), respectively.

3.5. The Interleaver

As the IT processes the representation from V4, the neural response is reformatted to be more selective for feature conjunctions (Rust and DiCarlo, 2010). In CTOR such processing is modeled via an interleaving operation. Interleavers are discussed in communication theoretic works such as Ramsey (1970), and have found application in computer science as well (Andrews et al., 1997). The biological motivation behind the interleaver lies in the necessity for the information output by the decoder to be parsed into a set representing the attributes and also the importance of the attributes for recognition. The interleaver shall arrange the decoder output into a sequence where the ordering has neurological significance for the efferent circuitry

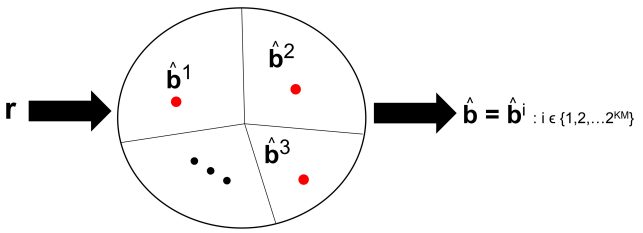

FIGURE 4 | An example of the decision regions at the IT's decoder at a specific time instant. Suppose $KM = 15$ with $2^{KM} = 32,768$ possible representations that the decoder can declare as having been the viewed object. The circle denotes the possible space of received $\mathbf{r}$ sequences with $\mathbf{r} \in \mathbb{R}^{KN}$. The red dots denote the representation of such objects in a decision space. The closed region that surrounds a dot represents the region where the VVS would declare $\widehat{\mathbf{b}} = \mathbf{b}^i$ if $\mathbf{r}$ were to lie in that region. The sequences $\mathbf{b}^1$, $\mathbf{b}^2$, and $\mathbf{b}^3$ are three prototypical objects.

(Brady et al., 2009; Nassar et al., 2018). Since there is a need for considering the notion of feature grouping within the visual system (Olshausen, 2013), the interleaving operation in CTOR is a functional equivalent to the IT deciding the order of importance given to the features by consciousness and attention. In the example of the viewed object being a brown bear, the identity, size, color, and shape are ranked according to their importance. More important features such as identity appear before the less important features such as color (**Figure 2**). **Figure 5** provides an example of the interleaving with the decoded message $\widehat{\mathbf{b}}$ being partitioned into smaller groups that correspond to the object's features. The ordering of the bits that comprise the interleaver output via the vector $\tilde{\mathbf{b}}$ signify the order-of-importance of the features. This parsing and segmentation into components has been motivated by Biederman (1987). In **Figure 5**, the KM bits in the decoded message have been partitioned into F features with the variables $M_1, M_2, \ldots, M_F$ denoting the number of bits attributed to each feature. There is an obvious constraint that $\sum_{i=1}^{F} M_i = KM$. The F features that we allude to correspond to the stimulus dimensions introduced by the feature-integration theory of attention (Treisman and Gelade, 1980) that has been further elaborated upon in works such as Van Essen and Gallant (1994). Since the plasticity of the IT is responsible for refining the basic vocabulary of features (Serre et al., 2007; Rust and DiCarlo, 2010) it is expected that the interleaver is vastly distinct among different brains. It is also logical to posit that the interleaving operation is a highly dynamic process within a subject. With respect to neurophysiology, works such as Poggio and Kreiman (2013) and Meyers et al. (2008) have discussed the prefrontal cortex (PFC) guiding the IT (via a top-down signal) in the activation of subgroups of neurons to specific object features. It has been shown that PFC neurons also exhibit an increase in dendritic and spine complexity that is seen in the latter stages of visual cortical processing (Jacobs et al., 2001; Jacobs and Scheibel, 2002; Elston et al., 2011), and that the complexity is amenable to the progressive increase in sophistication of the computational operations. This was a motivation for the PFC-IT interaction considered in **Figure 1** as the mechanism driving the interleaving operation. The interleaving operation constitutes a computation that is performed by populations of neurons acting collectively. Thus, the M-to-M component mapping of $\widehat{\mathbf{b}} \rightarrow \tilde{\mathbf{b}}$ entails the coordinated firing among a population of neurons rather than the autonomous firing of neurons that may occur in populations at the lower visual layers. The output of the interleaver is comprised of F clusters with each cluster distinguishing a feature of the viewed object. In effect, the sequence $\tilde{\mathbf{b}}$ is the information that the VVS has extracted (i.e., untangled) from the scene during object recognition via the decoding and interleaving operations.

3.6. Declaration of the Object Category

Despite the advancements in the study of primate vision, it has not been ascertained at what specific juncture in the VVS a viewed object can be said to have been recognized. The authors in (Neri and Heeger, 2002) advocate the presence of two stages in the VVS with the first performing object detection $\sim$ 100ms prior to the second stage performing identification of the object's features. **Figure 6** is a more detailed depiction of the operations associated with declaring the object category that was alluded to in **Figure 1**. In **Figure 6**, a classifier deciphers the object category by processing the decoded output. The CTOR model considers the progression of the decoder output into a classifier and an interleaver. Such parallel processing reflects the VVS's capability to classify the object category concomitant to discerning its features. A computationally simple model for object categorization is the inner product of the decoder output with a weight vector $\mathbf{w}$ via

$$f(\widehat{\mathbf{b}}) = \mathbf{w}^T \widehat{\mathbf{b}}. \tag{9}$$

This is essentially the linear classifier readout advocated in (Rust and DiCarlo, 2010) although it is expected that the dimensionality $\dim(\mathbf{w}) = \dim(\widehat{\mathbf{b}}) = KM$ for CTOR will be significantly larger than what has been previously considered. It is important to note that the output of $f(\cdot)$ is not sensitive to the order of the elements in the column vector $\widehat{\mathbf{b}}$ since $\mathbf{w}$ can be adjusted accordingly. Works such as Rust and DiCarlo (2010) and Pagan et al. (2013) have determined a realization of the vector $\mathbf{w}$ for every presented image in a set. While the selection of a classification technique for determining $\mathbf{w}$ is not the objective of this work, we remark on a crucial point. The assignment $\mathbf{w}^T = [1, 1, \ldots, 1]$ would lead to the discernment of the object identity being solely a function of the Hamming weight of the decoded message. The above consideration for $f(\cdot)$ also instantiates CTOR as exhibiting the invariance attribute since the components in $\widehat{\mathbf{b}}$ can be re-arranged without a change in a declaration of the identified object's category. The distinction between an object's category and identity should be apparent. In the presented example, "brown bear" is the identity of the input object while "bear" is a declared category.

4. THE NECESSITY OF ATTENTION, COMPRESSION, AND MEMORY

Seminal works such as Biederman (1987) and Treisman and Gelade (1980) have motivated the importance of

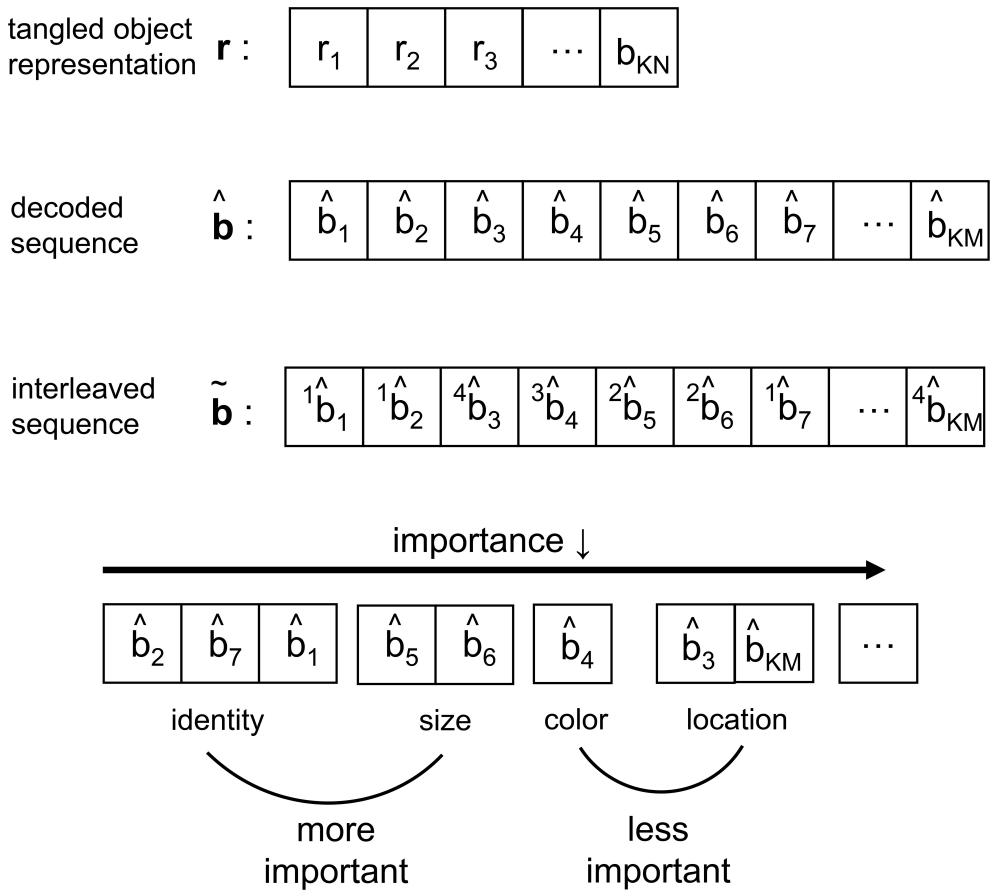

FIGURE 5 | Depiction of the signal structure as it evolves along the VVS model presented in **Figure 1**. The *KN*-component vector $\mathbf{r}$ with $r_i \in \mathbb{R}$ has been decoded into a message $\widehat{\mathbf{b}}$ with *KM* components, where $\widehat{b}_i \in \{0, 1\}$ and $N \gg M$. At the output of the interleaver the *KM* components are partitioned into a message $\widetilde{\mathbf{b}}$ with *F* features—or feature dimensions as denoted in works such as Kanwisher and Wojciulik (2000)—of variable bit lengths. In the above example, $F = 4$ feature groupings are presented.

considering memory, attention, and object recognition within a unified model. The patent biological interplay between the aforementioned processes leads one to believe that an incomplete analysis would result by not considering such processes as interacting either via feedforward or feedback connections. The authors of Usher and Niebur (1996) have also advocated the concurrent consideration of attention and memory with the neural activity associated with the early visual stages. The model presented in the aforementioned work considers the necessity of a top-down feedback projection when a subject is searching for an expected target in a scene. This section will discuss how CTOR accounts for the interaction of attention and memory to provide a unified model for object recognition.

4.1. Attention as a Top-Down Modulatory Signal

The incorporation of attention as the modulator of the neural processes associated with object recognition is crucial. A review of the neural circuitry in the visual cortex that is actively modulated by attentional feedback has been presented in Reynolds and Chelazzi (2004). From analysis in monkeys it is natural to suggest that the attention module in **Figures 1**, **6** would contain the lateral intraparietal area (LIP) and frontal eye fields (FEF). CTOR posits attention as modulating components such as the encoder, channel, decoder, and compression via a top-down regulatory mechanism (**Figure 6**). Attention affects the encoder via the multiplicative factors $\{G_i(t)\}$ in (1) that drive the spiking rates of the retinal and V1 neurons. This reflects a role associated with the projection from the attentional circuitry to the encoder. A subject's attentional state will also influence the channel by affecting the conditional distribution $P[\mathbf{r}|\mathbf{s}]$. In the case of a continuous channel the effect may be seen on the SNR values $\{\text{SNR}_i\}$ which are a function of a subject's vigilance as well as the inherent neural noise along the VVS. It is sensible to assume that the SNR values increase with greater levels of attention. In the case of a discrete channel a similar modulation is expected with the conditional entropy values being affected by attention. The decoder is immanently influenced by a subject's attentional state through the vector $\mathbf{r}$ that the decoder must process during each epoch. This is seen by noticing that the transition metrics, path metrics, and survivor paths computed at the decoder en-route to declaring a message $\widehat{\mathbf{b}}$ are determined by the channel and the

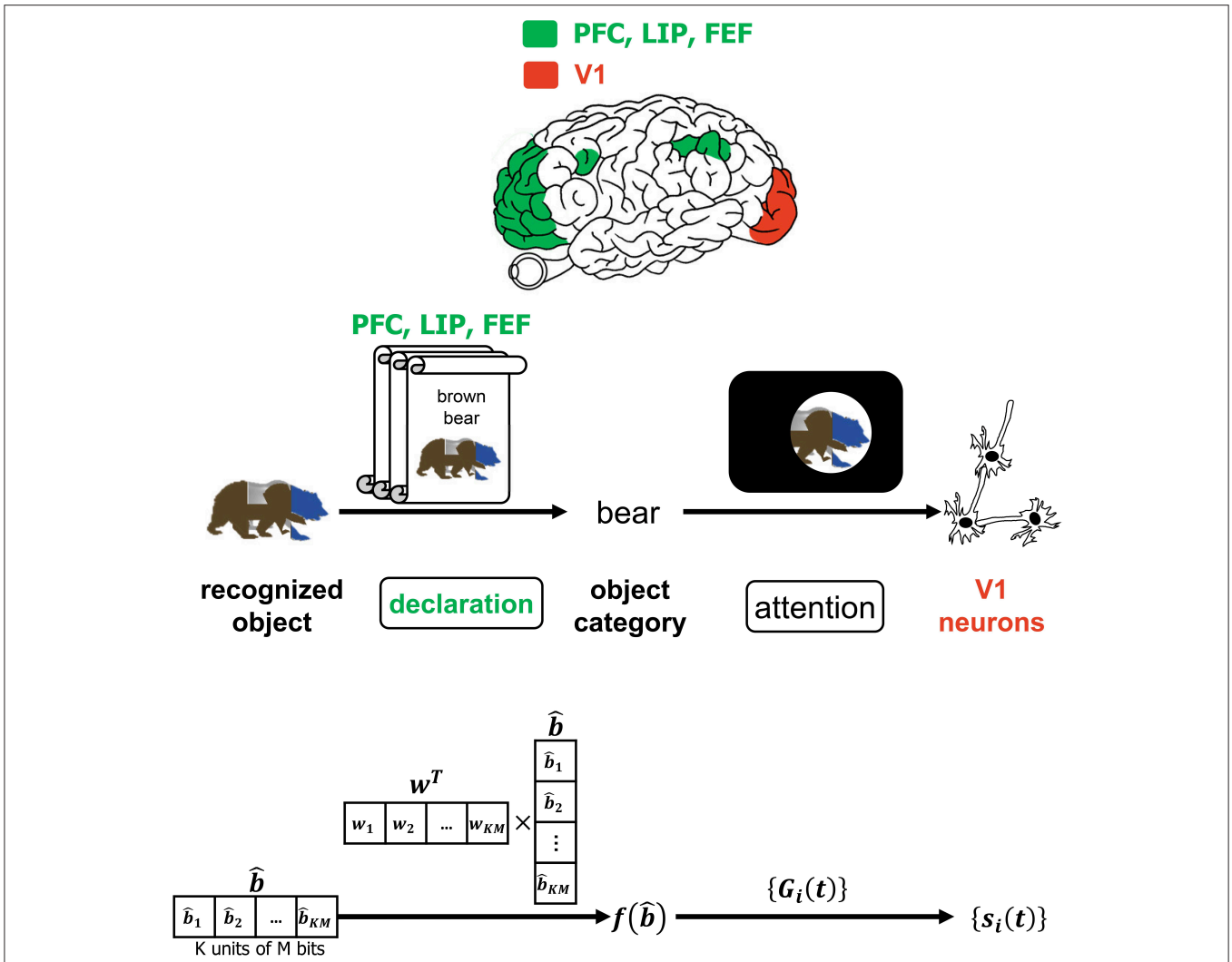

FIGURE 6 | The interaction of attention and the declaration of object category. The top row depicts PFC, LIP, and FEF regions as they govern attention which is affected by the declared object category. The middle and bottom row provide an example of the CTOR operations and the dimensionality of the associated signals during the declaration of a viewed object's category. The scrolls denote different categories that viewed objects may be classified into prior to affecting attention. Attending to a categorized object is believed to modulate the firing rates of V1 neurons as discussed in Equation (1) and depicted via the projection (a) of **Figure 1**.

encoder. We have mentioned that with CTOR the number of bits attributed to each feature by the interleaver is a dynamic process modulated by attention. Works such as Cukura et al. (2013) and Huth et al. (2012) provide experiments that illustrate attention driving the degree of compression applied to what constitutes the F interleaved features in CTOR.

Attention also modulates the goals of object recognition. Consider the general scenario of a subject knowing that he/she must espy a scene before making a critical decision on an object in the scene. A nominal example of this is a driver checking a blind-spot immediately before changing lanes on a highway. The brain will have a snapshot view of the scene and, due to the heightened level of attention necessary for this task, perform object recognition much more quickly than during typical visual tasks. In such a pedestrian example the IT's decoder would recognize a car but the brain would allocate significantly more importance to the location and proximity of the car than its color or luminance. Brain imaging neuropsychological studies conducted in works such as Kanwisher and Wojciulik (2000) and Turk-Browne et al. (2013) have explored attentional modulation of visual encoding, memory formation, and the brain's capability to prioritize the sensory information that is most relevant for a task. It is necessary that a computational vision model also incorporate such notions. The CTOR model currently considers attentional selection by the increased firing of V1 neurons, while not accounting for the more sophisticated scenario of overlapping objects as described in works such as Baldauf and Desimone (2014). The incorporation of the biological functions associated with the capability of the VVS to separate attended and unattended objects is an avenue for the advancement of CTOR as its constituent portions are expanded upon. For instance, it can not be claimed that the entire

VVS would consist of a single realization of **Figure 1**. Rather, it is more likely that there would be multitudes of such an architecture acting in parallel prior to a convergence. In Kersten et al. (2004), the parallel implementation of Bayesian models is mentioned and the authors advocate decomposing a scene or concurrently viewed objects into m features. **Figures 1, 2**, which have been a thrust of this work, will need to be cascaded into parallel streams to form a more comprehensive scheme that accounts for the case of overlapping objects competing for attention.

4.2. A Compressed Representation of Recognized Objects

It is infeasible to conceive that the brain will commit every feature of each identified object to memory. The CTOR model allows for the incorporation of a compressive operation to proceed the interleaving process. The degree of compression will be a dynamic process modulated by attention and will shorten the representation of each object based on its most important features. This may be achieved by prioritizing the features that have been highly ranked by the interleaver while summarizing or even discarding the less-important features of a viewed object. **Figure 7** depicts compression taking place in the hippocampus and amygdala where all objects' features such as identity, size, and color are subject to compression prior to being committed to memory. In CTOR, this process is achieved by combining multiple occurrences of similar objects into a single representation in memory as a sequence $\widetilde{\mathbf{b}}^c$. The memory circuitry is also driven by attention and will be presumed to have a fundamental role of providing the IT with the top-down a priori probabilities necessary for the IT to perform inference. The hippocampus's storage and rapid consolidation of object representations has been considered for decades with works such as O'Reilly and McClelland (1994) suggesting that the hippocampus is constructed to perform such a function. From a reverse engineering perspective, it is highly efficient that an object viewed at the highest frequency be allocated the smallest number of bits in memory. Different from the compression technique in CTOR, this alternative strategy would minimize storage and be akin to compression in the sense of Huffman coding or more recent proposals that suggest the hippocampus is performing even more sophisticated compression techniques (Petrantonakis and Poirazi, 2014).

4.3. The Consideration of Memory

For an object to be accurately recognized, a representation of the object must have been previously compressed and stored at an acceptable fidelity. There has been substantial evidence that memory-associated brain regions such as the hippocampus and amygdala are crucial for the neural processing underlying object recognition. Classical studies have referred to the area TE as containing "neural traces" associated with previously viewed stimuli (Mishkin et al., 1983). Such traces serve as stored representations against which subsequently viewed stimuli are compared. CTOR subsumes the comparisons into the decoding operation performed at the IT. The formation and storage of the traces are deemed as occurring at the hippocampus-amygdala circuitry that **Figure 7** portrays communicating with the IT via feedforward and feedback connections. This is also illustrated in **Figure 1** as the feedback connection from memory to decoder. The prevalence of the signaling from the memory circuitry to the IT and the neural circuits governing attention has also been justified in works such as Chelazzi et al. (1998) where the authors considered feedback provided by memory as a top-down signal for modulating the attention allocated to the object's attributes. The CTOR model considers two interactions between the memory and decoding circuitry that will propagate the transition probabilities $\{P[\mathbf{r}_k|\mathbf{b}_{k+1}, \mathbf{b}_k]\}$ and the a-priori probabilities $\{P[\mathbf{b}_{k+1}|\mathbf{b}_k]\}$ between the two entities. Firstly, the feedforward signal from the decoder that enters the hippocampus reflects memory formation following the recognition of an object and its associated features. Conversely, when a subject is processing a scene and attempting to recognize an object within the scene, the brain vests attention and draws upon stored memories to perform the recognition. It is expected that memory provides the a priori probabilities $\{P[\mathbf{b}_{k+1}|\mathbf{b}_k]\}$ to the decoder during decoding (**Figure 7**). Works such as Olshausen (2013) have discussed the importance of feedback in the visual system as a potential means of communicating, via a top-down signal, the a priori probabilities that the brain uses when performing inference in stimulus space. The feedback connection considered by CTOR from memory is a means of enabling the decoder portion of the IT to operate in Bayesian fashion by providing the decoder with updated a priori probabilities. Secondly, object recognition can not occur without the IT having access to an itemized list of objects and attributes. We posit that such a dictionary exists and is continuously updated via the feedforward and feedback signaling discussed herein. The components of the dictionary are compressed versions of the previously viewed representations. The work of Mishkin has provided analytical motivation and experimental results on the notion of recognition memory. The interaction of the PFC in guiding working memory and visual search has also been considered in a model presented in Usher and Niebur (1996) that was further advocated in Poggio and Kreiman (2013). For a decoder at the IT to implement the Viterbi algorithm it must have knowledge of the encoder and the channel statistics. We can explain this as synaptic plasticity that occurs between neural populations of various brain regions that share connection. That is how an upstream population in IT could learn about some properties of the V4 and V1 neurons that constitutes the transition probabilities. In other words, the transition probabilities $\{P[\mathbf{r}_k|\mathbf{b}_{k+1}, \mathbf{b}_k]\}$ must be conveyed to the decoder from the memory circuitry. The hippocampus and amygdala will continuously update their account of the transition probabilities by repeated interaction with the decoder in the IT. It is conceivable that during a developmental or training phase—that a subject may be agnostic to—the memory circuitry extensively communicates with the IT in order to update its estimates of the transition probabilities. Works such as Van Essen and Gallant (1994) and Miyashita (1993) have also cited IT neuron responses in primates as being markedly changed through repeated exposure to a limited set of stimuli. Accordingly, with

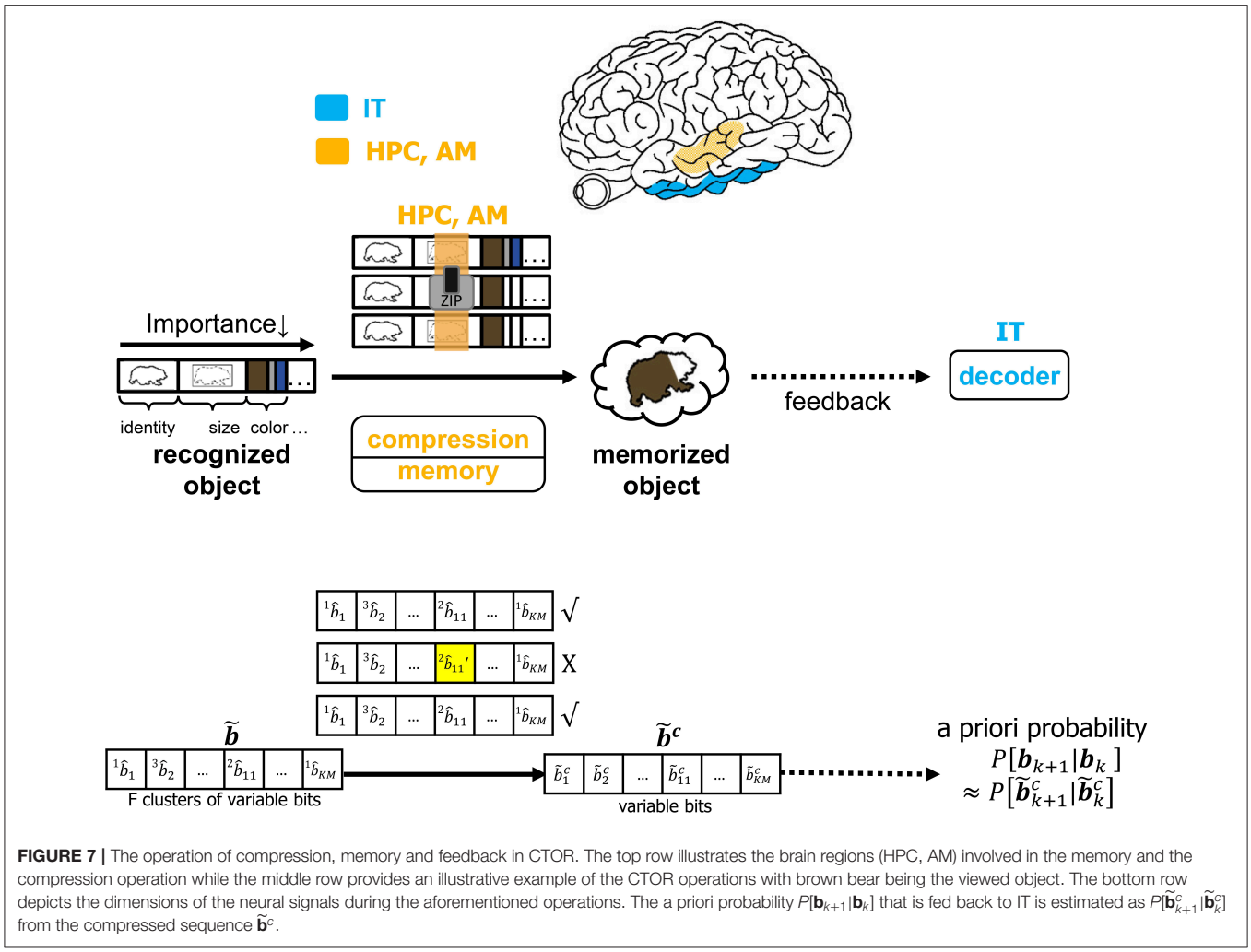

FIGURE 7 | The operation of compression, memory and feedback in CTOR. The top row illustrates the brain regions (HPC, AM) involved in the memory and the compression operation while the middle row provides an illustrative example of the CTOR operations with brown bear being the viewed object. The bottom row depicts the dimensions of the neural signals during the aforementioned operations. The a priori probability $P[\mathbf{b}_{k+1}|\mathbf{b}_k]$ that is fed back to IT is estimated as $P[\tilde{\mathbf{b}}^c_{k+1}|\tilde{\mathbf{b}}^c_k]$ from the compressed sequence $\tilde{\mathbf{b}}^c$.

CTOR the IT-hippocampus interaction will be an iterative process—if the decoded output is such that $\widehat{\mathbf{b}} \approx \mathbf{b}$, then the VVS may maintain the transition probabilities as legitimate estimates for ensuing epochs until $\widehat{\mathbf{b}}$ deviates sufficiently from $\mathbf{b}$ (Kersten et al., 2004). In statistical communication theory the above procedure is referred to as the decoder learning the channel and is implemented via means such as the Baum–Welch algorithm (Hastie et al., 2009).

5. THE OPERATION OF CTOR

It is insightful to consider an example of CTOR operation that commences with the tangling of the stimulus representation and concludes with a decoding, interleaving, and commitment to memory of the untangled object identity. We consider an example where at each time instant the early visual stages will tangle $M = 1$ bits of the object identity into a $N = 3$ bit sequence. We also consider $K = 4$ and thus the identity of the viewed objects will lie in a space with a cardinality of 16. As part of this toy example, suppose that the viewed object has the representation $\mathbf{b} = 1100$ at the IT. Of course this constitutes a highly synthetic stimulus signal with M, N, and K values small enough for the analysis to be tractable while still elucidating the computations advocated by CTOR. We caution that although computational intractability is avoided in this example, it is by no means reflective of the VVS avoiding such intractabilities—obviously the VVS's prospective implementation of the encoding and decoding would encompass significantly larger K and N values. The interleaving, categorization, and compression operations will also be instantiated in the toy example of this section.

5.1. Object Tangling and Manifestation of the Channel

From a biologically-inspired perspective, an encoder of rate 1/3 signifies that every bit from the representation that the viewed object should evoke at the IT has been tangled by the retina, LGN, and V1 into three bits. We shall consider the encoder in **Figure 3A** since it has already been discussed in section 3.2. Communication theorists would describe this encoder via a so-called algebraic generator sequence $G(D) =$

$[1 + D, 1 + D^2, 1 + D + D^2]$ and recognize that the encoder has a maximal memory of $L = 2$ that allows the encoder to take $2^L = 4$ possible states at each time instant. The encoder in **Figure 3A** has been extensively discussed in Lin and Costello (1983) and will be dubbed "simple-encoder" for the remainder of the paper. For clarity, the four states shall be referenced via $\{S_0, S_1, S_2, S_3\}$ as shown in the state diagram in **Figure 3B**. The number of possible transitions in the encoder state diagram is $2^N = 8$. For instance, $S_0 \rightarrow S_2 : 111, 1$ and $S_0 \rightarrow S_0 : 000, 0$ denote two of the transitions. The $N = 3$ bits written above each transition is the encoder output that is generated due to the combination of that transition and the $M = 1$ bit input to the encoder (e.g., "1" and "0" for the $S_0 \rightarrow S_2$ and $S_0 \rightarrow S_0$ transitions, respectively). The encoding operation has structure that is modeled via a state-machine - this reflects that the tangled signals converging at the IT via afferent projections are not completely random patterns. For instance, regardless of the nature of the viewed object, it is obvious that the encoder in **Figure 3** would prohibit an encoded sequence of 111 to be followed by an encoded sequence of 001. It can be verified that the considered sequence $\mathbf{b} = 1100$ would be encoded into $\mathbf{s} = 111\ 010\ 110\ 011$ according to the state diagram of simple-encoder.

During the encoding or tangling operation the neural representation that a viewed object should evoke at the IT is perturbed by a channel that encompasses the visual impairments inherent to the scene as well as neural noise inherent to the VVS. We model this via each element in $\mathbf{s}$ being stochastically transformed into one of four values denoted by A, B, C, and D. The four values reflect different intervals for the neural activity produced by the circuitry that projects to the IT. Consider the discrete memoryless channel quantified by the following conditional probabilities:

$$\begin{aligned} P[r_i = A|s_i = 0] &= 0.4 \\ P[r_i = B|s_i = 0] &= 0.3 \\ P[r_i = C|s_i = 0] &= 0.2 \\ P[r_i = D|s_i = 0] &= 0.1 \\ P[r_i = A|s_i = 1] &= 0.1 \\ P[r_i = B|s_i = 1] &= 0.2 \\ P[r_i = C|s_i = 1] &= 0.3 \\ P[r_i = D|s_i = 1] &= 0.4. \end{aligned} \tag{10}$$

The above is one of the channels considered in Lin and Costello (1983) and it is used in this illustrative example for its relative simplicity—it is easy to verify that $H(r_i|s_i = 0) = H(r_i|s_i = 1) = 1.846$ bits. The channel output is presumed to be the sequence

$$\mathbf{r} = (DCA,\ DDB,\ DDA,\ DDD). \tag{11}$$

The decoder considered in the following section will process the sequence $\mathbf{r}$ via the Viterbi algorithm in order to attain a MAP estimate of the viewed object's representation. The accuracy of the recovery process will quantify the fidelity at which the object identity $\mathbf{b}$ has been untangled at the IT.

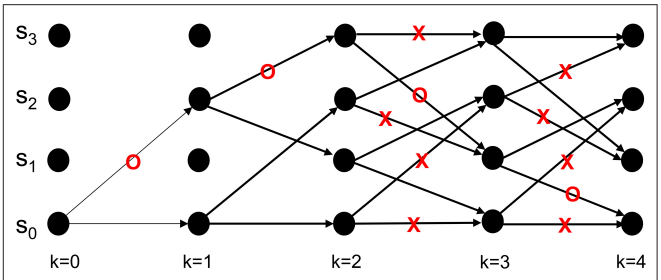

FIGURE 8 | An example from Lin and Costello (1983) to illustrate the dynamics of Viterbi decoding. The path marked with "O" denotes the final survivor path selected based on the smallest path length which corresponds to the MAP estimate of the sequence that the encoder desired to convey to the decoder. The transitions marked with "X" denote non-survivor transitions, while the unlabeled transitions denote survivors that were not part of the final survivor path.

5.2. Decoding Dynamics

Neural activity at the IT is believed to correspond to the untangled identity of the object that has been communicated to the IT (in tangled form) by the lower layers of the VVS. Cortical computation presentations such as Rao and Ballard (1999), Olshausen (2013), Lee and Mumford (2003), and Kersten et al. (2004) have advocated a hierarchical Bayesian model with top-down and bottom-up information flow. Such dynamics are at the heart of proposed decoding operation for the IT. The decoder uses bottom-up information from the encoder in conjunction with top-down information from memory to recover the object identity. The top-down information quantified by the a priori probabilities $\{P[\mathbf{b}_{k+1}|\mathbf{b}_k]\}$ will be assumed as uniform (i.e., equally-probable) among the different competing stimuli representations, and thus will not affect the transition lengths in (6). The decoding procedure applied to the sequence in (11) is shown via the trellis diagram of **Figure 8** with the decoder's initial condition given by $\mathbf{b}_0 = S_0$. At time $k = 1$ the decoder computes the transition lengths via (6) as

$$\begin{aligned} \lambda(\mathbf{b}_1 = S_0, \mathbf{b}_0 = S_0) &= -\ln(P[DCA|\mathbf{b}_1 = S_0, \mathbf{b}_0 = S_0]) \\ &= -\ln(P[r_1 = D|\text{input} = 0]) \\ &\quad -\ln(P[r_1 = C|\text{input} = 0]) \\ &\quad -\ln(P[r_1 = A|\text{input} = 0]) = 4.82 \\ \lambda(\mathbf{b}_1 = S_2, \mathbf{b}_0 = S_0) &= -\ln(P[DCA|\mathbf{b}_1 = S_2, \mathbf{b}_0 = S_0]) \\ &= -\ln(P[r_1 = D|\text{input} = 1]) \\ &\quad -\ln(P[r_1 = C|\text{input} = 1]) \\ &\quad -\ln(P[r_1 = A|\text{input} = 1]) = 4.42. \end{aligned} \tag{12}$$

The two possible transitions above are considered by the Viterbi algorithm because of the encoder's state diagram in **Figure 3B**. It should be noted that at $k = 1$ there are two rather than $2^L = 4$ survivors because the decoding has just commenced. Subsequently, the decoder computes

$$\begin{aligned} \Gamma(\mathbf{b}_1 = S_0, \mathbf{b}_0 = S_0) &= \Gamma(\mathbf{b}_0 = S_0) + \lambda(\mathbf{b}_1 = S_0, \mathbf{b}_0 = S_0) = 0 + 4.82 \\ \Gamma(\mathbf{b}_1 = S_2, \mathbf{b}_0 = S_0) &= \Gamma(\mathbf{b}_0 = S_0) + \lambda(\mathbf{b}_1 = S_2, \mathbf{b}_0 = S_0) = 0 + 4.42. \end{aligned} \tag{13}$$

This process is repeated for time $k = 1, 2, \ldots K$ (see the steps in **Appendix B**). At $k = K$ the Viterbi algorithm proceeds backwards in the trellis of **Figure 8** to arrive at $(\mathbf{b}_1 = S_2, \mathbf{b}_0 = S_0)$, $(\mathbf{b}_2 = S_3, \mathbf{b}_1 = S_2)$, $(\mathbf{b}_3 = S_1, \mathbf{b}_2 = S_3)$, $(\mathbf{b}_4 = S_0, \mathbf{b}_3 = S_1)$ as the final survivor path because it has the smallest metric among all of the candidates. The decoded sequence corresponding to this survivor path is

$$\widehat{\mathbf{s}} = 111\ 010\ 110\ 011 \tag{14}$$

which corresponds to the decoder's estimate of the encoded message being

$$\widehat{\mathbf{b}} = 1100. \tag{15}$$

By decoding, the stream $\mathbf{r}$ in the above example into the sequence $\widehat{\mathbf{b}}$, the IT has untangled the object's representation that was propagated along the VVS. In the above example $\widehat{\mathbf{b}} = \mathbf{b}$ which indicates perfect object recognition at the IT.

The assignment of uniform priors to the metric in (6) has the biological ramification of the IT having no prior memory, or synonymously, an unbiased account of what object to expect. If the VVS were to have identified an object in the prior $K = 4$ discrete time instances, then it would be sensible for the IT to have non-uniform priors with the first term in (6) biasing the transition metrics toward a particular representation. In organizing the CTOR model to emulate VVS operation, an updating rule should be presented to adjust the priors based on the object that was decoded in previous epochs, or is expected during the current viewing interval. As discussed in section 4.3, the a priori probabilities will be communicated from memory to the decoder (via the feedback signal in **Figure 1**) to be used in the ensuing decoding. Formulation of an updating rule for the priors that are stored in memory is an important future avenue because it would further substantiate the model's biological feasibility.

Properties such as poor visibility and a subject's inattention are factors that can adversely affect the decoding process by bringing about a channel with a low CQM. This will affect the decoding process in a conspicuous manner regardless of the decoder's proficiency. For instance, consider a case where (10) is replaced with the following channel

$$\begin{aligned} P[r_i = j | s_i = 0] &= 0.25 \\ P[r_i = j | s_i = 1] &= 0.25 \quad \text{for} \quad j = A, B, C, D \end{aligned} \tag{16}$$

that has a conditional entropy of $H(r_i|s_i = 0) = H(r_i|s_i = 1) = 2$ bits. Assuming uniform priors, it can be confirmed from (6) that the above channel would yield transition metrics of equal value at each time instant in the decoding process. The consequence of this is that the IT will have no choice but to arbitrarily select one of the possible $1/2^{KM}$ sequences. The preceding is an example of how recognition can be obscured by a catastrophically bad channel. The properties associated with the stimulus, environment, and neural circuitry that may bring about such a channel are not immediately obvious, but this is a question that warrants scrutiny. Following the decoding operation it is possible that the IT is indecisive as to the stimulus identity. In such a case the VVS may declare an erasure (Forney, 1968) as a means of requiring additional time to decide upon the identity or attributes of the viewed object. From a psychophysics perspective it is expected that the erasure is reflected by a higher reaction time and degraded processing speed for recognizing the object. The dynamics and threshold associated with the declaration of an erasure by the IT after decoding is an avenue for future consideration. It is interesting to note that an erasure may not be a complete waste of time and resources by the VVS since information may be gained and used about the viewed object at subsequent time epochs. This is expected of an adaptive system that has been optimized through continuous training and evolution.

5.3. Declaration of Object Category

In this example, we assume $\mathbf{w}^T = [1, 1, 1, 1]$ which would result in $f(\widehat{\mathbf{b}}) = \mathbf{w}^T\widehat{\mathbf{b}} = 2$ via (9). This operation is perhaps too elementary in this toy example because we except more than four object categories to exist during the viewing of a stimulus. It is more insightful to examine the scenario given by (8). The assignment of $\mathbf{w}$ as a 15-dimensional vector of 1's yields $f(\widehat{\mathbf{b}}) = 8$ for the three decoded sequences of (8), and hence the three stimuli would be categorized into the same category. There are three important points that follow with respect to this dubious outcome. First, although the three objects would be classified under the same category, their differing features can be still discerned by the IT assuming a sufficient degree of redundancy at the encoder, a channel that is not too dispersive, and adequate processing at the decoder. Second, the choice of $\mathbf{w}$ has not been determined via an SVM or even a correlation-based classifier as considered in works such as Rust and DiCarlo (2010) and Meyers et al. (2008), respectively. Both techniques would provide a $\mathbf{w}$ that has been acquired via a training process on already-viewed stimuli. For instance, it is easy to confirm that the (non-unique) choice $\mathbf{w}^T = [1, 0, 1, 1, 0, 1, 0, 1, 0, 0, 1, 0, 0, 1, 1]$ would yield perfect classification of category for this example. Third, a more biologically realistic scenario would have a significantly larger N value. This would lead to greater granularity among the object categories and significantly better classification capability.

5.4. Interleaver Operation, Message Compression, and Memory

The interleaving operation is considered as a means for identifying the important features in the decoded sequence. In this example we consider two features via $F = 2$ with $M_1 = 2$ and $M_2 = 2$. The static interleaving operation given by $\mathcal{I} = [[1, 3], [2, 4]]$ will be assumed where $[1, 3]$ signifies the bit indices corresponding to the more important feature M_1, and $[2, 4]$ refers to the bit indices of a less important feature M_2. Thus, the interleaving will lead to the following grouping

$$\widehat{\mathbf{b}} = 1100 \ \rightarrow \ \widetilde{\mathbf{b}} = [{}^1\widehat{b}_1, {}^2\widehat{b}_2, {}^1\widehat{b}_3, {}^2\widehat{b}_4] = {}^1 1 {}^2 1 {}^1 0 {}^2 0$$

where the left superscript of each bit indicates its importance level as dictated by the interleaver. The operation $\mathcal{I}$ is equivalent to a mapping that ranks the importance of the bits in the decoded stream via the features that they correspond to. In the present

example, the more important feature, M_1, represents the object identity while M_2 will correspond to the object's size. As shown in **Figure 1**, the interleaving operation is dynamically guided by attention in the prioritization of sensory information that is necessary for a task. To consider an instance of compression, suppose that the same object is presented to a subject during the next three viewing intervals with the corresponding representation given by $\widehat{\mathbf{b}} = 1100\ 1100\ 1110\ 1101$. The third representation has a different object identity than the other three while the fourth representation differs from the first two due to the object's size varying as a result of a change in viewing distance. The 16-bit decoded version of this sequence $\widehat{\mathbf{b}}$ would be interleaved to $\tilde{\mathbf{b}}$. With the hippocampus performing a compressive operation, the most important features of the representation $\tilde{\mathbf{b}} = {}^1 1^2 1^1 0^2 0 \quad {}^1 1^2 1^1 0^2 0 \quad {}^1 1^2 1^1 1^2 0 \quad {}^1 1^2 1^1 0^2 1$ are committed to memory with less compression than the less-important features. More specifically, the features with a left superscript of "1" are 10 and 11 while the less important features have been labeled via a superscript of "2," i.e., 10 and 11. A compression mechanism may entail the less important features being compressed among viewed objects that share the same important features. In the present example, the first and second representations 1100, and the fourth representation 1101 share the same value of important feature 10. These three objects can be compressed in memory as 1100 since it appears more frequently than 1101. The other value for the important feature (i.e., 11) occurs only in the third viewing of the object via 1110 which is stored in memory as well. Thus, for the considered compression, the 16-bit input $\widehat{\mathbf{b}} = 1100\ 1100\ 1110\ 1101$ is compressed into the 8-bit representation $\tilde{\mathbf{b}}^c = 1100\ 1110$.

Having attained a compressed representation of the viewed object, it is possible for the a priori probabilities $\{P[\mathbf{b}_{k+1}|\mathbf{b}_k]\}$ to be computed at the hippocampus and amygdala. The neural circuitry can estimate $P(\mathbf{b}_{k+1} = S_1|\mathbf{b}_k = S_3)$ from the sequence $\tilde{\mathbf{b}}^c = 1100\ 1110$ by counting the occurrences of $S_1 = 10$ after $S_3 = 11$ and normalizing that value by the occurrences of $S_3 = 11$. In the present example, $S_3 = 11$ occurs three times with two occurrences followed by $S_1 = 10$, therefore $P(\mathbf{b}_{k+1} = S_1|\mathbf{b}_k = S_3) = 2/3$. A more descriptive analysis of how the VVS may perform such a calculation for the remaining a priori probabilities is provided in **Appendix C**.

6. THE PERFORMANCE OF CTOR, AND USE OF PRIOR KNOWLEDGE IN OBJECT RECOGNITION

The previous section provided an instantiation of CTOR operation. It is also necessary to have an idea of the performance that is possible with this model. Accordingly, a more realistic scenario must be considered than the toy example of the previous section; clearly a larger stream is present as the input to the primate VVS. Since CTOR is not based on a neural network or a SVM, the metrics used to assess the performance of models based on the aforementioned methods are by-in-large not applicable here. To assess the performance of CTOR, several metrics must be discussed within the object recognition paradigm.

- *Bit Correct Rate (BCR)*: According to (7) the Hamming distance $d(\widehat{\mathbf{b}}, \mathbf{b}^i)$ was derived between the decoded sequence at the IT and the representation that the object should induce at the IT. The expression

$$\mathrm{BCR} = 1 - \frac{1}{T}\sum_{t=1}^{T}\left[\frac{d(\widehat{\mathbf{b}}, \mathbf{b}^i)}{KM}\right]_t \tag{17}$$

  provides a measure of the deviation between the expected and decoded representations over T viewed sequences. In the above expression $[X]_t$ denotes the value of the argument X at the t-th iteration. It is not difficult to observe that at chance $\mathrm{BCR} = 1/2$.
- *Symbol Correct Rate (SCR)*: A more stringent measure of correct object recognition is given by

$$\mathrm{SCR} = \frac{1}{T}\sum_{t=1}^{T}\left[\mathbb{1}(d(\widehat{\mathbf{b}}, \mathbf{b}^i) = 0)\right]_t. \tag{18}$$

  In the above expression $\mathbb{1}(\cdot)$ denotes the indicator function and the condition inside the indicator function is only satisfied for perfect recovery of the object identity. It can be verified that at chance $\mathrm{SCR} = 1/2^{KM}$.
- *Category Correct Rate (CCR)*: A distinction between object recognition and categorization has been made in the presentation of CTOR. Accordingly, we consider a measure for the correct identification of object category via

$$\mathrm{CCR} = \frac{1}{T}\sum_{t=1}^{T}\left[\mathbb{1}(\mathbf{w}^T\widehat{\mathbf{b}} = \mathbf{w}^T\mathbf{b}^i)\right]_t. \tag{19}$$

  It should be apparent that the classification vector $\mathbf{w}$ used in (19) is derived based on a classifier cost function rather than the CCR metric, otherwise the trivial solution $\mathbf{w} = \mathbf{0}$ would result. At chance this metric will equal the reciprocal of the number of categories considered i.e., $\mathrm{CCR} = 1/KM$.
- *Approximate Category Correct Rate (ACCR)*: A less stringent measure of categorization accuracy follows from considering the metric

$$\mathrm{ACCR} = \frac{1}{T}\sum_{t=1}^{T}\left[\mathbb{1}(|\mathbf{w}^T\widehat{\mathbf{b}} - \mathbf{w}^T\mathbf{b}^i| \le c_1)\right]_t. \tag{20}$$

  The constant $c_1 > 0$ is the maximum tolerable difference between the expected and recovered representation for the category to be determined at an acceptable fidelity. By its definition it can be noted that $\mathrm{AACR} \geq \mathrm{CCR}$.

The performance of the CTOR model for the VVS will be analyzed for all of the aforementioned metrics. We are, in effect, attempting to justify the utility of the BCR, SCR, CCR, and ACCR within the object recognition paradigm. It is interesting that the presented dialogue has provided a means to

quantitatively decipher between categorization performance and object recognition performance. To the best of our knowledge prior works have not made such a quantitative distinction and this may be viewed as a void in object recognition models.

The performance of CTOR with the simple-encoder in **Figure 3A** and its corresponding decoder implementing the Viterbi algorithm will be analyzed via simulation. The object representation $\mathbf{b}$ will be comprised of $K = 6$, 12, 24, 36, 48, or 60 bits meaning that the encoder will entangle such representations into a sequence $\mathbf{s}$ consisting of $3K$ bits. Without loss of generality, we specify the viewed stimulus as having a representation at the IT given by an alternating sequence of 1 and 0, e.g., for $K = 6$, $\mathbf{b} = 101010$. For each object $T = 10^6$ iterations will be considered in a Monte-Carlo (MC) simulation. Each iteration entails the components of the encoded sequence being probabilistically perturbed by the channel (10). The dispersive nature of (10) will be shown by examining object recognition in a less-dispersive discrete memoryless channel where the transition probabilities are given by

$$\begin{aligned} P[r_i = A|s_i = 0] &= 0.65 \\ P[r_i = B|s_i = 0] &= 0.2 \\ P[r_i = C|s_i = 0] &= 0.1 \\ P[r_i = D|s_i = 0] &= 0.05 \\ P[r_i = A|s_i = 1] &= 0.05 \\ P[r_i = B|s_i = 1] &= 0.1 \\ P[r_i = C|s_i = 1] &= 0.2 \\ P[r_i = D|s_i = 1] &= 0.65 \end{aligned} \tag{21}$$

and lead to $H(r_i|s_i = 0) = H(r_i|s_i = 1) = 1.416$ bits. **Figures 9A,B** contain simulation results of the object recognition metrics for CTOR with simple-encoder and the channels given by (10) and (21). A value of $c_1 = 1$ was used when computing the ACCR metric, in other words, a disagreement in Hamming distance of one between $\widehat{\mathbf{b}}$ and $\mathbf{b}^i$ was deemed tolerable in the recovery of the object category. The representation of more complicated stimuli would require a larger number of bits and may result in a degradation in the VVS's capability to accurately perform object recognition and classification. It is also expected that the more complex stimuli will require increased amounts of neural processing leading to longer message lengths (i.e., larger K values). The more convoluted an object, the worse a subject's performance in recognizing, categorizing, and parsing the attributes of the object. The CTOR model is capable of reflecting this aspect that seems fundamental to the working of the VVS. Indeed, the BCR, SCR, CCR, and ACCR metrics in **Figures 9A,B** show a degradation with increasing K values. It is interesting that the SCR shows the most precipitous degradation with increasing object complexity. This is attributed to the correct decoding of the entire object representation being more difficult, and hence more sensitive to the viewed object complexity, than a partial or a category-only recovery. A comparison of the four metrics shown in **Figures 9A,B** confirm an improvement in recognition and categorization performance when the channel is represented by (21) instead of (10). Thus, the impact of a degradation in CQM on object recognition is patent since the recognition and classification accuracies are lower for the more dispersive channel. The consideration of the retina, LGN, and V1 stage via a more sophisticated encoder shall be referred to as "complex-encoder." The rate of the encoder is maintained at 1/3, but a maximal memory order of $L = 8$ is considered via the following generator sequence

$$\begin{aligned} G(D) = \big[&1 + D^2 + D^3 + D^5 + D^6 + D^7 + D^8, 1 + D + D^3 + D^4 \\ &+ D^7 + D^8, 1 + D + D^2 + D^5 + D^8\big]. \end{aligned}$$

The above encoder has been studied in Lin and Costello (1983); its shift-register structure and state diagram are not shown because of their involved nature in comparison to simple-encoder. For instance, the decoding would consist of a trellis with $2^8 = 256$ states at time k and two prospective transitions out of each state. Via a higher constraint length (L), there are a larger number of paths to compare at each stage of the trellis and this leads to an increase in resolution when making a decision on every encoded bit. Thus, a decoder that would accommodate complex-encoder will generally be more accurate in recovering representations than the decoder accommodating simple-encoder. The performance of CTOR with complex-encoder shall now be assessed. Due to the increased complexity and run time, rather than using the complete Viterbi algorithm that was used for simple-encoder, MATLAB's vitdec(·) function with soft-decision decoding and 4 levels (i.e., nsdec = 2) were used in the simulations with complex-encoder. It shall still be assumed that the viewed object has a representation at the IT given by an alternating sequence of 1 and 0. The evaluation of the BCR, SCR, CCR, and ACCR in **Figures 9C,D** show that similar conclusions can be drawn for CTOR with complex-encoder as with simple-encoder. With the exception of the BCR for the channel of (21), the metrics in **Figure 9** exhibit a degradation with increasing K values.

There are findings to discuss in light of the simulation results shown in **Figure 9**. The BCR appears to be the most robust of the metrics with respect to increasing degrees of stimulus complexity. By definition the BCR is restricted to the interval [0.5, 1], and the observed limited range in comparison to the other metrics in the simulations indicates that the BCR may not be as insightful of a metric. A comparison of the performance of simple- vs. complex-encoder shows that the latter exhibits a clear improvement across all of the metrics for the less-dispersive channel. Interestingly, the affect of the channel is more pronounced on the metrics for complex-encoder than for simple-encoder. In the case of the highly dispersive channel, however, the two systems yield similar performance. This is attributed to the increased processing not being able to overcome the detriments brought forth by the high dispersion. For an engineered system, so long as the channel is not overly dispersive, a higher L is desirable because it yields more reliable communication (i.e., higher BCR and SCR), the tradeoff is that an increase in constraint length leads to an increase in complexity and processing. Of course the VVS is not subject to the same tradeoffs that exist in engineered systems, thus it may be presumed that a CTOR implementation of the VVS will entail a large L value and accommodate the decoder (i.e., IT

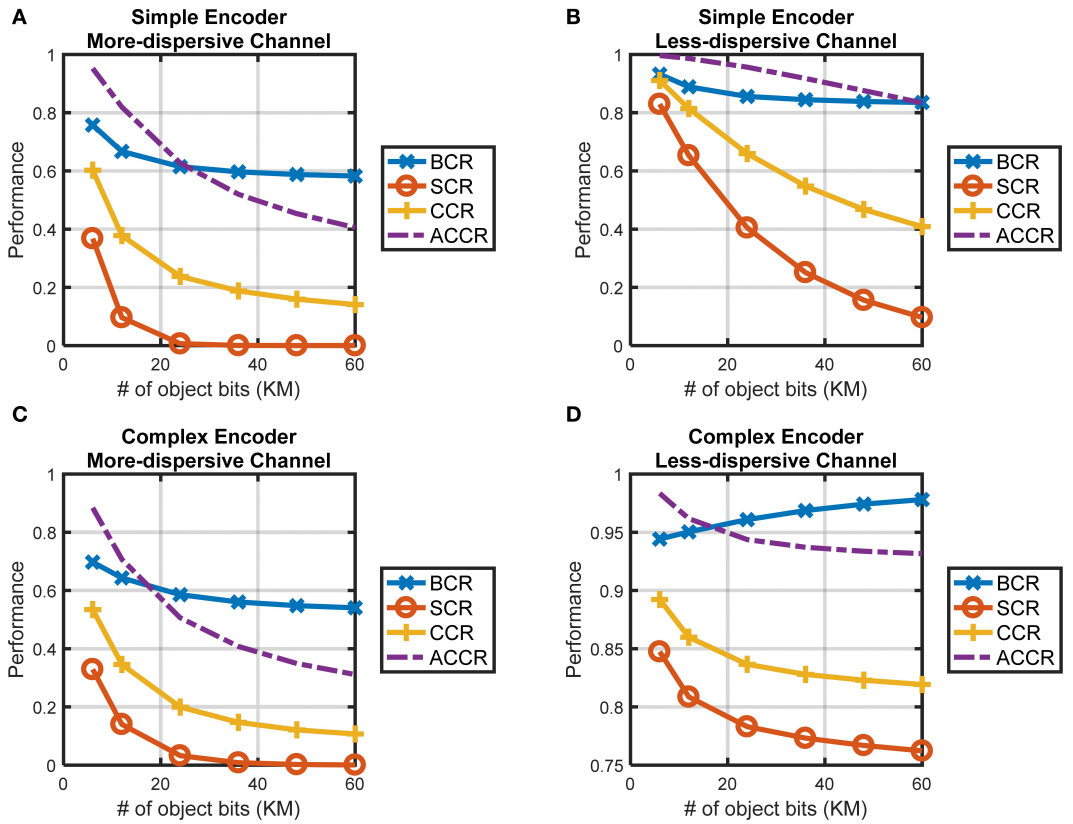

FIGURE 9 | The performance of CTOR with the encoding, channel, and decoding structures discussed in this work. **(A)** The performance attained with simple encoder and the more-dispersive channel in (10), and **(B)** with the less-dispersive channel in (21). The decoding for the simple encoder was implemented with custom code in MATLAB. **(C)** The performance attained for complex encoder with the more-dispersive channel, and **(D)** with the less-dispersive channel. The decoding for the complex decoder was implemented using MATLAB's vitdec(·) function with 4 levels specified in the Viterbi soft-decision decoding algorithm.

circuitry) being able to process an immense number of states in the prospective trellis.

The operation of the CTOR model has been studied with the view of uniform a priori probabilities for the state transitions that are used at the IT for untangling the representation of the viewed object. While this has been done for ease in presentation, such an assumption reduces maximum a posteriori probability (MAP) decoding to maximum likelihood estimation (MLE). We shall now consider the scenario of non-uniform a priori probabilities for the state transitions. The simple encoder, more dispersive channel in (10), and an input sequence of $K = 60$ bits shall be assumed in considering the Viterbi algorithm operation with MLE (**Figure 10A**) and MAP (**Figure 10B**) decoding. The simulations were performed in the same manner as in **Figure 9** except $T = 10^3$ iterations were considered for each object. With MAP decoding, the a priori probabilities are estimated from the previous iteration using the technique described in section 5.4. The BCR attained with MLE fluctuates between 0.4 and 0.8 and has a mean of 0.581, while the mean BCR attained with MAP is 0.695 with the BCR equating to 1 at several iterations (**Figure 10C**). The SCR results show a similar trend as the MLE does not correctly recover the entire object at any iteration (i.e., mean SCR = 0) whereas MAP is able to do so (mean SCR = 0.032) (**Figure 10D**). A kymograph of the decoded bits across each iteration illustrates that the correctly recovered bits are more clustered for MAP decoding than with MLE (**Figure 10E**). This is because knowledge of the a priori probabilities guides the fidelity with which consecutive bits are decoded. It is observed that the a priori probabilities computed at the decoder during MAP decoding are rather constant across the 10^3 iterations (**Figure 10F**). This is expected because the same input sequence was used for each iteration. Furthermore, we note that the state transitions $S_2 \rightarrow S_1$ and $S_3 \rightarrow S_1$ are assigned the highest a priori probabilities. This is also expected since it can be verified that the transitions $11 \rightarrow 10$ and $01 \rightarrow 10$ will be the most frequent transitions for the considered input stream. In summary, the CTOR formulation with MAP decoding surpasses the performance noted with MLE (**Figures 10C,D**), thus confirming the value of the feedforward-feedback interaction between the IT and hippocampus during object recognition.

CONCLUSION

Three communities are concurrently involved in the comprehension of visual object recognition: neuroscientists, computer vision scientists, and visual psychophysicists. The presented CTOR model has drawn upon elements advocated

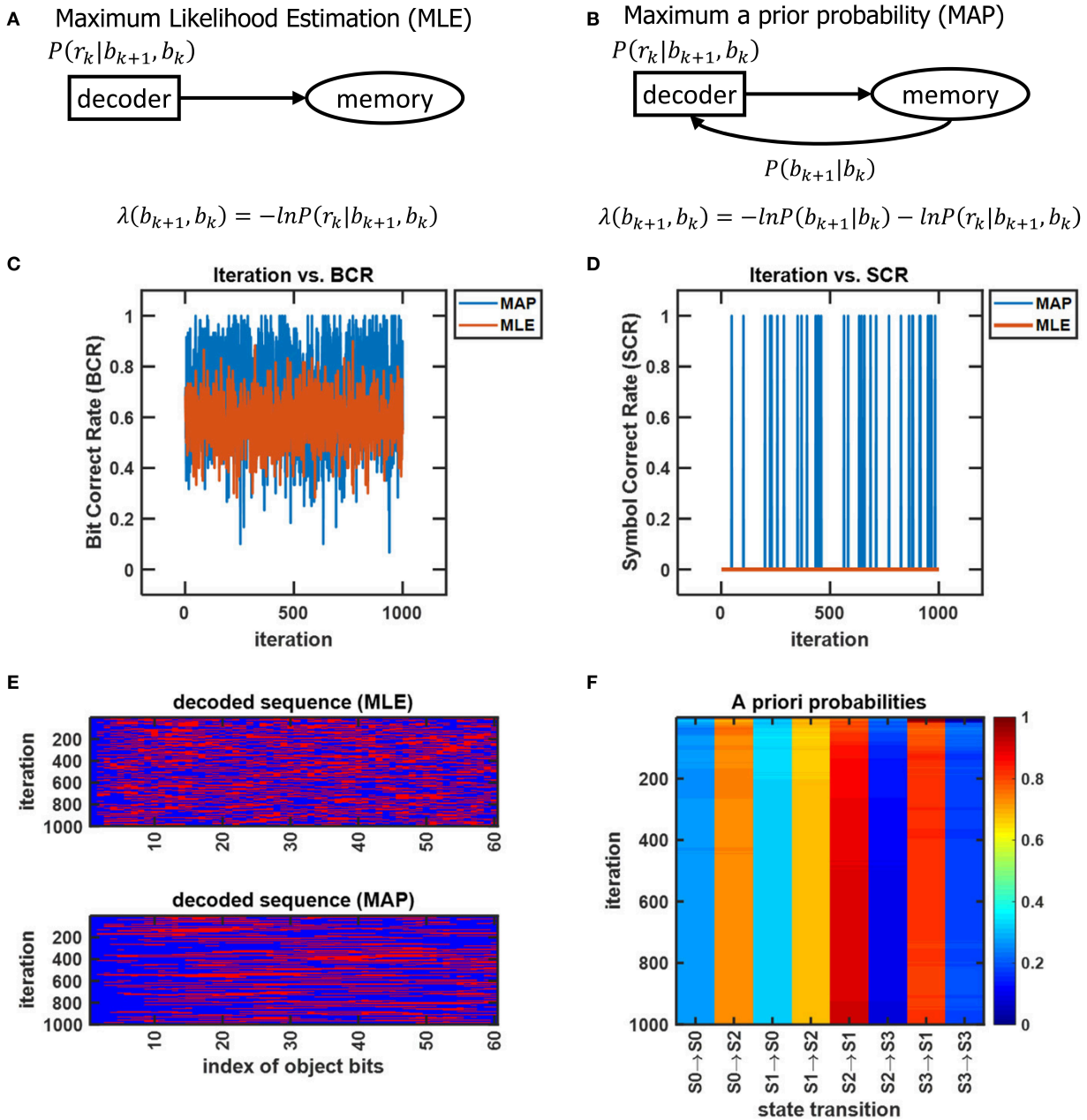

FIGURE 10 | A comparison of MLE and MAP with a simple encoder, the dispersive channel of (10) and $K = 60$ bit object identity. **(A,B)** Algorithmic depictions of the MLE and MAP decoding schemes. The transition length $\lambda(\cdot, \cdot)$ is computed for both schemes, and for MAP decoding, includes the a priori probability $P[\mathbf{b}_{k+1}|\mathbf{b}_k]$ in addition to the transition probability $P[\mathbf{r}_k|\mathbf{b}_{k+1}, \mathbf{b}_k]$. **(C)** The progression of the BCR at each iteration of CTOR when considering MLE and MAP for the same viewed object. **(D)** The progression of the SCR at each iteration of CTOR when considering MLE and MAP for the same viewed object. It should be noted that the SCR can only take on values of $\{0, 1\}$ at any iteration. **(E)** A kymograph of the decoded sequence over different viewing intervals (i.e., iterations). Red represents a bit correctly identified by the decoder and blue indicates that the bit was erroneously identified. **(F)** A kymograph of the a priori probabilities calculated by the attention and memory circuitry over the duration of each viewing interval.

from the three realms. Previously considered for lower visual areas, dynamic inference via an on-line algorithm for MAP sequence estimation has been proposed for the higher visual areas implicated during object recognition. Although the primary motivation for CTOR is to provide an account for the proficiency of the IT, the formulation is also a starting point for a more comprehensive scrutiny of the computations performed by the VVS during real-time object recognition. The performance of the model was evaluated by presenting several metrics to assess categorization accuracy and object identity recognition. The simulation results provide insight into the dynamics and capabilities of CTOR. The role of attention and memory

have been incorporated via top-down signaling that guides the inference, and is also affected by the cognition. Empirical corroboration of CTOR would entail presentation of data to support or verify the algorithmic notions discussed in this work. In order to test or affirm aspects of CTOR in the framework of current knowledge, it is crucial to consider primate neuroscience studies that have already amassed high-dimensional recordings from multiple brain regions and pursued computational questions. The study in Shinomoto et al. (2009) considers neural spike data from 15 cortical areas in awake, behaving monkeys that were collected at different labs. The authors used this data to make statements about the functional category of the cortical area. A similar methodology can be used to assess aspects of CTOR. For instance, an experiment could entail showing the same objects to subjects, recording V1, V4, and IT neural responses, and amassing the collected data among different labs into one dataset to evaluate the encoding and decoding operations. Initially, the IT neural population responses would be compared to the V1 responses in order to determine the encoder. In effect, a code rate, constraint length, and encoder structure would be assumed, evaluated, and altered in iterative fashion until a candidate has been deemed as fitting the data appropriately. Such iterative searches are routinely performed by coding theorists—e.g., for convolutional codes see Conan, 1984; Chang et al., 1997; Katsiotis et al., 2010—to discover encoders that satisfy a criterion. The considered scenario is unique in selecting the code that best fits the data in connecting the IT response to the V1 response. Subsequently, parameters associated with the decoder and the channel can be evaluated or fit to the V4 neural population responses from the same viewed objects. Such analysis would also require initial assumptions about the channel (e.g., continuous vs. discrete) and the decoder prior to performing the iterative searches over their associated parameter spaces. In a different study, Lehky et al. (2011) recorded responses of 674 IT neurons across two monkeys as they were shown 806 objects. The authors analyzed the data in holistic fashion to determine that the heavy tails of the population responses are suggestive of different neurons being tuned to different features. More recently, Dong et al. (2017) incorporated the 806×674 data matrix of the aforementioned work to develop simulations for a large number of neuronal responses with various settings for neuron number, stimulus number and identity, and noise level. Through their simulations, the authors justify the findings in Lehky et al. (2011) and also provide an instance of how information can be extracted from a dataset to test additional hypotheses with different assumptions for the underlying processes. Similar to the analysis of Dong et al. (2017), the CTOR hypotheses can be scrutinized by simulating the neuronal responses of the populations in **Figures 1**, **2** with the variables listed in **Appendix D**. Although CTOR is a proposition; it is biologically inspired, motivated by prior empirical discussions, and mirrors the tangling-untangling notion that has been accredited within the primate vision community.

DATA AVAILABILITY STATEMENT

The source code for this study can be found on github (https://github.com/shenghuanjie/ctor).

AUTHOR CONTRIBUTIONS

This work was conceived by extensive discussions between all authors. SS and HS conceived the hypothesis and performed the simulations. SS and HS wrote the manuscript with feedback from HP.

ACKNOWLEDGMENTS

This work was supported in part by the U.S. National Science Foundation under Grant CCF-0939370. The authors would like to thank Kendrick Kay and Bruno Olshausen for introducing them to the general problem statement and providing several helpful discussions. We also thank Max Riesenhuber for valuable feedback on the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncom.2020.00046/full#supplementary-material

REFERENCES

Andrews, K., Heegard, C., and Kozen, D. (1997). *A Theory of Interleavers*. Technical report, Cornell University.

Baldauf, D., and Desimone, R. (2014). Neural mechanisms of object-based attention. *Science* 344, 424–427. doi: 10.1126/science.1247003

Biederman, I. (1987). Recognition-by components: a theory of human image understanding. *Psychol. Rev.* 94, 115–147. doi: 10.1037/0033-295X.94.2.115

Brady, T. F., Konkle, T., and Alvarez, G. A. (2009). Compression in visual working memory: using statistical regularities to form more efficient memory representations. *J. Exp. Psychol. Gen.* 138:487. doi: 10.1037/a0016797

Cadieu, C., Kouh, M., Pasupathy, A., Connor, C., Riesenhuber, M., and Poggio, T. (2007). A model of v4 shape selectivity and invariance. *J. Neurophysiol.* 98, 1733–1750. doi: 10.1152/jn.01265.2006

Chang, J.-J., Hwang, D.-J., and Lin, M.-C. (1997). Some extended results on the search for good convolutional codes. *IEEE Trans. Inform. Theory* 43, 1682–1697. doi: 10.1109/18.623175

Chelazzi, L., Duncan, J., Miller, E., and Desimone, R. (1998). Responses of neurons in inferior temporal cortex during memory-guided visual search. *J. Neurophysiol.* 80, 2918–2940. doi: 10.1152/jn.1998.80.6.2918

Conan, J. (1984). The weight spectra of some short low-rate convolutional codes. *IEEE Trans. Commun.* 32, 1050–1053. doi: 10.1109/TCOM.1984.1096180

Cukura, T., Nishimoto, S., Huth, A., and Gallant, J. (2013). Attention during natural vision warps semantic representation across the human brain. *Nat. Neurosci.* 16, 763–770. doi: 10.1038/nn.3381

DeFelipe, J., and Farinas, I. (1992). The pyramidal neuron of the cerebral cortex: morphological and chemical characteristics of the synaptic inputs. *Progr. Neurobiol.* 39, 563–607. doi: 10.1016/0301-0082(92)90015-7

DiCarlo, J., and Cox, D. (2007). Untangling invariant object recognition. *Trends Cogn. Sci.* 11, 333–341. doi: 10.1016/j.tics.2007.06.010

DiCarlo, J., Zoccolan, D., and Rust, N. (2012). How does the brain solve visual object recognition? Neuron 73, 415–434. doi: 10.1016/j.neuron.2012.01.010

Dong, Q., Liu, B., and Hu, Z. (2017). Comparison of it neural response statistics with simulations. *Front. Comput. Neurosci.* 11:60. doi: 10.3389/fncom.2017.00060

Eddy, S. R. (2004). What is a hidden Markov model? *Nat. Biotechnol.* 22:1315. doi: 10.1038/nbt1004-1315

Elston, G. N. (2002). Cortical heterogeneity: implications for visual processing and polysensory integration. *J. Neurocytol.* 31, 317–335. doi: 10.1023/A:1024182228103

Elston, G. N. (2003). Cortex, cognition and the cell: new insights into the pyramidal neuron and prefrontal function. *Cereb. Cortex* 13, 1124–1138. doi: 10.1093/cercor/bhg093

Elston, G. N., Benavides-Piccione, R., Elston, A., Manger, P., and DeFelipe, J. (2011). Pyramidal cells in prefrontal cortex of primates: marked differences in neuronal structure among species. *Front. Neuroanat.* 5:2. doi: 10.3389/fnana.2011.00002

Elston, G. N., Elston, A., Kass, J. H., and Casagrande, V. (2005). Regional specialization in pyramidal cell structure in the visual cortex of the galago: an intracellular injection study of striate and extrastriate areas with comparative notes on new world and old world monkeys. *Brain Behav. Evol.* 66, 10–21. doi: 10.1159/000085044

Elston, G. N., and Fujita, I. (2014). Pyramidal cell development: postnatal spinogenesis, dendritic growth, axon growth, and electrophysiology. *Front. Neuroanat.* 8:78. doi: 10.3389/fnana.2014.00078

Elston, G. N., Tweedale, R., and Rosa, M. (1999). Cortical integration in the visual system of the macaque monkey: large scale morphological differences of pyramidal neurons in the occipital, parietal and temporal lobes. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 266, 1367–1374. doi: 10.1098/rspb.1999.0789

Fano, R. (1963). A heuristic discussion of probabilistic decoding. *IEEE Trans. Inform. Theory* 9, 64–74. doi: 10.1109/TIT.1963.1057827

Feldman, M. L. (1984). "Morphology of the neocortical pyramidal neuron," in *Cerebral Cortex: Cellular Components of the Cerebral Cortex*, eds A. Peters and E. G. Jones (New York, NY: Plenum Press), 123–200.

Forney, G. (1968). Exponential error bounds for erasure, list, and decision feedback schemes. *IEEE Trans. Inform. Theory* 14, 206–220. doi: 10.1109/TIT.1968.1054129

Forney, G. (1973). The Viterbi algorithm. *Proc. IEEE* 61, 268–278. doi: 10.1109/PROC.1973.9030

Hastie, T., Tibshirani, R., and Friedman, J. (2009). *The Elements of Statistical Learning*. New York, NY: Springer. doi: 10.1007/978-0-387-84858-7

Hubel, D., and Wiesel, T. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex? *J. Physiol.* 160, 106–154. doi: 10.1113/jphysiol.1962.sp006837

Huth, A., Nishimoto, S., Vu, A., and Gallant, J. (2012). A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron* 76, 1210–1224. doi: 10.1016/j.neuron.2012.10.014

Jacobs, B., Schall, M., Prather, M., Kapler, E., Driscoll, L., Baca, S., et al. (2001). Regional dendritic and spine variation in human cerebral cortex: a quantitative golgi study. *Cereb. Cortex* 11, 558–571. doi: 10.1093/cercor/11.6.558

Jacobs, B., and Scheibel, A. B. (2002). "Regional dendritic variation in primate cortical pyramidal cells," in *Cortical Areas: Unity and Diversity*, eds A. Schuez and R. Miller (New York, NY: CRC Press), 111–131. doi: 10.1201/9780203299296.pt2

Kanwisher, N., and Wojciulik, E. (2000). Visual attention: insights from brain imaging. *Nat. Rev. Neurosci.* 1, 91–100. doi: 10.1038/35039043

Karklin, Y., and Lewicki, M. (2009). Emergence of complex cell properties by learning to generalize in natural scenes. *Nature* 475, 83–87. doi: 10.1038/nature07481

Katsiotis, A., Rizomiliotis, P., and Kalouptsidis, N. (2010). New constructions of high-performance low-complexity convolutional codes. *IEEE Trans. Commun.* 58, 1950–1961. doi: 10.1109/TCOMM.2010.07.090149

Kersten, D., Mamassian, P., and Yuille, A. (2004). Object perception as Bayesian inference. *Annu. Rev. Psychol.* 55, 271–304. doi: 10.1146/annurev.psych.55.090902.142005

Lamme, V., and Roelfsema, P. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends Neurosci.* 23, 571–579. doi: 10.1016/S0166-2236(00)01657-X

Lee, T., and Mumford, D. (2003). Hierarchical Bayesian inference in the visual cortex. *J. Opt. Soc. Am.* 20, 1434–1448. doi: 10.1364/JOSAA.20.001434

Lehky, S. R., Kiani, R., Esteky, H., and Tanaka, K. (2011). Statistics of visual responses in primate inferotemporal cortex to object stimuli. *J. Neurophysiol.* 106, 1097–1117. doi: 10.1152/jn.00990.2010

Lin, S., and Costello, D. (1983). *Error Control Coding: Fundamentals and Applications*. Englewood Cliffs, NJ: Prentice-Hall.

Luebke, J. I. (2017). Pyramidal neurons are not generalizable building blocks of cortical networks. *Front. Neuroanat.* 11, 1–8. doi: 10.3389/fnana.2017.00011

Mar, D., Chow, C., Gerstner, W., Adams, R., and Collins, J. (1999). Noise shaping in populations of coupled model neurons. *Proc. Natl. Acad. Sci. U.S.A.* 96, 10450–10455. doi: 10.1073/pnas.96.18.10450

McAdams, C., and Maunsell, J. (1999). Effects of attention on orientation-tuning functions of single neurons in macaque cortical area v4. *J. Neurosci.* 19, 431–441. doi: 10.1523/JNEUROSCI.19-01-00431.1999

Meyers, E., Freedman, D., Kreiman, G., Miller, E., and Poggio, T. (2008). Dynamic population coding of category information in inferior temporal and prefrontal cortex. *J. Neurophysiol.* 100, 1407–1419. doi: 10.1152/jn.90248.2008

Mishkin, M. (1982). A memory system in the monkey. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 298, 85–95. doi: 10.1098/rstb.1982.0074

Mishkin, M., Ungerleider, L., and Macko, K. (1983). Object vision and spatial vision: two cortical pathways. *Trends Neurosci.* 6, 414–417. doi: 10.1016/0166-2236(83)90190-X

Miyashita, Y. (1993). Inferior temporal cortex: where visual perception meets memory. *Annu. Rev. Neurosci.* 16, 245–263. doi: 10.1146/annurev.ne.16.030193.001333

Nassar, M. R., Helmers, J. C., and Frank, M. J. (2018). Chunking as a rational strategy for lossy data compression in visual working memory. *Psychol. Rev.* 125:486. doi: 10.1037/rev0000101

Neri, P., and Heeger, D. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nat. Neurosci.* 5, 812–816. doi: 10.1038/nn886

Olshausen, B. (2013). "Perception as an inference problem," in *The Cognitive Neurosciences*, eds V. M. Gazzaniga and R. Mangun (Cambridge, MA: MIT Press).

Olshausen, B., and Field, D. (2005). How close are we to understanding v1. *Neural Comput.* 17, 1665–1699. doi: 10.1162/0899766054026639

O'Reilly, R., and McClelland, J. (1994). Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. *Hippocampus* 4, 661–682. doi: 10.1002/hipo.450040605

O'Reilly, R., Wyatte, D., Herd, S., Mingus, B., and Jilk, D. (2013). Recurrent processing during object recognition. *Front. Psychol.* 4, 1–14. doi: 10.3389/fpsyg.2013.00124

Pagan, M., Urban, L., Wohl, M., and Rust, N. (2013). Signals in inferotemporal and perirhinal cortex suggest an "untangling" of visual target information. *Nat. Neurosci.* 16, 1132–1139. doi: 10.1038/nn.3433

Petrantonakis, P., and Poirazi, P. (2014). A compressed sensing perspective of hippocampal function. *Front. Syst. Neurosci.* 8:141. doi: 10.3389/fnsys.2014.00141

Poggio, T., and Kreiman, G. (2013). *Biologically Inspired Circuits for Visual Search and Recognition in Complex Scenes*. MIT-Harvard Technical Report. doi: 10.21236/ADA579012

Ramsey, J. (1970). Realization of optimum interleavers. *IEEE Trans. Inform. Theory* 16, 338–345. doi: 10.1109/TIT.1970.1054443

Rao, R., and Ballard, D. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nat. Neurosci.* 2, 79–87. doi: 10.1038/4580

Reynolds, J., and Chelazzi, L. (2004). Attentional modulation of visual processing. *Annu. Rev. Neurosci.* 27, 611–647. doi: 10.1146/annurev.neuro.26.041002.131039

Riesenhuber, M., and Poggio, T. (1999). Hierarchical models of object recognition in cortex. *Nat. Neurosci.* 2, 1019–1025. doi: 10.1038/14819

Rockland, K. S. (1997). "Elements of cortical architecture," in *Cerebral Cortex: Extrastriate Cortex in Primates*, eds K. S. Rockland, J. H. Kaas, and A. Peters (Boston, MA: Springer), 243–293. doi: 10.1007/978-1-4757-9625-4_6

Roe, A. W. (2019). Columnar connectome: toward a mathematics of brain function. *Netw. Neurosci.* 3, 779–791. doi: 10.1162/netn_a_00088

Rust, N., and DiCarlo, J. (2010). Selectivity and tolerance ("invariance") both increase as visual information propagates from cortical area v4 to it. *J. Neurosci.* 39, 12978–12995. doi: 10.1523/JNEUROSCI.0179-10.2010

Salinas, E., and Abbott, L. (1997). Invariant visual responses from attentional gain fields. *J. Neurophysiol.* 77, 3267–3272. doi: 10.1152/jn.1997.77.6.3267

Serre, T., and Riesenhuber, M. (2004). Realistic modeling of simple and complex cell tuning in the HMAX model, and implications for invariant object recognition in cortex. *MIT CSAIL Memo* 3, 1–13. doi: 10.21236/ADA 459692

Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M., and Poggio, T. (2007). Robust object recognition with cortex-like mechanisms. *IEEE Trans Pattern Anal. Mach. Intell.* 29, 411–426. doi: 10.1109/TPAMI.2007.56

Shannon, C. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 379–423. doi: 10.1002/j.1538-7305.1948.tb01338.x

Shinomoto, S., Kim, H., Shimokawa, T., Matsuno, N., Funahashi, S., Shima, K., et al. (2009). Relating neuronal firing patterns to functional differentiation of cerebral cortex. *PLoS Comput. Biol.* 5:e1000433. doi: 10.1371/journal.pcbi.1000433

Simoncelli, E., and Olshausen, B. (2001). Natural image statistics and neural representation. *Annu. Rev. Neurosci.* 24, 1193–1216. doi: 10.1146/annurev.neuro.24.1.1193

Spruston, N. (2008). Pyramidal neurons: dendritic structure and synaptic integration. *Nat. Rev. Neurosci.* 9, 206–221. doi: 10.1038/nrn2286

Stevens, C. (2001). An evolutionary scaling law for the primate visual system and its basis in cortical function. *Nature* 411, 193–195. doi: 10.1038/35075572

Thorpe, S. D. A., and Van Rullen, R. (2001). Spike-based strategies for rapid processing. *Neural Netw.* 3, 715–725. doi: 10.1016/S0893-6080(01)00083-1

Treisman, A., and Gelade, G. (1980). A feature-integration theory of attention. *J. First Reference* 12, 97–136. doi: 10.1016/0010-0285(80)90005-5

Turk-Browne, N., Golomb, J., and Chun, M. (2013). Complementary attentional components of successful memory encoding. *J. First Reference* 66, 553–562. doi: 10.1016/j.neuroimage.2012.10.053

Usher, M., and Niebur, E. (1996). Modeling the temporal dynamics of it neurons in visual search: a mechanism for top-down selective attention. *J. Cogn. Neurosci.* 8, 311–327. doi: 10.1162/jocn.1996.8.4.311

Van Essen, D., and Gallant, J. (1994). Neural mechanisms of form and motion processing in the primate visual system. *Neuron* 10, 1–10. doi: 10.1016/0896-6273(94)90455-3

Wozencraft, J., and Jacobs, I. (1965). *Principles of Communication Engineering*. New York, NY: Wiley.

Wyatte, D., Curran, T., and O'Reilly, R. (2012). The limits of feedforward vision: recurrent processing promotes robust object recognition when objects are degraded. *J. Cogn. Neurosci.* 24, 2248–2261. doi: 10.1162/jocn_a_00282

Zhang, Y., Meyers, E., Bichot, N., Serre, T., Poggio, T., and Desimone, R. (2011). Object decoding with attention in inferior temporal cortex. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8850–8855. doi: 10.1073/pnas.1100999108

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.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 24 June 2020
doi: 10.3389/fncom.2020.00056

Serotonergic Axons as Fractional Brownian Motion Paths: Insights Into the Self-Organization of Regional Densities

Skirmantas Janušonis[1]*, *Nils Detering*[2], *Ralf Metzler*[3] and *Thomas Vojta*[4]

[1] Department of Psychological and Brain Sciences, University of California, Santa Barbara, Santa Barbara, CA, United States, [2] Department of Statistics and Applied Probability, University of California, Santa Barbara, Santa Barbara, CA, United States, [3] Institute of Physics and Astronomy, University of Potsdam, Potsdam, Germany, [4] Department of Physics, Missouri University of Science and Technology, Rolla, MO, United States

All vertebrate brains contain a dense matrix of thin fibers that release serotonin (5-hydroxytryptamine), a neurotransmitter that modulates a wide range of neural, glial, and vascular processes. Perturbations in the density of this matrix have been associated with a number of mental disorders, including autism and depression, but its self-organization and plasticity remain poorly understood. We introduce a model based on reflected Fractional Brownian Motion (FBM), a rigorously defined stochastic process, and show that it recapitulates some key features of regional serotonergic fiber densities. Specifically, we use supercomputing simulations to model fibers as FBM-paths in two-dimensional brain-like domains and demonstrate that the resultant steady state distributions approximate the fiber distributions in physical brain sections immunostained for the serotonin transporter (a marker for serotonergic axons in the adult brain). We suggest that this framework can support predictive descriptions and manipulations of the serotonergic matrix and that it can be further extended to incorporate the detailed physical properties of the fibers and their environment.

Keywords: brain, 5-hydroxytryptamine, serotonin, fibers, density, stochastic process, anomalous diffusion, fractional Brownian motion

OPEN ACCESS

Edited by:
Thomas Nowotny,
University of Sussex, United Kingdom

Reviewed by:
Massimo Pasqualetti,
University of Pisa, Italy
Lianchun Yu,
Lanzhou University, China

***Correspondence:**
Skirmantas Janušonis
janusonis@ucsb.edu

Received: *27 December 2019*
Accepted: *19 May 2020*
Published: *24 June 2020*

Citation:
Janušonis S, Detering N, Metzler R and Vojta T (2020) Serotonergic Axons as Fractional Brownian Motion Paths: Insights Into the Self-Organization of Regional Densities. Front. Comput. Neurosci. 14:56. doi: 10.3389/fncom.2020.00056

INTRODUCTION

All cells in vertebrate brains are surrounded by a matrix of highly tortuous fibers that release serotonin (5-hydroxytryptamine, 5-HT), a major neurotransmitter. Presently the self-organization and dynamics of this matrix are not understood beyond neuroanatomical descriptions. Altered densities of serotonergic (serotonin-releasing) fibers have been associated with many mental disorders and conditions, including Autism Spectrum Disorder (Azmitia et al., 2011), Major Depressive Disorder (Numasawa et al., 2017), epilepsy (Maia et al., 2019), and exposure to 3,4-methylenedioxymethamphetamine (MDMA, Ecstasy) (Adori et al., 2011). Predictive computational models can significantly advance this research and support its biomedical applications. Motivated by this potential, we used a stochastic process framework to develop a model of serotonergic fibers and reproduced some key features of their density distribution in the mouse brain.

Serotonergic fibers are axons of neurons whose bodies are located in several brainstem clusters known as the raphe nuclei (Stuesse et al., 1991; Jacobs and Azmitia, 1992; Hornung, 2003). In mammals, these neurons mature early in development. They begin synthesizing serotonin around embryonic day 11–13 in the mouse and rat brains (Lidov and Molliver, 1982; Hendricks et al., 1999; Hawthorne et al., 2010) and around 5 weeks of gestation in the human brain (Sundstrom et al., 1993; Mai and Ashwell, 2004). In the adult mammalian brain, serotonergic axons are unusual in their ability to regenerate, with potential implications for the efforts to restore other axon systems after injury (Hawthorne et al., 2011; Jin et al., 2016; Kajstura et al., 2018). A recent study has shown that they may share this property with other axons in the ascending reticular activating system (Dougherty et al., 2019).

Recent studies have revealed great diversity of serotonergic neurons (Okaty et al., 2015, 2019), the functional significance of which is an active area of research (Ren et al., 2018). Paradoxically, transgenic mouse models with no serotonin synthesis in the brain during development have no gross neuroanatomical abnormalities and show only mild behavioral deficits (Mosienko et al., 2015; Pratelli and Pasqualetti, 2019). Serotonergic neurons can release other major neurotransmitters, such as glutamate (Okaty et al., 2019) and perhaps GABA (Stamp and Semba, 1995; Okaty et al., 2019). In the raphe nuclei, serotonergic neurons coexist with many other neurons (Cardozo Pinto et al., 2019; Schneeberger et al., 2019), some of which may participate in stereotypic synaptic arrangements (Soiza-Reilly et al., 2013).

The development of serotonergic fibers is currently conceptualized to proceed in two or three stages: the initial growth in well-defined fiber tracts, followed by extensive arborization and eventual dispersal in "terminal" fields (Lidov and Molliver, 1982; Carrera et al., 2008; Kiyasova and Gaspar, 2011; Jin et al., 2016; Maddaloni et al., 2017; Donovan et al., 2019). This orderly sequence is generally consistent with experimental observations at the level of local fiber populations, visualized with tract-tracing techniques or quantified with density measures. It is also theoretically appealing in that it mirrors the development of brain projections that connect two well-defined brain regions (e.g., the lateral geniculate nucleus and the primary visual cortex).

There is little doubt that the initial serotonergic projections form well-defined paths, which we have studied in our own research (Janušonis et al., 2004; Slaten et al., 2010). These paths are well-described because they are followed by axon bundles, thus facilitating their visualization in time and space. In contrast, the processes that lead to the formation of regional fiber densities remain poorly understood. Fundamentally, they require a rigorous description of the behavior of single fibers, each one of which has a unique, meandering trajectory. Since these processes are the main focus of the present study, we note several important challenges.

The first detailed morphological description of single serotonergic axons has become available only recently (Gagnon and Parent, 2014). This study has reconstructed a small set of axons originating in the dorsal raphe nucleus and has concluded that they travel through multiple brain regions, rarely branching in some of them and producing profuse arborizations in others. However, true branching points are difficult to distinguish from highly tortuous fiber segments that simply pass each other, especially in bright field microscopy (used in this study). A branching point can be unambiguously demonstrated only by examining an individual fiber at high resolution in all three-dimensions, within the physical section (Pratelli et al., 2017). Even when a confocal system with high-power objectives is used, a branching point can be difficult to distinguish from fibers crossing each other at sub-micrometer distances (Janušonis and Detering, 2019; Janušonis et al., 2019). The extent of local sprouting in regeneration also remains an open problem (Hawthorne et al., 2010; Jin et al., 2016), but current evidence suggests that it may not be significant (Jin et al., 2016). The rapidly developing methods of super-resolution microscopy and tissue expansion are well-positioned to provide definitive answers to some of these questions (Janušonis et al., 2019; Wassie et al., 2019).

If serotonergic fiber densities are determined by local arborization, it is unclear how fibers can be restricted to specific "terminal" regions, such as cerebral cortical layers (Linley et al., 2013). Because of their high degree of tortuosity, they are likely to cross over to adjacent areas, suggesting a subtle balance between region-specific branching and a diffusion-like process. Furthermore, the concept of "terminal region" is ambiguous for the serotonergic axons that typically do not form conventional synapses and can release serotonin at virtually any segment of their trajectory, based on *in vivo* and *in vitro* observations of axon varicosities (Benzekhroufa et al., 2009; Gagnon and Parent, 2014; Quentin et al., 2018). Serotonergic neurons also can release serotonin from the soma, dendrites, and growth cones, effectively making their entire membrane surface active (Ivgy-May et al., 1994; Quentin et al., 2018). It should be noted that serotonergic axons may also form conventional synapses (Papadopoulos et al., 1987), but the extent of this "wiring" transmission (Agnati and Fuxe, 2014) is currently unknown and continues to be debated.

In this study, we model individual serotonergic fibers as paths of a stochastic process that reflects their physical properties and show that regional arborization or other local control is not necessary to arrive at a good approximation of the observed fiber densities. Instead, these densities may strongly depend on the geometry of the brain. Our novel approach may offer insights into the self-organization of serotonergic densities in development and may also explain their stability in adulthood, without assuming the permanence of individual fiber trajectories. Consistent with this hypothesis, a recent study has shown that in the adult mouse brain regenerating serotonergic fibers do not follow the pathways left by degenerated fibers but can still restore the layer-specific densities after cortical injury (Jin et al., 2016).

We focus on Fractional Brownian Motion (FBM), a process first described under this name by Mandelbrot and Van Ness in 1968 (Mandelbrot and Van Ness, 1968). FBM and related stochastic processes have emerged as flexible and theoretically rich models in a variety of physical and biological systems. In particular, they have been used to understand the behavior of filamentous objects, such as biopolymer chains and chromosomes (Polovnikov et al., 2018, 2019).

The application of Brownian random walks to polymer chains dates back to the seminal works of Paul Flory, a founder of modern polymer physics and the winner of the 1974 Nobel Prize in Chemistry (Flory, 1986). In this theoretical approach, a fully flexible polymer in a poor solvent can be represented by a discrete random walk, in which each step corresponds to a monomer. However, these models are insufficient to reproduce polymer behavior in good solvents. Self-avoiding walks (SAWs) on lattices may be useful in these systems, but they are difficult to treat analytically and in computer simulations. FBM offers a more powerful theoretical solution, in which the fractal dimension of the path (the polymer chain in space) can be reproduced by a scaling exponent that controls the decay of the long-range correlation of the FBM steps (Chakravarty and Sebastian, 1997; Qian et al., 1998). This fractal dimension is a measure of how often the chain intersects itself (Hu and Nualart, 2005; Polovnikov et al., 2019). The same mathematical framework also allows modeling active motion, for instance, in living biological cells (Reverey et al., 2015). In the polymer language, these trajectories correspond to long chains that can show persistence in their spatial direction.

Theoretically, FBM extends the normal Brownian motion (BM), which for over a century has served as a standard model to describe simple diffusion and other similar processes (e.g., simple polymer dynamics and stock markets). While BM assumes independence between non-overlapping increments, FBM expands this model by allowing non-zero correlations. The sign and strength of the increment correlation is determined by the Hurst index (H), which defines two fundamentally different FBM regimes. If $0 < H < ½$, two neighboring increments are negatively correlated, which produces highly jittery, "anti-persistent" trajectories (also known as "rough paths"). If $½ < H < 1$, two neighboring increments are positively correlated, which produces "persistent" trajectories that tend to maintain their current direction. In precise terms, the correlation between two neighboring increments is given by $2^{2H-1} - 1$ and the mean-square displacement follows the power law $\langle x^2 \rangle \sim n^{2H}$, where n is the number of performed steps. In this framework, BM becomes a special case of FBM, represented by $H = ½$. According to the scaling of the mean-square displacement, which can be (in the number or steps) sublinear (for $0 < H < ½$) or superlinear (for $½ < H < 1$), FBM can be classified as *subdiffusion* or *superdiffusion*, respectively. Since the trajectories of serotonergic fibers are considerably less jittery than BM and have the tendency to maintain their current direction, one can expect to capture their behavior with superdiffusive FBM. In the polymer language, $H < ½$ would increase the tendency to create coiled configurations, whereas $H > ½$ would lead to more stretched configurations. The ballistic limit $H = 1$ would correspond to a fully stretched chain.

In addition, FBM has four convenient properties that make it a natural choice in this context. First, it is a continuous process, which is consistent with the time-continuity of axon growth. Second, it has stationary increments, which informally means that its statistical properties do not change as the process evolves (this assumption is reasonable from the biological perspective). Third, it is a self-similar process, which ensures that the estimation of H does not depend on the discretization grid of experimental observations. Since time-dependent information is difficult to obtain in growing fibers [e.g., with time-lapse imaging in live animals Jin et al., 2016], this property ensures robustness. Fourth, its increments are normally distributed. Assuming that randomness in the fiber trajectory arises from collision-like events in its microenvironment and that each of these events has a small effect on the trajectory, the total effect of these collisions inevitably leads to a normal distribution (by the Central Limit Theorem). Importantly, FBM is the only stochastic process with all of these properties (assuming mean-zero increments).

Although FBM was introduced over a half-century ago, it poses major challenges in theoretical analyses. This is due to the fact that FBM is neither a Markovian process nor a semimartingale (in contrast to BM). A particularly important problem for biological sciences is the behavior of FBM in bounded domains (e.g., in two- or three-dimensional shapes). Reflected BM is well-understood (Ito and McKean, 1965), but it is not until very recently that the first description of the properties of reflected FBM (rFBM) have become available, in one-dimensional domains (Wada and Vojta, 2018; Guggenberger et al., 2019; Wada et al., 2019). The present study is the first application of this theoretical framework to serotonergic fiber distributions, on the whole-brain scale. We determine the steady state distributions of superdiffusive rFBM in constrained brain-like domains and show that they approximate neuroanatomical observations.

MATERIALS AND METHODS

Immunohistochemistry and Imaging

Two adult male mice (C57BL/6J, 8 months of age, The Jackson Laboratory) were deeply anesthetized with a mixture of ketamine (200 mg/kg) and xylazine (20 mg/kg) and perfused transcardially with saline, followed by 4% paraformaldehyde. Their brains were dissected, postfixed in 4% paraformaldehyde overnight at 4°C, cryoprotected in 30% sucrose overnight at 4°C, and sectioned coronally at 40 μm thickness on a freezing microtome. The sections were rinsed in 0.1 M phosphate-buffered saline (PBS, pH 7.2), incubated in 0.3% H_2O_2 in PBS for 20 min to suppress endogenous peroxidase activity, rinsed in PBS (3 times, 5 min each), blocked in 5% normal donkey serum (NDS) and 0.3% Triton X-100 (TX), and incubated in rabbit anti-serotonin transporter (SERT) IgG (1:5000; ImmunoStar, #24330) with 5% NDS and 0.3% TX in PBS for 2.5 days at 4°C on a shaker. The sections were rinsed in PBS (3 times, 10 min each), incubated in biotinylated donkey anti-rabbit IgG (1:1000; Jackson ImmunoResearch, #711-065-152) with 2% NDS and 0.3% TX in PBS, rinsed in PBS (3 times, 10 min each), incubated in the avidin-biotin-peroxidase complex (1:100; Vector Laboratories, #PK-6100), rinsed in PBS (3 times, 10 min each), developed with 3,3'-diaminobenzidine and H_2O_2 using nickel intensification for 5 min (Vector Laboratories, #SK-4100), rinsed in PBS (3 times, 5 min each), mounted onto gelatin/chromium-subbed slides, allowed to air-dry, and coverslipped with Permount. They were imaged on the Zeiss Axio Imager Z1 system with the following

objectives: 1 × (NA = 0.025; used for whole sections), 10 × (NA = 0.45), 20 × (NA = 0.80), and 40 × (NA = 1.30, oil). All procedures have been approved by the UCSB Institutional Animal Care and Use Committee.

Preparation of 2D-Shapes

The outer border, ventricular spaces, and major white matter tracts of imaged sections were hand-traced in Adobe Illustrator CC by an expert trained in neuroanatomy. To reduce the number of contours, major tracts at the edge of the section were left out of the border contour. The contours were imported into Wolfram Mathematica 12. The outer border contour was split along the median (sagittal) symmetry line. The right side of the contour was smoothed with a moving average and reflected on the left side. The same procedure was used for internal contours symmetric with respect to the median line (e.g., the cerebral aqueduct). Inner contours away from the median line (e.g., the fornix) were smoothed on the right side with a moving average and reflected on the left side. Therefore, the final digitized contours (rational-valued arrays of X- and Y-coordinates) were perfectly bilaterally symmetric, compensating for minor sectioning plane deviations and real (minor) brain asymmetries.

The obtained contours were next reformatted for FBM simulations. They were transformed into $N \times 2$ matrices, the rows of which represented consecutive, integer-valued Y-coordinates and the two columns of which represented the leftmost and rightmost X-coordinates of the contour (also integer-valued). To arrive at this format, the original contour coordinates were divided by an integer factor that after rounding produced at least four X-values for each consecutive Y-value, and the minimal and maximal X-values were chosen. Since this procedure effectively reduced the size of the contour, it was enlarged back to its original size by multiplying the integer coordinates by the same factor and filling in the new, empty rows with X-values obtained by linear interpolation between the nearest available X-coordinates. Because this format cannot encode concavities oriented along the Y-axis (e.g., the third ventricle), such concavities were stored as separate inner contours. In the study, all such concavities were centered on the median line, which allowed their easy capture with the maximal X-value left to the line and the minimal X-value right to the line. For the purpose of this study, all inner contours were treated as impenetrable obstacles, irrespective of their physical nature (ventricular spaces, outer border concavities, white matter tracts).

The computer simulations described in detail below were performed on the Frontera supercomputing system (NSF, Texas Advanced Computing Center).

Discrete Reflected FBM

In the simulations, each individual serotonergic fiber was represented as the trajectory of a discrete two-dimensional FBM (Qian, 2003). Consider a random walker starting at position r_0 and moving according to the recursion relation $r_{n+1} = r_n + \xi_n$, where ξ_n is a two-component fractional Gaussian noise. This means that the x and y components of ξ_n are Gaussian random numbers with zero average and variance σ^2, and that each component features long-range correlations between the steps (but the x and y motion are independent of each other). The corresponding covariance function is given by $\langle \xi_{x,m}\, \xi_{x,n+m} \rangle = \langle \xi_{y,m}\, \xi_{y,n+m} \rangle = \frac{1}{2}\, \sigma^2\, [|n+1|^{2H} - 2\,|n|^{2H} + |n-1|^{2H}]$, where H is the Hurst index. The Fourier-filtering method (Makse et al., 1996) was employed to generate these long-range correlated random numbers on the computer (Wada and Vojta, 2018).

If the random walker encounters a boundary (i.e., an outer or inner contour), it is reflected. This reflection can be implemented in several different ways, including a repulsive potential force or a simple reflection condition that either prevents the trajectory from entering the forbidden region or mirrors it at the reflecting wall. Extensive test calculations (Vojta et al., 2020) have demonstrated that the choice of the reflection condition does not affect the resulting probability density for rFBM outside a narrow region (just a few steps wide) at the boundary. The following simulations thus employed the condition that a step that would lead the random walker into the forbidden region was simply not carried out.

Reflected FBM-Paths in Brain-Like Shapes

The goal of the FBM simulations in brain-like shapes was to model the two-dimensional distributions of serotonergic fiber densities at different rostro-caudal levels. The brain is a three-dimensional structure, but two-dimensional simulations are sufficient to approximate the densities, provided that the brain geometry does not change too rapidly in the direction perpendicular to the selected section. Serotonergic fibers do not form fascicles and show no preferred orientation (i.e., they are spatially isotropic) in terminal fields (Janušonis et al., 2019). As the three spatial coordinates in FBM are independent of each other, this implies that the motion perpendicular to the selected section can be neglected [for further information, see Vojta et al. (2020)].

Most simulations were performed with $H = 0.8$, but values between $H = 0.3$ and 0.9 were also studied for comparison. Each dataset consisted of 960 individual fibers (FBM trajectories) starting from random positions inside the shape. Each trajectory had $2^{22} \approx 4$ million steps of size (standard deviation) $\sigma = 1.29\,\mu$m. Sample trajectories are shown in **Figure 1**. The local density (d_s) was determined by counting the total number of random walk segments inside each cell (of linear size 12.9 μm) of a square grid covering the shape. The trajectories were sufficiently long for the relative densities to reach a steady state (i.e., they did not change if the trajectory lengths were increased further).

For comparison between the simulated fiber densities and the densities observed in the actual (immunostained) sections, the simulated densities were scaled to "optical densities" by the (Beer-Lambert law-like) transformation $d_o = 1 - \exp(-k d_s)$, where the attenuation parameter k was chosen such that the mean pixel value in the simulated section matched the mean pixel value in the image of the immunostained section. This transformation removed the dependence on the arbitrarily chosen number and length of the FBM trajectories, and it also realistically constrained density values to a finite interval.

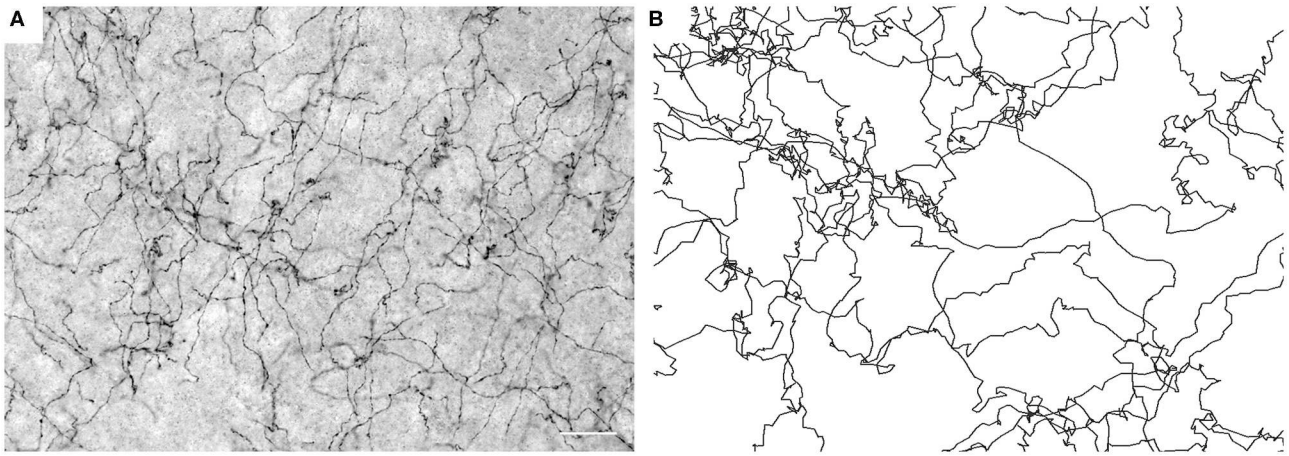

FIGURE 1 | (A) Serotonergic fibers visualized with SERT-immunohistochemistry at high resolution (40 ×, NA = 1.3 oil) in the mouse primary somatosensory cortex (S1). The image represents five focal levels (in a 40 μm-thick coronal section) that have been digitally merged. Scale bar = 20 μm. **(B)** Simulated FBM sample paths ($H = 0.8$).

Reflected FMB-Paths in a Ring With a Varying Curvature

The boundary curvature may strongly affect the accumulation of fibers, which has direct implications for brain neuroanatomy. For example, in the coronal plane the curvature of the cerebral cortex can be high near the sagittal plane, as the outer border turns toward the corpus callosum, and low in the lateral convexities. The cortical gyri and sulci of the human brain create additional variability of local curvature, which can further modulate the accumulation of serotonergic fibers. Therefore, a *qualitative* relationship between boundary curvatures and fiber densities can allow theoretical predictions in neuroanatomy and may obviate the need for detailed computer simulations for each individual shape. However, caution should be exercised in shapes with complex geometry because superdiffusive rFBM is not a local process, due to its long-range correlations. Therefore, its properties can be potentially affected by the geometry of the entire domain. A systematic quantitative analysis of rFBM in geometrically idealized spatial domains (in two and three dimensions) falls outside the scope of the present analysis and is presented elsewhere (Vojta et al., 2020).

In order to gain an insight into how the density of FBM-trajectories varies as a function of H and the contour curvature, especially in neuroanatomically-relevant shapes, we investigated a ring-like shape bounded by the contours $R_{outer} = 100(1 + 0.5\cos^2(4\varphi))$ and $R_{inner} = 50$ (defined in polar coordinates, where $\varphi \in [0, 2\pi)$). This shape can represent a cross section through an abstracted vertebrate brain. Vertebrate brains develop from the neural tube and remain topologically tube-like in adulthood, where the "hollow" inside of the tube is the ventricular space filled with the cerebrospinal fluid (CSF).

These simulations were performed with $H = 0.3$, 0.5, and 0.8. Each data set consisted of 100 individual fibers (FBM trajectories), starting from random positions inside the shape. Each trajectory had $2^{20} \approx 1$ million steps of size (standard deviation) $\sigma = 0.4$. In the subdiffusive case ($H = 0.3$), it was necessary to increase the step size to 1 to ensure that the simulations reached the stationary regime. As before, the local fiber density (d_s) was determined by counting the total number of random walk segments inside each cell (of linear size 1) of a square grid covering the shape.

Simulation of Reflected FBM-Paths in a 2D-Disk With Crowding

The brain tissue is a highly crowded space, but little quantitative information is available about the extent and geometry of the extracellular space in different brain regions (Hrabetova et al., 2018). In order to assess the sensitivity of our results to crowding effects, we performed FBM simulations in a large disk of radius $R = 100$, filled with 1,013 small disk-shaped obstacles of radius $R_{obs} = 2$. The obstacles were located within a circle of radius 90, leaving an outer ring of width 10 unoccupied. This shape was intended to represent a highly abstracted neocortical region, with the empty outer rim representing cortical layer I. In adulthood, this layer is nearly devoid of neuron somata (but contains many dendrites and axons that for simplicity were ignored in the simulation).

The simulations were performed with $H = 0.8$. Each dataset consisted of 192 individual fibers (FBM trajectories) starting from random positions inside the shape. Each trajectory had $2^{25} \approx 34$ million steps of size (standard deviation) $\sigma = 0.4$. As before, the local fiber density (d_s) was determined by counting the total number of random walk segments inside each cell (of linear size 0.5) of a square grid covering the shape.

RESULTS

Serotonergic fiber densities are typically described with regard to specific neuroanatomical brain regions, with the assumption that they reflect the region's functional demands and are supported by

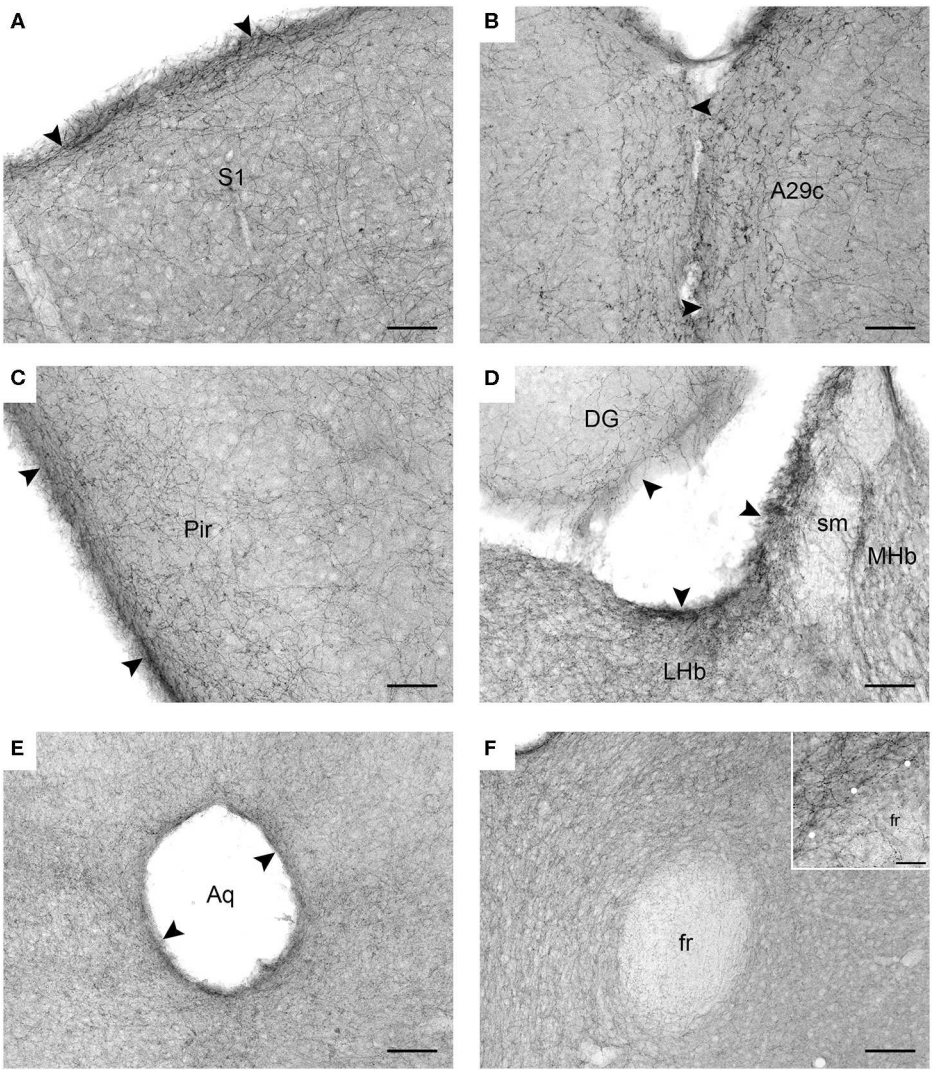

FIGURE 2 | Serotonergic fibers visualized with SERT-immunohistochemistry in coronal sections of the mouse brain. **(A)** The primary somatosensory cortex (S1). **(B)** The cingulate cortex, area 29c (A29c). **(C)** The piriform cortex (Pir). **(D)** The habenula (MHb, medial habenula; LHb, lateral labenula; sm, stria medullaris) and the dentate gyrus (DG) of the hippocampus. **(E)** The periaqueductal region in the mesencephalon (Aq, cerebral aqueduct). **(F)** The region around the fasciculus retroflexus (fr) in the diencephalon. The inset is a high-magnification image of the border (marked with circles) between the fasciculus retroflexus and the surrounding brain tissue. The arrowheads indicate increased fiber densities. Scale bars = 50 μm in **(A–D)**, 100 μm in **(E,F)**, and 20 μm in the inset of F.

local biological factors. Our immunolabeling results (**Figure 2**) are consistent with the previously reported density maps in the rat and hamster brains (Steinbusch, 1981; Morin and Meyer-Bernstein, 1999). In particular, they show high fiber densities in the superficial layers of the cerebral cortex, consistent with observations in the rat and ferret (Voigt and de Lima, 1991a; Linley et al., 2013).

In contrast to previous interpretations, we suggest that these data indicate a *general* tendency of serotonergic fibers to accumulate near the borders of neural tissue (at the pial or ependymal surfaces) (**Figures 2A–E**). This observation has been previously made in some brain regions [e.g., in the hamster thalamus Morin and Meyer-Bernstein, 1999] but to our knowledge has never been extended to the entire brain. We next show that this tendency is consistent with the behavior of rFBM in the superdiffusion regime.

Supercomputing FBM simulations (with $H = 0.8$) were performed in four two-dimensional shapes that closely approximated the shape of actual coronal sections. Since major white matter tracts may act as forbidden regions (**Figure 2F**), we included them as "obstacles" in the simulations. Qualitatively, FBM sample paths closely resembled the trajectories of individual serotonergic fibers (**Figure 1**).

The comparisons between the simulated and actual fiber densities are shown in **Figures 3–6**. These densities showed similar increases at the outer brain border, irrespective of the

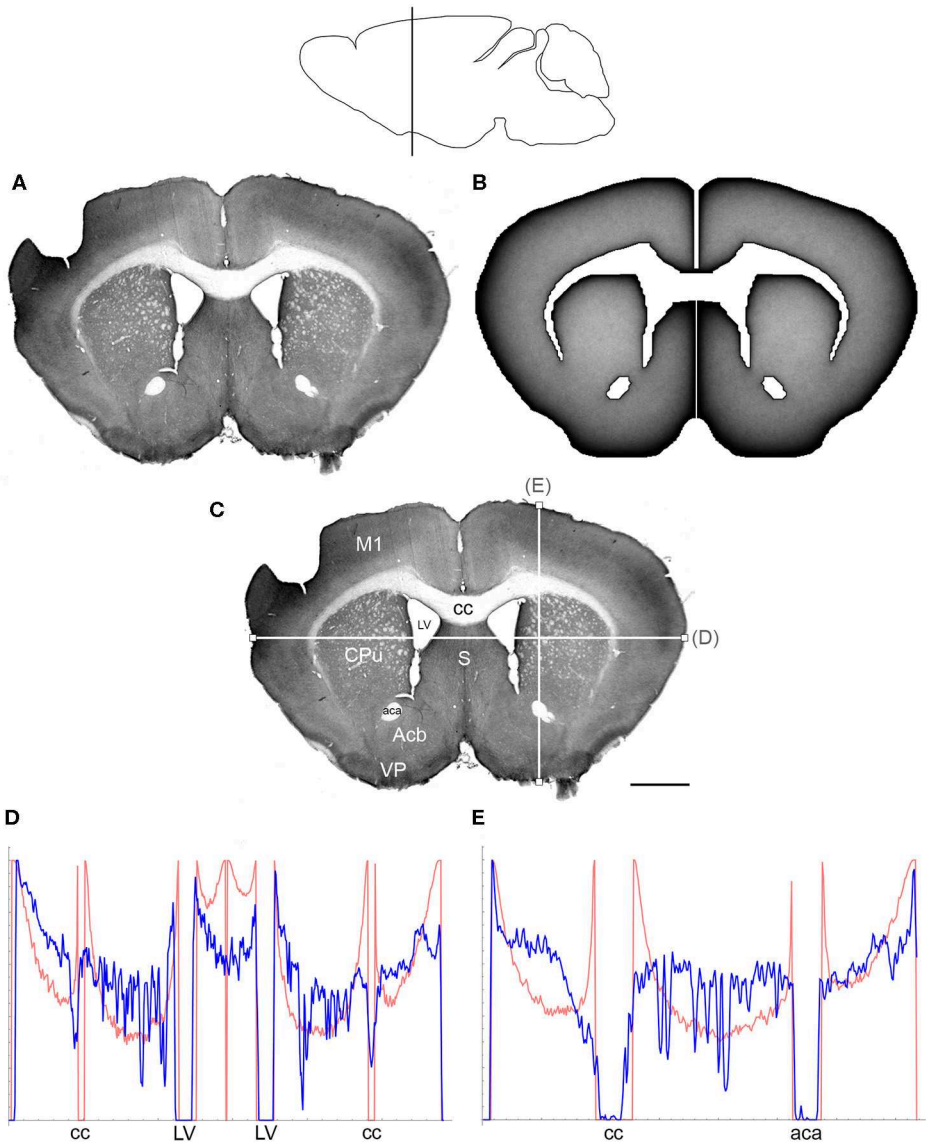

FIGURE 3 | (A) The density of serotonergic fibers (immunostained for SERT) in a coronal section of the mouse telencephalon. **(B)** The equilibrium density of simulated fibers performing FBM-walks in the same 2D-shape ($H = 0.8$; 960 fibers). In both **(A,B)**, darker regions represent higher densities. **(C)** The main neuroanatomical structures and two density cuts, plotted in **(D,E)**. The end points of the plotted segments are marked with small squares. Acb, nucleus accumbens; aca, anterior commissure; cc, corpus callosum; CPu, caudate/putamen; LV, lateral ventricle; M1, primary motor cortex; S, septum; VP, ventral pallidum. Scale bar = 1 mm. In **(D,E)**, the experimental and simulated densities are shown in blue and red, respectively. The simulated densities have been transformed to "optical densities," as described in Materials and Methods (the Y-axis ranges from 0 to 1). The rostro-caudal level in the sagittal view is shown at the top.

rostro-caudal level. A similar trend was observed around the ventricles. Interestingly, high simulated fiber densities were obtained in some neuroanatomically-defined brain regions, such as the lateral geniculate nucleus and the hypothalamus (**Figures 4**, **5**), even though these regions were not specifically modeled and the increase was induced purely by the contour geometry. It should be noted that both of these regions have a convex border with a relatively high curvature. We investigated this potential association using a simple shape and a range of H values (**Figure 7**), which further supported this conjecture. It leads to verifiable predictions of how serotonergic fiber densities can vary across brain regions and may also support comparative neuroanatomy, where differences in the fiber densities across mammalian species may be caused, at least in part, by differences in the brain shapes (despite the highly similar neuroanatomical plans).

At the outer border, a considerable mismatch was observed between some *gradients* of densities (e.g., **Figure 4E**). Considering the neuroanatomical simplicity of the simulated shape (e.g., it contained no "cells"), this result is not surprising.

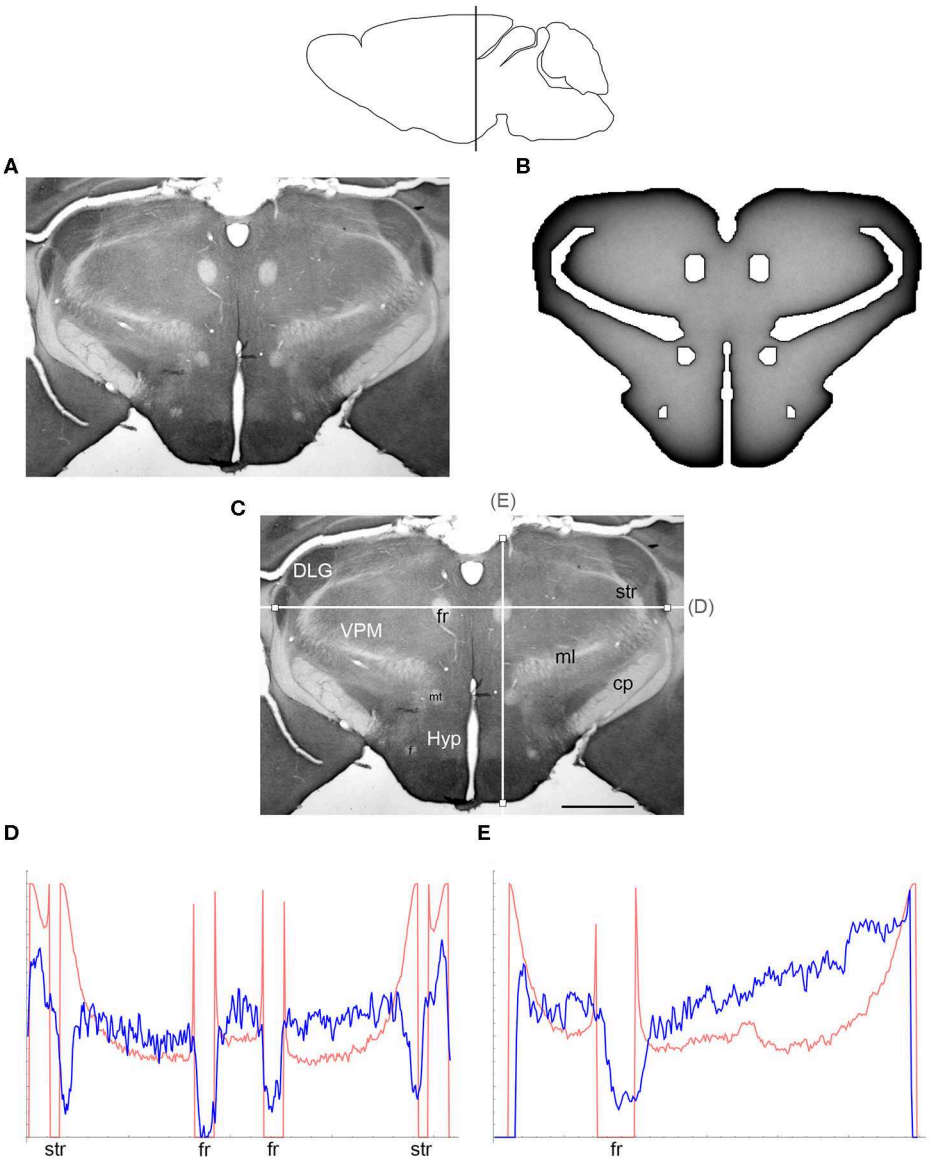

FIGURE 4 | (A) The density of serotonergic fibers (immunostained for SERT) in a coronal section of the mouse rostral diencephalon. **(B)** The equilibrium density of simulated fibers performing FBM-walks in the same 2D-shape ($H = 0.8$; 960 fibers). In both **(A,B)**, darker regions represent higher densities. **(C)** The main neuroanatomical structures and two density cuts, plotted in **(D,E)**. The end points of the plotted segments are marked with small squares. DLG, dorsal lateral geniculate nucleus; cp, cerebral peduncle; f, fornix; fr, fasciculus retroflexus; Hyp, hypothalamus; ml, medial lemniscus; mt, mammillothalamic tract; str, superior thalamic radiation; VPM, ventral posteromedial nucleus. Scale bar = 1 mm. In **(D,E)**, the experimental and simulated densities are shown in blue and red, respectively. The simulated densities have been transformed to "optical densities," as described in Materials and Methods (the Y-axis ranges from 0 to 1). The rostro-caudal level in the sagittal view is shown at the top.

Also, the simulated gradients depend on the value of H and the attenuation parameter of "optical transformation" (e.g., they can be made less steep, with an effect on the overall density intensity). Matching the simulated and actual gradients precisely is difficult because the true fiber density cannot be determined in immunostained sections without tracing every fiber. Also, many non-linear effects can take place between the section and the image sensor, even at optimal illumination settings.

The strongest discrepancy was found between the strong density spikes around white matter tracts in the simulations (where the tracts were modeled as impenetrable "obstacles") and the virtual absence of such spikes in the actual (immunostained) sections. This suggests that these tracts cannot be modeled as "hard" obstacles and that other FBM-reflection models may reflect their properties more accurately.

Since a fixed H was used in the simulations, we investigated the sensitivity of the obtained results to a range of H

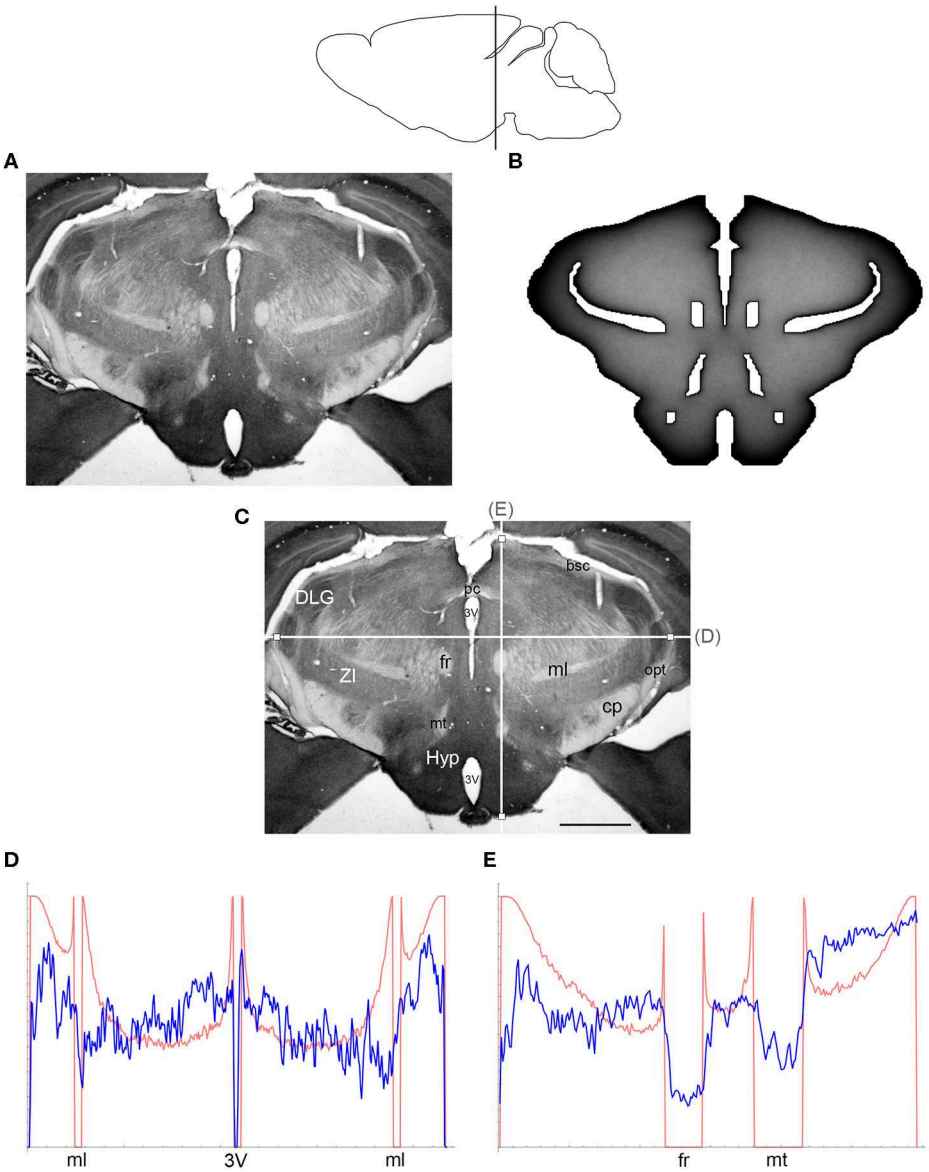

FIGURE 5 | (A) The density of serotonergic fibers (immunostained for SERT) in a coronal section of the mouse caudal diencephalon. **(B)** The equilibrium density of simulated fibers performing FBM-walks in the same 2D-shape ($H = 0.8$; 960 fibers). In both **(A,B)**, darker regions represent higher densities. **(C)** The main neuroanatomical structures and two density cuts, plotted in **(D,E)**. The end points of the plotted segments are marked with small squares. 3V, third ventricle; bsc, brachium of the superior colliculus; cp, cerebral peduncle; DLG, dorsal lateral geniculate nucleus; fr, fasciculus retroflexus; Hyp, hypothalamus; ml, medial lemniscus; mt, mammillothalamic tract; opt, optic tract; pc, posterior commissure; ZI, zona incerta. Scale bar = 1 mm. In **(D,E)**, the experimental and simulated densities are shown in blue and red, respectively. The simulated densities have been transformed to "optical densities," as described in Materials and Methods (the Y-axis ranges from 0 to 1). The rostro-caudal level in the sagittal view is shown at the top.

values (**Figures 6**, **7**). The density distribution patterns varied dramatically across the three diffusion regimes, as anticipated (Wada and Vojta, 2018). However, they were robust within the superdiffusion regime, suggesting that the results can be safely generalized to other H values, beyond the one that was used in the simulations (0.8).

Finally, we examined the sensitivity of the results to cell packing. Heterogeneous crowding is an essential property of neural tissue (Hrabetova et al., 2018), but currently little is known about the variability of the crowding density in brain regions. In the present study, we used a relatively simple model and tested whether the accumulation of fibers at borders could be reversed by many cell-like obstacles that can potentially retard or trap fibers in the interior region of the bounded domain (**Figure 8**). Despite the presence of many obstacle surfaces, the simulation produced only thin, high-density layers around the obstacles, with no major effect on the overall density distribution. This result supports the robustness of our findings

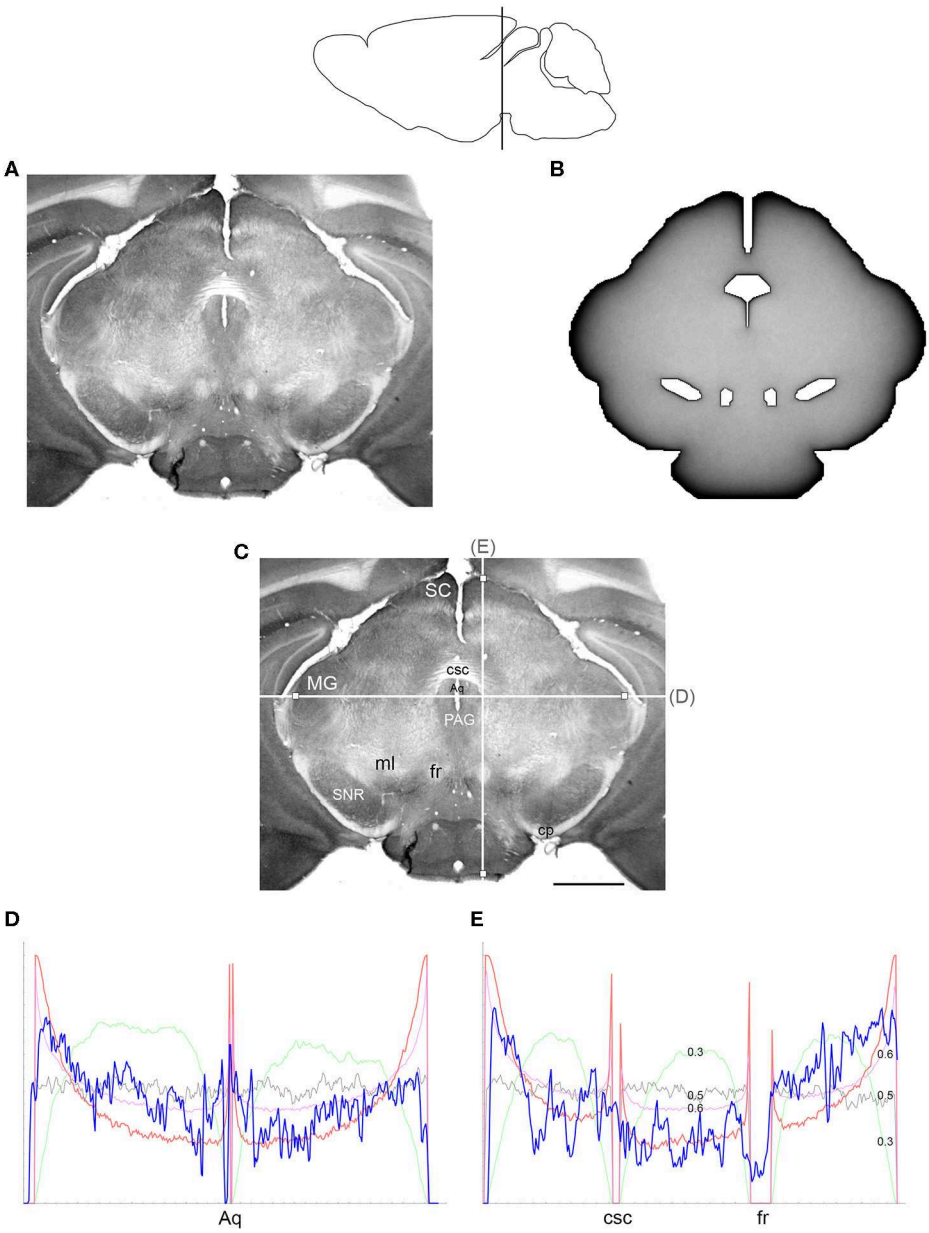

FIGURE 6 | (A) The density of serotonergic fibers (immunostained for SERT) in a coronal section of the mouse mesencephalon. **(B)** The equilibrium density of simulated fibers performing FBM-walks in the same 2D-shape ($H = 0.8$; 960 fibers). In both **(A,B)**, darker regions represent higher densities. **(C)** The main neuroanatomical structures and two density cuts, plotted in **(D,E)**. The end points of the plotted segments are marked with small squares. Aq, aqueduct; cp, cerebral peduncle; csc, commissure of the superior colliculus; fr, fasciculus retroflexus; MG, medial geniculate nucleus; ml, medial lemniscus; PAG, periaqueductal gray; SC, superior colliculus; SNR, substantia nigra pars reticulata. Scale bar = 1 mm. In **(D,E)**, the experimental and simulated densities are shown in blue and red, respectively. For comparison, this figure shows simulated densities obtained with three other H values [thin lines: $H = 0.3$ (green), $H = 0.5$ (gray), $H = 0.6$ (pink)]. Note that $H = 0.3$ produces drastically reduced densities at the border and obstacles and that $H = 0.5$ produces densities that remain constant along the cut, with no change at the border or obstacles. Neither of these results is consistent with the real densities, in contrast to H values in the superdiffusion regime ($H > 0.5$). For visualization of these patterns in two dimensions, see **Figure 7**. All simulated densities have been transformed to "optical" densities, as described in Materials and Methods (the Y-axis ranges from 0 to 1). The rostro-caudal level in the sagittal view is shown at the top.

and is generally consistent with experimental observations in the cerebral cortex, where the high density of serotonergic fibers may not be restricted to cortical layer I (which is virtually devoid of neuron somata) and may also include layers II-III (which contain densely packed neurons) (Voigt and de Lima, 1991b). However, this finding should not be overgeneralized without future studies of the interactions among different shape geometries, heterogeneous crowding, and the physical constraints of the

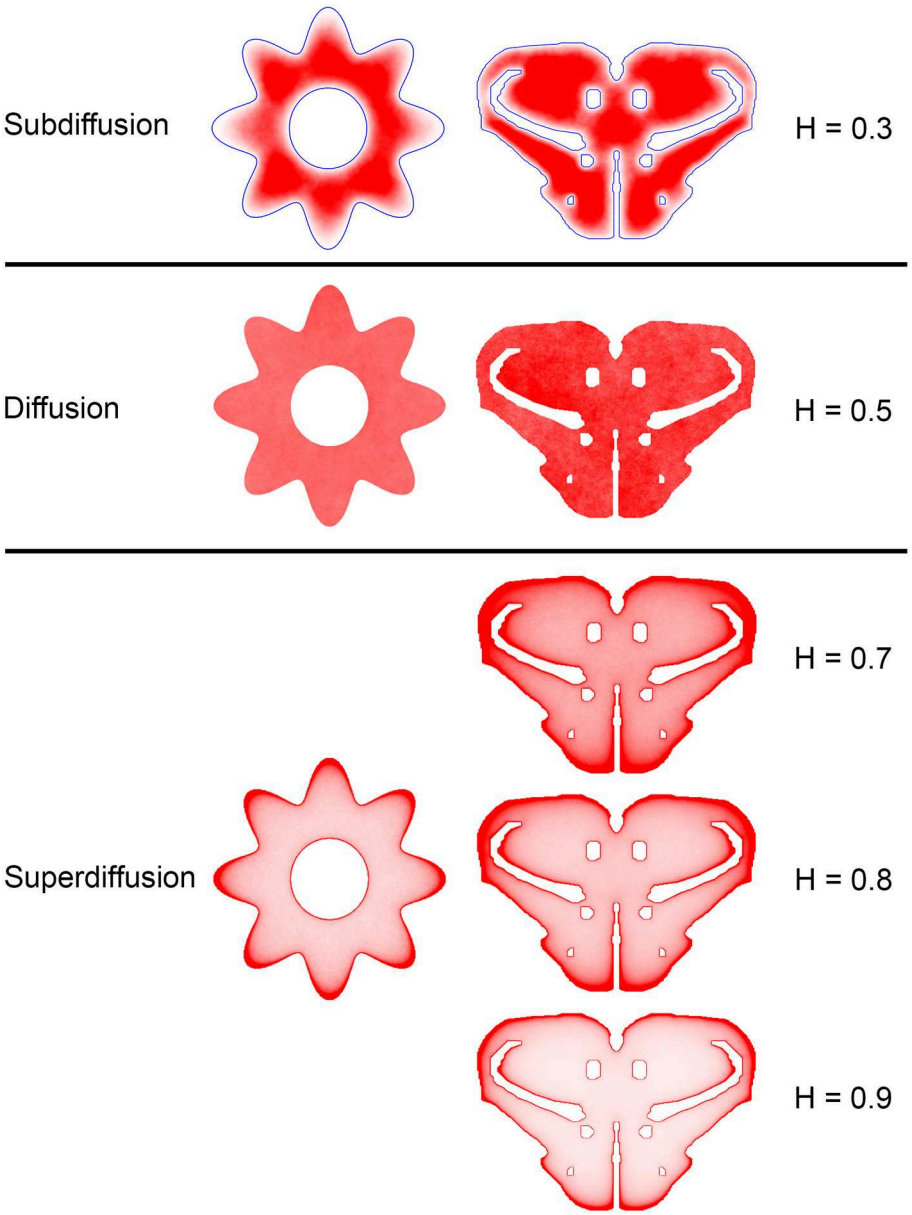

FIGURE 7 | The dependence of equilibrium fiber density on the Hurst index (*H*). Two shapes were used: the cross-section of a cylinder with an outer boundary that has a non-constant curvature (representing an abstracted neural tube; left column) and a shape representing the coronal section of the mouse rostral diencephalon (also used in **Figure 4**; right column). The contour borders are shown in blue in the subdiffusion regime ($H = 0.3$); no borders are marked in the other regimes (but they are clearly visible due to the accumulation of fibers). The simulated densities are plotted with no transformation. One hundred fibers were used for the abstracted shape in the left column, while 960 trajectories were used for the coronal section in the right column. The simulated densities are mapped linearly to the color scale, such that darker regions represent higher densities.

serotonergic fibers (e.g., the simulated trajectories in our study had no physical width).

DISCUSSION

We introduce a novel approach to the self-organization of serotonergic fibers in the brain and demonstrate that rFBM can replicate the behavior of these fibers at "hard" borders (such as the pial and ependymal surfaces of the brain). In contrast, major white matter tracts do not appear to significantly "reflect" serotonergic fibers, even though they constrain their trajectories as obstacles. This phenomenon may be due to the fact that many individual fibers can penetrate (and perhaps traverse) these tracts (**Figure 2F**). The presence of serotonergic fibers in some major pathways, such as the fasciculus retroflexus, has been noted

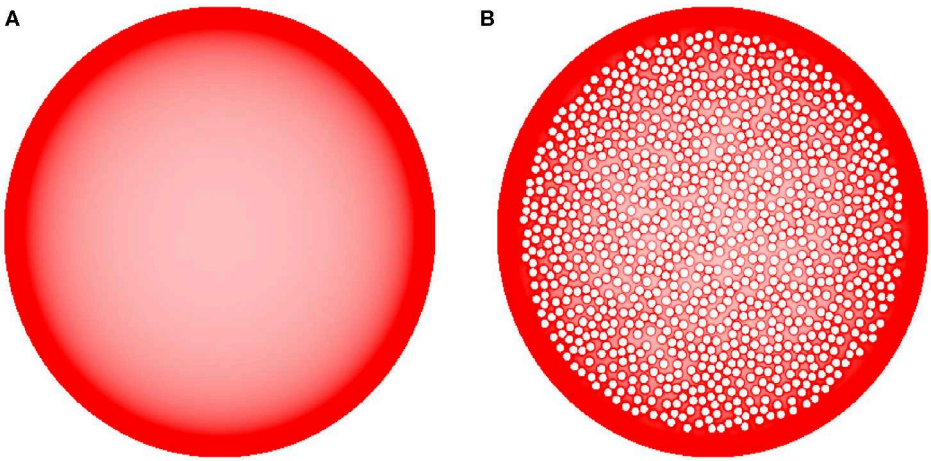

FIGURE 8 | The sensitivity of equilibrium fiber density **(A)** to densely packed internal obstacles **(B)** (represented by the small, impenetrable white discs inside the large shape). Even though fibers accumulate around the border of each individual obstacle, the obstacles have only a minor effect on the overall density distribution, especially near the outer border (where the density is higher). In the simulations, $H = 0.8$ and 192 fibers were used. The simulated densities are mapped linearly to the color scale, such that darker regions represent higher densities.

in early neuroanatomical studies (Lidov and Molliver, 1982). These phenomena can be included in our model by treating the boundaries of white matter tracts differently from other boundaries, such as modeling them as soft repulsive potentials (Vojta et al., 2020).

Our model does not include the rich brain architecture and biological signals that may act on serotonergic fibers. While the simulated fiber densities are a good approximation of the actual fiber densities, some of these factors may be important for more accurate predictions. We briefly mention some of them.

We used two-dimensional shapes, even though the brain is a three-dimensional object. As explained in the section Reflected FBM-paths in Brain-like Shapes, the reduction to two dimensions unlikely produced major distortions. However, some discrepancies between the actual and simulated densities may be due to the geometry of coronal sections just rostral or caudal to the selected section, with some fibers "spilling into" or "leaking from" the examined section. This problem may be particularly significant where the geometry of ventricular spaces rapidly changes in the rostro-caudal direction (e.g., ventricular spaces "fuse" or "separate" in two-dimensional projections). Conceptually, the computational approach can be easily extended to three dimensions, provided a fully three-dimensional model of the brain geometry is available. However, the numerical effort for such simulations would be significantly higher. Specifically, it would require the parametrization of a two-dimensional manifold (the boundary), leading to increased computational costs close to the boundary (where the decision has to be made whether the walker is in the allowed region or not). One potential solution is using two-dimensional brain shapes obtained in two or three perpendicular planes (e.g., coronal, sagittal, axial), which may allow a more accurate reconstruction of densities in all three dimensions.

The brain can be viewed as a highly heterogeneous material. It is densely packed with cells, their processes, microvasculature, and other elements that serotonergic fibers cannot penetrate. Axons can be assembled into major white matter tracts which may require detours; for example, the number of axons in the primate corpus callosum can be on the order of 10^7-10^8 (Doty, 2007). The fine structure of the extracellular space (ECS) in the brain requires state-of-the-art experimental methods and currently is an active area of research (Nicholson and Hrabetova, 2017; Hrabetova et al., 2018). The stochastic geometry of the ECS may impose a particular covariance structure on traveling fibers (or a "memory effect") (Morgado et al., 2002; Bénichou et al., 2013), possibly in a region-specific manner. Some of this geometry can be modeled with sphere packing (to represent cell bodies) (Picka, 2012). Similar problems arise in intracellular environments (Smith et al., 2017). In addition to the "hard" geometry, the stochastic properties of fibers may be affected by the viscoelastic properties of their environment (Cherstvy et al., 2019). In the case of FBM, it may lead to different H values in different brain regions, with implications for local fiber densities. Conversely, the estimated H values of fibers can be potentially used to obtain information about the structure and integrity of the ECS, with possible biomedical applications.

Since individual serotonergic fibers can be visualized, estimates of H can be obtained from experimental data. However, it requires overcoming a few technical challenges, which are a focus of our larger research program (in the present study, we assumed $H = 0.8$ and demonstrated the robustness of the main results). For example, tracing a single fiber in immunostained brain sections (e.g., 40 μm in thickness) is difficult because most fibers exit the section in the Z-direction, before advancing substantially in the X-Y plane (Janušonis et al., 2019). Modern tissue clearing methods and light-sheet microscopy allow direct

3D-imaging with no sectioning (Mano et al., 2018; Hillman et al., 2019), but light-sheet microscopy is only now approaching the sub-micrometer resolution needed for the imaging of individual serotonergic fibers (Chakraborty et al., 2019). The high fiber densities in most brain regions presents another problem, where single-fiber tracing has to be performed in the presence of interfering signals, such as other fibers and potential branching points. Advances in computer image analysis (Kayasandik et al., 2018; Falk et al., 2019), combined with transgenic technologies that allow labeling individual neurons and their processes with unique combinations of fluorophores [such as Brainbow 3.2 Cai et al., 2013], are well positioned to advance these efforts.

Our model assumes no self-avoidance or biological feedback signals that depend on fiber density. Several factors have been reported to control the growth of distribution of serotonergic fibers, but this research has been heavily influenced by the notion that orderly distribution cannot be achieved without tight biological control [see, e.g., Chen et al. (2017)]. The simulation results provide evidence that a considerable degree of self-organization can be achieved with simple assumptions, but it does not rule out these factors. Early studies have found that S100β, a biologically active protein, may promote the development of serotonergic fibers. Interestingly, S100β can be released from astrocytes, in response of activation of serotonin 5-HT_{1A} receptors, suggesting a positive feedback loop (Whitaker-Azmitia, 2001). Also, the absence of brain serotonin synthesis alters the development of normal fiber densities, but this effect appears to be strongly region-dependent (Migliarini et al., 2013). It may be mediated by the brain-derived neurotrophic factor (BDNF) which has long been implicated in the growth of serotonergic fibers (Mamounas et al., 1995; Migliarini et al., 2013). It remains unclear how serotonin affects fiber densities under physiological conditions because a different genetic model (with a less severe reduction of brain serotonin levels) has failed to reproduce these effects (Donovan et al., 2019). Recently, protocadherin-αc2 has been strongly implicated in the distribution of serotonergic fibers, through homophilic interaction between individual axons (Katori et al., 2009, 2017; Chen et al., 2017). Intriguingly, protocadherin-α mutants show pronounced increases in the fiber densities in layer I of the primary motor cortex and in the lacunosum-moleculare layer of the hippocampus, both of which are at the border of their respective brain regions. This study also has noted that in some other regions the "distribution of serotonin axonal terminals was [. . .] dense at the periphery of each region but sparse in the center" (Katori et al., 2009). Our results suggest that this experimental result may reflect either a more pronounced rFBM behavior (in the absence of fiber interaction) or an rFBM with a higher H (induced by the genetic mutation).

It is important to note that FBM is not the only mathematical model that allows superdiffusion (or anomalous diffusion, more generally). Superdiffusion can also be modeled with Lévy flights or with continuous-time random walks (CTRWs) that have a heavy-tailed displacement probability density (Codling et al., 2008; Schulz et al., 2013; Metzler et al., 2014). These two processes allow large, instantaneous (spatially discontinuous) jumps from one location to another, but CTRWs are additionally parametrized by the waiting-time probability density (therefore, a Lévy flight can be viewed as a special CTRW). They are highly appropriate for some physical processes [e.g., the dynamics of molecular complexes jumping from one segment of a polymer to another, facilitated by folding-induced physical proximity Lomholt et al., 2005], but they do not well-represent axon growth where instantaneous jumps are not biologically realistic. Also, the trajectories of growing axons are likely to show long-range temporal correlations that are an inherent property of FBM (in addition to other useful properties reviewed in the Introduction; here we assume $H \neq 1/2$). It should be noted that the long-range correlations in FBM extend to arbitrarily large distances (Biagini et al., 2010), which may exceed biological reality, but a possible theoretical refinement may be provided by stochastic processes in which the long-range correlations are cut off at a large but finite distance (Molina-Garcia et al., 2018). In addition, branching FBM-like processes may offer insights into how the bifurcation or arborization of serotonergic fibers can affect their steady state distribution. Direct simulations of macromolecular dynamics in confined domains can further enrich these studies; for example, in some simulations particles near the wall tend to stay near the wall (Chow and Skolnick, 2015), which may explain the tendency of serotonergic fibers to orient parallel to the edge immediately below the pia (for depths up to 25–50 μm; **Figure 2**). Finally, it has been recently demonstrated (Vojta et al., 2019) that the increased density close to a boundary arises from the non-equilibrium nature of FBM. A similar anomalous diffusion process in thermal equilibrium, modeled by the fractional Langevin equation, does not lead to accumulation at the boundary. This property of FBM is consistent with the active growth of the serotonergic fibers.

In summary, the present study demonstrates that FBM offers a promising theoretical framework for the modeling of serotonergic fibers. Since serotonin-releasing fibers are a part of the larger ascending reticular activating system, which releases other major neurotransmitters and has widespread projections, this framework may also be useful in advancing the understanding of other stochastic axon systems in vertebrate brains.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

ETHICS STATEMENT

The animal study was reviewed and approved by UCSB Institutional Animal Care and Use Committee.

AUTHOR CONTRIBUTIONS

SJ proposed the hypothesis of serotonergic fibers as paths of a stochastic process, performed the immunostaining, prepared the 2D-shape matrices for simulations, and wrote the first draft of the manuscript. TV guided the computational analyses of

reflected FBM and performed all supercomputing simulations. ND suggested FBM as a potential model that allows scale-invariance and made other theoretical contributions. RM led the development of the model within the theoretical framework of anomalous diffusion processes. All authors are the Principal Investigators of their respective research programs.

FUNDING

This research was supported by the National Science Foundation (grants #1822517 and #1921515 to SJ and ND), the National Institute of Mental Health (grant #MH117488 to SJ and ND), the California NanoSystems Institute (Challenge grants to SJ and ND), the Research Corporation for Science Advancement (a Cottrell SEED Award to TV), and the German Research Foundation (DFG grant #ME 1535/7-1 to RM). RM acknowledges support from the Foundation of Polish Science through an Alexander von Humboldt Polish Honorary Research Scholarship.

ACKNOWLEDGMENTS

We thank Kasie Mays (UCSB) for the assistance with tissue preparation.

REFERENCES

Adori, C., Low, P., Ando, R. D., Gutknecht, L., Pap, D., Truszka, F., et al. (2011). Ultrastructural characterization of tryptophan hydroxylase 2-specific cortical serotonergic fibers and dorsal raphe neuronal cell bodies after MDMA treatment in rat. *Psychopharm.* 213, 377–391.

Agnati, L. F., and Fuxe, K. (2014). Extracellular-vesicle type of volume transmission and tunnelling-nanotube type of wiring transmission add a new dimension to brain neuro-glial networks. *Philos. Trans. R Soc. Lond. B Biol. Sci.* 369:20130505. doi: 10.1098/rstb.2013.0505

Azmitia, E. C., Singh, J. S., and Whitaker-Azmitia, P. M. (2011). Increased serotonin axons (immunoreactive to 5-HT transporter) in postmortem brains from young autism donors. *Neuropharmacology* 60, 1347–1354. doi: 10.1016/j.neuropharm.2011.02.002

Bénichou, O., Bodrova, A., Chakraborty, D., Illien, P., Law, A., Mejía-Monasterio, C., et al. (2013). Geometry-induced superdiffusion in driven crowded systems. *Phys. Rev. Lett.* 111:260601. doi: 10.1103/PhysRevLett.111.260601

Benzekhroufa, K., Liu, B., Tang, F., Teschemacher, A. G., and Kasparov, S. (2009). Adenoviral vectors for highly selective gene expression in central serotonergic neurons reveal quantal characteristics of serotonin release in the rat brain. *BMC Biotechnol.* 9:23. doi: 10.1186/1472-6750-9-23

Biagini, F., Hu, Y., Oksendal, B., and Zhang, T. (2010). *Stochastic Calculus for Fractional Brownian Motion and Applications.* London: Springer.

Cai, D., Cohen, K. B., Luo, T., Lichtman, J. W., and Sanes, J. R. (2013). Improved tools for the Brainbow toolbox. *Nat. Methods* 10, 540–547. doi: 10.1038/nmeth.2450

Cardozo Pinto, D. F., Yang, H., Pollak Dorocic, I., de Jong, J. W., Han, V. J., Peck, J. R., et al. (2019). Characterization of transgenic mouse models targeting neuromodulatory systems reveals organizational principles of the dorsal raphe. *Nat. Commun.* 10:4633. doi: 10.1038/s41467-019-12392-2

Carrera, I., Molist, P., Anadon, R., and Rodriguez-Moldes, I. (2008). Development of the serotoninergic system in the central nervous system of a shark, the lesser spotted dogfish *Scyliorhinus canicula*. *J. Comp. Neurol.* 511, 804–831. doi: 10.1002/cne.21857

Chakraborty, T., Driscoll, M. K., Jeffery, E., Murphy, M. M., Roudot, P., Chang, B. J., et al. (2019). Light-sheet microscopy of cleared tissues with isotropic, subcellular resolution. *Nat. Methods* 16, 1109–1113. doi: 10.1038/s41592-019-0615-4

Chakravarty, N., and Sebastian, K. L. (1997). Fractional Brownian motion models for polymers. *Chem. Phys. Lett.* 267, 9–13. doi: 10.1016/S0009-26149700075-4

Chen, W. V., Nwakeze, C. L., Denny, C. A., O'Keeffe, S., Rieger, M. A., Mountoufaris, G., et al. (2017). Pcdhαc2 is required for axonal tiling and assembly of serotonergic circuitries in mice. *Science* 356, 406–411. doi: 10.1126/science.aal3231

Cherstvy, A. G., Thapa, S., Wagner, C. E., and Metzler, R. (2019). Non-Gaussian, non-ergodic, and non-Fickian diffusion of tracers in mucin hydrogels. *Soft. Matter.* 15, 2526–2551. doi: 10.1039/C8SM02096E

Chow, E., and Skolnick, J. (2015). Effects of confinement on models of intracellular macromolecular dynamics. *Proc. Natl. Acad. Sci. U.S.A.* 112, 14846–14851. doi: 10.1073/pnas.1514757112

Codling, E. A., Plank, M. J., and Benhamou, S. (2008). Random walk models in biology. *J. R Soc. Interface* 5, 813–834. doi: 10.1098/rsif.2008.0014

Donovan, L. J., Spencer, W. C., Kitt, M. M., Eastman, B. A., Lobur, K. J., Jiao, K., et al. (2019). Lmx1b is required at multiple stages to build expansive serotonergic axon architectures. *Elife* 8:e48788. doi: 10.7554/eLife.48788.028

Doty, R. W. (2007). "Cortical commissural connections in primates," in *Evolution of Nervous Systems: A Comprehensive Reference,* eds J. H. Kaas and T. M. Preuss (New York, NY: Academic Press), 277–289. doi: 10.1016/B0-12-370878-8/00011-2

Dougherty, S. E., Kajstura, T. J., Jin, Y., Chan-Cortes, M. H., Kota, A., and Linden, D. J. (2019). Catecholaminergic axons in the neocortex of adult mice regrow following brain injury. *Exp. Neurol.* 323:113089. doi: 10.1101/737072

Falk, T., Mai, D., Bensch, R., Cicek, O., Abdulkadir, A., Marrakchi, Y., et al. (2019). U-Net: deep learning for cell counting, detection, and morphometry. *Nat. Methods* 16, 67–70. doi: 10.1038/s41592-018-0261-2

Flory, P. (1986). Development of concepts in polymer science. A half century in retrospect. *Makromol. chem. Macromol. Symposia* 1, 5–16. doi: 10.1002/masy.19860010103

Gagnon, D., and Parent, M. (2014). Distribution of VGLUT3 in highly collateralized axons from the rat dorsal raphe nucleus as revealed by single-neuron reconstructions. *PLoS ONE* 9:e87709. doi: 10.1371/journal.pone.0087709

Guggenberger, T., Pagnini, G., Vojta, T., and Metzler, R. (2019). Fractional Brownian motion in a finite interval: correlations effect depletion or accretion zones of particles near boundaries. *New J. Phys.* 21:022002. doi: 10.1088/1367-2630/ab075f

Hawthorne, A. L., Hu, H., Kundu, B., Steinmetz, M. P., Wylie, C. J., Deneris, E. S., et al. (2011). The unusual response of serotonergic neurons after CNS injury: lack of axonal dieback and enhanced sprouting within the inhibitory environment of the glial scar. *J. Neurosci.* 31, 5605–5616. doi: 10.1523/JNEUROSCI.6663-10.2011

Hawthorne, A. L., Wylie, C. J., Landmesser, L. T., Deneris, E. S., and Silver, J. (2010). Serotonergic neurons migrate radially through the neuroepithelium by dynamin-mediated somal translocation. *J. Neurosci.* 30, 420–430. doi: 10.1523/JNEUROSCI.2333-09.2010

Hendricks, T., Francis, N., Fyodorov, D., and Deneris, E. S. (1999). The ETS domain factor Pet-1 is an early and precise marker of central serotonin neurons and interacts with a conserved element in serotonergic genes. *J. Neurosci.* 19, 10348–10356. doi: 10.1523/JNEUROSCI.19-23-10348.1999

Hillman, E. M. C., Voleti, V., Li, W., and Yu, H. (2019). Light-sheet microscopy in neuroscience. *Annu. Rev. Neurosci.* 42, 295–313. doi: 10.1146/annurev-neuro-070918-050357

Hornung, J. P. (2003). The human raphe nuclei and the serotonergic system. *J. Chem. Neuroanat.* 26, 331–343. doi: 10.1016/j.jchemneu.2003.10.002

Hrabetova, S., Cognet, L., Rusakov, D. A., and Nagerl, U. V. (2018). Unveiling the extracellular space of the brain: from super-resolved microstructure to *in vivo* function. *J. Neurosci.* 38, 9355–9363. doi: 10.1523/JNEUROSCI.1664-18.2018

Hu, Y., and Nualart, D. (2005). Renormalized self-intersection local time for fractional Brownian motion. *Ann. Prob.* 33, 948–983. doi: 10.1214/009117905000000017

Ito, K., and McKean, H. P. (1965). *Diffusion Processes and Their Sample Paths.* (New York, NY: Springer).

Ivgy-May, N., Tamir, H., and Gershon, M. D. (1994). Synaptic properties of serotonergic growth cones in developing rat brain. *J. Neurosci.* 14, 1011–1029. doi: 10.1523/JNEUROSCI.14-03-01011.1994

Jacobs, B. L., and Azmitia, E. C. (1992). Structure and function of the brain serotonin system. *Physiol. Rev.* 72, 165–229. doi: 10.1152/physrev.1992.72.1.165

Janušonis, S., and Detering, N. (2019). A stochastic approach to serotonergic fibers in mental disorders. *Biochimie* 161, 15–22. doi: 10.1016/j.biochi.2018.07.014

Janušonis, S., Gluncic, V., and Rakic, P. (2004). Early serotonergic projections to Cajal-Retzius cells: relevance for cortical development. *J. Neurosci.* 24, 1652–1659. doi: 10.1523/JNEUROSCI.4651-03.2004

Janušonis, S., Mays, K. C., and Hingorani, M. T. (2019). Serotonergic axons as 3D-walks. *ACS Chem. Neurosci.* 10, 3064–3067. doi: 10.1021/acschemneuro.8b00667

Jin, Y., Dougherty, S. E., Wood, K., Sun, L., Cudmore, R. H., Abdalla, A., et al. (2016). Regrowth of serotonin axons in the adult mouse brain following injury. *Neuron* 91, 748–762. doi: 10.1016/j.neuron.2016.07.024

Kajstura, T. J., Dougherty, S. E., and Linden, D. J. (2018). Serotonin axons in the neocortex of the adult female mouse regrow after traumatic brain injury. *J. Neurosci. Res.* 96, 512–526. doi: 10.1002/jnr.24059

Katori, S., Hamada, S., Noguchi, Y., Fukuda, E., Yamamoto, T., Yamamoto, H., et al. (2009). Protocadherin-α family is required for serotonergic projections to appropriately innervate target brain areas. *J. Neurosci.* 29, 9137–9147. doi: 10.1523/JNEUROSCI.5478-08.2009

Katori, S., Noguchi-Katori, Y., Okayama, A., Kawamura, Y., Luo, W., Sakimura, K., et al. (2017). Protocadherin-αC2 is required for diffuse projections of serotonergic axons. *Sci. Rep.* 7:15908. doi: 10.1038/s41598-017-16120-y

Kayasandik, C., Negi, P., Laezza, F., Papadakis, M., and Labate, D. (2018). Automated sorting of neuronal trees in fluorescent images of neuronal networks using NeuroTreeTracer. *Sci. Rep.* 8:6450. doi: 10.1038/s41598-018-24753-w

Kiyasova, V., and Gaspar, P. (2011). Development of raphe serotonin neurons from specification to guidance. *Eur. J. Neurosci.* 34, 1553–1562. doi: 10.1111/j.1460-9568.2011.07910.x

Lidov, H. G., and Molliver, M. E. (1982). An immunohistochemical study of serotonin neuron development in the rat: ascending pathways and terminal fields. *Brain Res. Bull.* 8, 389–430. doi: 10.1016/0361-92308290077-6

Linley, S. B., Hoover, W. B., and Vertes, R. P. (2013). Pattern of distribution of serotonergic fibers to the orbitomedial and insular cortex in the rat. *J. Chem. Neuroanat.* 48–49, 29–45. doi: 10.1016/j.jchemneu.2012.12.006

Lomholt, M. A., Ambjornsson, T., and Metzler, R. (2005). Optimal target search on a fast-folding polymer chain with volume exchange. *Phys. Rev. Lett.* 95:260603. doi: 10.1103/PhysRevLett.95.260603

Maddaloni, G., Bertero, A., Pratelli, M., Barsotti, N., Boonstra, A., Giorgi, A., et al. (2017). Development of serotonergic fibers in the post-natal mouse brain. *Front. Cell. Neurosci.* 11:202. doi: 10.3389/fncel.2017.00202

Mai, J. K., and Ashwell, K. W. S. (2004). "Fetal development of the central nervous system," in *The Human Nervous System,* eds G. Paxinos and J. K. Mai (San Diego, CA: Elsevier), 49–94. doi: 10.1016/B978-012547626-3/50004-1

Maia, G. H., Soares, J. I., Almeida, S. G., Leite, J. M., Baptista, H. X., Lukoyanova, A. N., et al. (2019). Altered serotonin innervation in the rat epileptic brain. *Brain Res. Bull.* 152, 95–106. doi: 10.1016/j.brainresbull.2019.07.009

Makse, H. A., Havlin, S., Schwartz, M., and Stanley, H. E. (1996). Method for generating long-range correlations for large systems. *Phys. Rev. E* 53, 5445–5449. doi: 10.1103/PhysRevE.53.5445

Mamounas, L. A., Blue, M. E., Siuciak, J. A., and Altar, C. A. (1995). Brain-derived neurotrophic factor promotes the survival and sprouting of serotonergic axons in rat brain. *J. Neurosci.* 15, 7929–7939. doi: 10.1523/JNEUROSCI.15-12-07929.1995

Mandelbrot, B. B., and Van Ness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev.* 10, 422–437. doi: 10.1137/1010093

Mano, T., Albanese, A., Dodt, H. U., Erturk, A., Gradinaru, V., Treweek, J. B., et al. (2018). Whole-brain analysis of cells and circuits by tissue clearing and light-sheet microscopy. *J. Neurosci.* 38, 9330–9337. doi: 10.1523/JNEUROSCI.1677-18.2018

Metzler, R., Jeon, J. H., Cherstvy, A. G., and Barkai, E. (2014). Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking. *Phys. Chem. Chem. Phys.* 16, 24128–24164. doi: 10.1039/C4CP03465A

Migliarini, S., Pacini, G., Pelosi, B., Lunardi, G., and Pasqualetti, M. (2013). Lack of brain serotonin affects postnatal development and serotonergic neuronal circuitry formation. *Mol. Psychiatr.* 18, 1106–1118. doi: 10.1038/mp.2012.128

Molina-Garcia, D., Sandev, T., Safdari, H., Pagnini, G., Chechkin, A., and Metzler, R. (2018). Crossover from anomalous to normal diffusion: truncated power-law noise correlations and applications to dynamics in lipid bilayers. *New J. Phys.* 20:103027. doi: 10.1088/1367-2630/aae4b2

Morgado, R., Oliveira, F. A., Batrouni, G. G., and Hansen, A. (2002). Relation between anomalous and normal diffusion in systems with memory. *Phys. Rev. Lett.* 89:100601. doi: 10.1103/PhysRevLett.89.100601

Morin, L. P., and Meyer-Bernstein, E. L. (1999). The ascending serotonergic system in the hamster: comparison with projections of the dorsal and median raphe nuclei. *Neuroscience* 91, 81–105. doi: 10.1016/S0306-45229800585-5

Mosienko, V., Beis, D., Pasqualetti, M., Waider, J., Matthes, S., Qadri, F., et al. (2015). Life without brain serotonin: reevaluation of serotonin function with mice deficient in brain serotonin synthesis. *Behav. Brain Res.* 277, 78–88. doi: 10.1016/j.bbr.2014.06.005

Nicholson, C., and Hrabetova, S. (2017). Brain extracellular space: the final frontier of neuroscience. *Biophys. J.* 113, 2133–2142. doi: 10.1016/j.bpj.2017.06.052

Numasawa, Y., Hattori, T., Ishiai, S., Kobayashi, Z., Kamata, T., Kotera, M., et al. (2017). Depressive disorder may be associated with raphe nuclei lesions in patients with brainstem infarction. *J. Affect. Disord.* 213, 191–198. doi: 10.1016/j.jad.2017.02.005

Okaty, B. W., Commons, K. G., and Dymecki, S. M. (2019). Embracing diversity in the 5-HT neuronal system. *Nat. Rev. Neurosci.* 20, 397–424. doi: 10.1038/s41583-019-0151-3

Okaty, B. W., Freret, M. E., Rood, B. D., Brust, R. D., Hennessy, M. L., deBairos, D., et al. (2015). Multi-scale molecular deconstruction of the serotonin neuron system. *Neuron* 88, 774–791. doi: 10.1016/j.neuron.2015.10.007

Papadopoulos, G. C., Parnavelas, J. G., and Buijs, R. (1987). Monoaminergic fibers form conventional synapses in the cerebral cortex. *Neurosci. Lett.* 76, 275–279. doi: 10.1016/0304-39408790414-9

Picka, J. (2012). Statistical inference for disordered sphere packings. *Stat. Surveys* 6, 74–112. doi: 10.1214/09-SS058

Polovnikov, K. E., Gherardi, M., Cosentino-Lagomarsino, M., and Tamm, M. V. (2018). Fractal folding and medium viscoelasticity contribute jointly to chromosome dynamics. *Phys. Rev. Lett.* 120:088101. doi: 10.1103/PhysRevLett.120.088101

Polovnikov, K. E., Nechaev, S., and Tamm, M. V. (2019). Many-body contacts in fractal polymer chains and fractional Brownian trajectories. *Phys. Rev. E* 99:032501. doi: 10.1103/PhysRevE.99.032501

Pratelli, M., Migliarini, S., Pelosi, B., Napolitano, F., Usiello, A., and Pasqualetti, M. (2017). Perturbation of serotonin homeostasis during adulthood affects serotonergic neuronal circuitry. *eNeuro* 4:ENEURO.0376-16.2017. doi: 10.1523/ENEURO.0376-16.2017

Pratelli, M., and Pasqualetti, M. (2019). Serotonergic neurotransmission manipulation for the understanding of brain development and function: learning from Tph2 genetic models. *Biochim* 161, 3–14. doi: 10.1016/j.biochi.2018.11.016

Qian, H. (2003). "Fractional Brownian motion and fractional Gaussian noise," in *Processes with Long-Range Correlations: Theory and Applications,* eds G. Rangarajan and M. Ding (Berlin/Heidelberg: Springer), 22–33. doi: 10.1007/3-540-44832-2_2

Qian, H., Raymond, G. M., and Bassingthwaighte, J. B. (1998). On two-dimensional fractional Brownian motion and fractional Brownian random field. *J. Phys. A Math. Gen.* 31:L527. doi: 10.1088/0305-4470/31/28/002

Quentin, E., Belmer, A., and Maroteaux, L. (2018). Somato-dendritic regulation of raphe serotonin neurons; a key to antidepressant action. *Front. Neurosci.* 12:982. doi: 10.3389/fnins.2018.00982

Ren, J., Friedmann, D., Xiong, J., Liu, C. D., Ferguson, B. R., Weerakkody, T., et al. (2018). Anatomically defined and functionally distinct dorsal raphe serotonin sub-systems. *Cell* 175, 472–487.e420. doi: 10.1016/j.cell.2018.07.043

Reverey, J. F., Jeon, J. H., Bao, H., Leippe, M., Metzler, R., and Selhuber-Unkel, C. (2015). Superdiffusion dominates intracellular particle motion in

the supercrowded cytoplasm of pathogenic *Acanthamoeba castellanii*. *Sci. Rep.* 5:11690. doi: 10.1038/srep11690

Schneeberger, M., Parolari, L., Das Banerjee, T., Bhave, V., Wang, P., Patel, B., et al. (2019). Regulation of energy expenditure by brainstem GABA neurons. *Cell* 178, 672–685.e612. doi: 10.1016/j.cell.2019.05.048

Schulz, J. H. P., Chechkin, A. V., and Metzler, R. (2013). Correlated continuous time random walks: combining scale-invariance with long-range memory for spatial and temporal dynamics. *J. Phys. A Math. Theor.* 46:475001. doi: 10.1088/1751-8113/46/47/475001

Slaten, E. R., Hernandez, M. C., Albay, R., Lavian, R., and Janušonis, S. (2010). Transient expression of serotonin 5-HT_4 receptors in the mouse developing thalamocortical projections. *Dev. Neurobiol.* 70, 165–181. doi: 10.1002/dneu.20775

Smith, S., Cianci, C., and Grima, R. (2017). Macromolecular crowding directs the motion of small molecules inside cells. *J. R Soc. Interface* 14:20170047. doi: 10.1098/rsif.2017.0047

Soiza-Reilly, M., Anderson, W. B., Vaughan, C. W., and Commons, K. G. (2013). Presynaptic gating of excitation in the dorsal raphe nucleus by GABA. *Proc. Natl. Acad. Sci. U.S.A.* 110, 15800–15805. doi: 10.1073/pnas.1304505110

Stamp, J. A., and Semba, K. (1995). Extent of colocalization of serotonin and GABA in the neurons of the rat raphe nuclei. *Brain Res.* 677, 39–49. doi: 10.1016/0006-89939500119-B

Steinbusch, H. W. (1981). Distribution of serotonin-immunoreactivity in the central nervous system of the rat-cell bodies and terminals. *Neuroscience* 6, 557–618. doi: 10.1016/0306-45228190146-9

Stuesse, S. L., Cruce, W. L., and Northcutt, R. G. (1991). Localization of serotonin, tyrosine hydroxylase, and leu-enkephalin immunoreactive cells in the brainstem of the horn shark, *Heterodontus francisci*. *J. Comp. Neurol.* 308, 277–292. doi: 10.1002/cne.903080211

Sundstrom, E., Kolare, S., Souverbie, F., Samuelsson, E. B., Pschera, H., Lunell, N. O., et al. (1993). Neurochemical differentiation of human bulbospinal monoaminergic neurons during the first trimester. *Brain Res. Dev. Brain Res.* 75, 1–12. doi: 10.1016/0165-380693 90059-J

Voigt, T., and de Lima, A. D. (1991a). Serotoninergic innervation of the ferret cerebral cortex. I. Adult pattern. *J. Comp. Neurol.* 314, 403–414. doi: 10.1002/cne.903140214

Voigt, T., and de Lima, A. D. (1991b). Serotoninergic innervation of the ferret cerebral cortex. II. Postnatal development. *J. Comp. Neurol.* 314, 415–428. doi: 10.1002/cne.903140215

Vojta, T., Halladay, S., Skinner, S., Janušonis, S., Guggenberger, T., and Metzler, R. (2020). Reflected fractional Brownian motion in one and higher dimensions. *ArXiv e-prints arXiv*. arXiv:2004.14511. Available online at: https://arxiv.org/abs/2004.14511

Vojta, T., Skinner, S., and Metzler, R. (2019). Probability density of the fractional Langevin equation with reflecting walls. *Phys. Rev. E* 100:042142. doi: 10.1103/PhysRevE.100.042142

Wada, A. H. O., and Vojta, T. (2018). Fractional Brownian motion with a reflecting wall. *Phys. Rev. E* 97:020102. doi: 10.1103/PhysRevE.97.020102

Wada, A. H. O., Warhover, A., and Vojta, T. (2019). Non-Gaussian behavior of reflected fractional Brownian motion. *J. Stat. Mech. Theory Exp. 2019*:033209. doi: 10.1088/1742-5468/ab02f1

Wassie, A. T., Zhao, Y., and Boyden, E. S. (2019). Expansion microscopy: principles and uses in biological research. *Nat. Methods* 16, 33–41. doi: 10.1038/s41592-018-0219-4

Whitaker-Azmitia, P. M. (2001). Serotonin and brain development: role in human developmental diseases. *Brain Res. Bull.* 56, 479–485. doi: 10.1016/S0361-92300100615-3

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.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 29 July 2020
doi: 10.3389/fncom.2020.00064

Necessary Conditions for Reliable Propagation of Slowly Time-Varying Firing Rate

Navid Hasanzadeh[1,2], *Mohammadreza Rezaei*[1,3], *Sayan Faraz*[1], *Milos R. Popovic*[3,4] and *Milad Lankarany*[1,3,4*]

[1] Clinical and Computational Neuroscience, Krembil Research Institute, University Health Network, Toronto, ON, Canada, [2] School of Electrical and Computer Engineering, College of Engineering, University of Tehran, Tehran, Iran, [3] KITE Research Institute, Toronto Rehabilitation Institute, University Health Network, Toronto, ON, Canada, [4] Institute of Biomedical Engineering, University of Toronto, Toronto, ON, Canada

Reliable propagation of slow-modulations of the firing rate across multiple layers of a feedforward network (FFN) has proven difficult to capture in spiking neural models. In this paper, we explore necessary conditions for reliable and stable propagation of time-varying asynchronous spikes whose instantaneous rate of changes—in fairly short time windows [20–100] msec—represents information of slow fluctuations of the stimulus. Specifically, we study the effect of network size, level of background synaptic noise, and the variability of synaptic delays in an FFN with all-to-all connectivity. We show that network size and the level of background synaptic noise, together with the strength of synapses, are substantial factors enabling the propagation of asynchronous spikes in deep layers of an FFN. In contrast, the variability of synaptic delays has a minor effect on signal propagation.

Keywords: time-varying rate coding, information propagation, feed-forward neural network, noise, network size, synaptic delays

OPEN ACCESS

Edited by:
Thomas Nowotny,
University of Sussex, United Kingdom

Reviewed by:
Changsong Zhou,
Hong Kong Baptist University, Hong Kong
Jorge F. Mejias,
University of Amsterdam, Netherlands

***Correspondence:**
Milad Lankarany
milad.lankarany@uhnresearch.ca

Received: *02 March 2020*
Accepted: *28 May 2020*
Published: *29 July 2020*

Citation:
Hasanzadeh N, Rezaei M, Faraz S, Popovic MR and Lankarany M (2020) Necessary Conditions for Reliable Propagation of Slowly Time-Varying Firing Rate. Front. Comput. Neurosci. 14:64. doi: 10.3389/fncom.2020.00064

INTRODUCTION

Information in the brain is encoded by either the number of spikes in a relatively long time window, i.e., rate code, or by their precise timing, i.e., temporal code (Abeles et al., 1994; Panzeri et al., 2001, 2017; Montemurro et al., 2007; Kremkow et al., 2010; London et al., 2010; Zuo et al., 2015; Runyan et al., 2017; Noble, 2019). The feasibility of utilizing both coding strategies has also been shown in different neural systems (Kumar et al., 2010; Lankarany et al., 2019). In temporal coding, information is carried by groups of neurons that fire *synchronously*, as in synfire chains (Abeles et al., 1994; Diesmann et al., 1999), whereas in rate coding, neuronal firing ideally remains asynchronous across neurons (Litvak et al., 2003). Information processing in a hierarchically organized cortical system relies on the reliable propagation of synchronous and asynchronous spikes (Joglekar et al., 2018).

The reliable propagation of synchronous spikes (temporal code) is well-understood and relatively easy to implement in computer models (Kumar et al., 2008, 2010; Joglekar et al., 2018). In contrast, the reliable propagation of rate-modulated asynchronous spiking (rate code) is poorly understood and remains challenging to implement in computer models (Litvak et al., 2003). Indeed, spikes may synchronize as the signal progresses through deeper layers or may tend toward an attractor state representing quiescence or a fixed rate. In all of the scenarios, rate-based coding is compromised.

Several studies have addressed the conditions required for spike-rate propagation (Shadlen and Newsome, 1994; van Rossum et al., 2002; Litvak et al., 2003; Wang et al., 2006; Kumar et al., 2008, 2010; Joglekar et al., 2018; Barral et al., 2019; Han et al., 2019). Shadlen and Newsome (Shadlen and Newsome, 1994) demonstrated the feasibility of rate transmission using leaky integrate and fire (LIF) models receiving balanced excitatory and inhibitory inputs. Nevertheless, Litvak et al. (2003) showed, using the same parameters quoted by Shadlen and Newsome (1994), that rate was not reliably transmitted when >2 layers were considered. They concluded that rate transmission in FFNs is highly unlikely. van Rossum et al. (2002) showed the feasibility of reliable transmission of instantaneous firing rate (asynchronous spikes) in un-balanced FFNs where the input to each layer is delivered as an injected current. Kumar et al. (2008) studied conditions for propagating synchronous and asynchronous spikes utilizing biologically realistic network models. The coexistence of firing rate and synchrony propagation was shown under precise combinations of synaptic strength and connection probability (Kumar et al., 2010). Cortes and van Vreeswijk (2015) showed that the pulvinar thalamic nucleus allows for asynchronous spike propagation through the cortex; they supply the input-output firing rate relationship between two cortical areas without manipulating synaptic strengths.

Neuronal networks with feedforward connections consisting of excitatory neurons are thought of as the model for linking upstream neurons with downstream neurons, either across different layers within the same cortical region or between different cortical regions. Recent studies have capitalized on the role of recurrent connections in reliable transmission of synchronous and asynchronous spikes (Joglekar et al., 2018; Barral et al., 2019). A recent study has demonstrated that networks with recurrent excitation and lateral inhibition stabilize signal transmission (Joglekar et al., 2018). As well, a recurrent network including AMPA and NMDA-mediated components that operates in an excitatory-inhibitory balanced regime (with respect to both magnitude and time of excitatory and inhibitory synaptic inputs) has been recently proposed as a novel model of information propagation (Barral et al., 2019).

Despite undoubtedly significant impacts of biologically realistic network architectures in conveying information across/between layers, neuronal networks with feedforward connections play a substantial role in understanding the mechanisms of faithful propagation of different types of spikes (Kumar et al., 2010). Due to the simplicity of feedforward networks (FFNs) compared to networks with recurrent connections, the impact of different factors like intrinsic properties of neurons in reliable propagation of spikes can be better studied using FFNs. For example, Han et al. (2019) showed that layer-to-layer heterogeneity arising from lamina-specific cellular properties facilitates propagation of synchronous and asynchronous spikes in FFNs.

In this paper, we focus to systematically explore the necessary conditions underlying which information of time-varying asynchronous spikes can be reliably transmitted in deep layers of an FFN. Using optimal synaptic weights (Faraz et al., 2020), we study the roles of (i) network size, (ii) the level of background synaptic noise, and (iii) the variability of synaptic delays in propagation of slowly time-varying asynchronous spikes. We show that unlike the variability of synaptic delays that has a minor effect on signal propagation, network size and the level of background synaptic noise are substantial factors enabling the propagation of asynchronous spikes in deep layers of an FFN.

METHODS

Network size, level of background synaptic noise, and variability of synaptic delays are the main factors that were investigated in a feedforward architecture. We created an FFN composed of excitatory neurons, modeled by leaky integrate and fire (LIF) model, receiving shared input from the previous layer plus background synaptic noise. We calculated coding fraction—representing goodness of propagation of time-varying asynchronous spikes—for different levels of background synaptic noise, network size, and variability of synaptic delays.

Optimal vs. Fixed Synaptic Weights

To test the effect of each factor, we estimated synaptic weights using a reduced network model whose function, in the propagation of a shared input, is equivalent to an FFN with all-to-all connectivity [see (Lankarany, 2019; Faraz et al., 2020) for more details]. Schematic representations of the reduced network and an FFN with all-to-all connectivity were shown in **Figures S1A,B**. To better distinguish between the performances of an FFN with and without optimal synaptic weight, we calculated the input signal of the second layer of an FFN given the spikes of the first layer. This input, the reconstructed stimulus, was shown in **Figures S1C,D** for optimal and fixed synaptic weights, respectively. The value of the fixed synaptic weights generated 0.5 mV postsynaptic potential (per spike), which is within a biologically realistic range. As can be seen in these figures, the stimulus reconstructed by optimal synaptic weights significantly better tracked the original stimulus (specifically for the optimal range of network size and noise level). For a reduced network model with a fixed number of neurons and a constant level of synaptic noise, the optimal weights were estimated to minimize the L2-norm error between the reconstructed (in the first layer) and original stimulus (Faraz et al., 2020). Of note in the reduced model, is the vector representation of the synaptic weights (rather than the matrix form) that enables the use of convex optimization techniques to calculate these weights. Given the estimated weights in the reduced network model, the mean and standard deviation (std) of the weights were calculated. To obtain synaptic weights for an FFN with all-to-all connectivity, we drew samples from a Gaussian distribution of the mean and std of the estimated weights. Codes for estimating synaptic weights are available at https://github.com/nsbspl/async-spike-propagation.

Slow Stimulus

The slow signal was delivered to the neurons in the first layer of an FFN and modeled by an Ornstein-Uhlenbeck (OU) process of the time constant of 50 msec. This slow signal can, for example,

represent the luminance of a natural stimulus (Lankarany et al., 2019). The OU process can be written as:

$$\frac{dx}{dt} = -\frac{x(t) - \mu}{\tau} + a\sqrt{\frac{2}{\tau}}\,\xi(t) \tag{1}$$

where ξ is a random number drawn from a Gaussian distribution with 0 average and unit variance. τ is the time constant, μ and α indicate the mean and standard deviation of variable x, respectively. The mean and variance of the slow signal is 16 and 15 pA, respectively.

Level of Background Synaptic Noise (OU Process)

Background synaptic noise was modeled by an OU process of the time constant of 5 msec (Destexhe et al., 2001). The mean of synaptic noise is 0 that balances the effect of background excitatory and inhibitory synaptic inputs. The standard deviation of noise varied in a range of (5, 10, 15, 20, 25, 30, 35, 40) pA, which generated background (no stimulus) spiking activity in a range of (1.2–10.4) Hz.

Neuron Model

We used leaky integrate and fire (LIF) model. The dynamics of membrane potential is expressed as follows.

$$\frac{dV}{dt} = \frac{-(V - E_L) + RI_{inj}}{\tau_V} \tag{2}$$

where $E_L = -70$ mV, $R = 1$ MΩ, and $\tau_V = 10$ msec. I_{inj} indicates the injected current (slow signal as the stimulus of the first layer). Spike occurs when V≥Vth, where $V_{th} = -40$ mV and the reset voltage is -90 mV. For the neurons in the subsequent layers, I_{inj} is equal to the total presynaptic input (weighted by synaptic strengths) plus independent synaptic noise. A double exponential function of $\tau_{rise} = 0.5$ *msec* & $\tau_{fall} = 5$ *msec* was used to model an identical synaptic waveform.

Network Size

Our computational study was performed by varying the number of neurons in an FFN. Network sizes were varied in a range of [50, 100, 200, 300, 400, 500, 750, 1000].

Variability in Synaptic Delays

Variability in synaptic delays introduced heterogeneities in the network. This variability can be interpreted as an un-equal distance between pre- and post-synaptic neurons (see Discussion). We modeled such variability by a Gaussian distribution whose mean represents the synaptic delay (3 msec) and standard deviation indicates the level of variability across neurons. The std of synaptic delays varies within an interval of (0:0.25:1.5) msec, which is in agreement with small variations of synaptic latencies in the cortex (Boudkkazi et al., 2007).

Instantaneous Firing Rate and Optimal Kernel Width

The instantaneous firing rate was calculated by convolving superimposed spikes (of each layer of an FFN) with a Gaussian kernel. To achieve a consistent comparison between the firing rates across layers, we used a Gaussian kernel of width = 25 msec. This kernel width was near to optimal for spikes in the first layer of the FFN (regardless the network size and noise level). We used a method proposed in Shimazaki and Shinomoto (2010) to calculate the optimum (Gaussian) kernel width of the spikes in the first layer. The optimal kernel width was equal to 21.8 msec. We plotted the instantaneous firing rates estimated by these widths against each other in **Figure S2** to demonstrate the similarities between these estimates.

Coding Fraction and Kullback–Leibler Divergence

To quantify how much network size and the level of background synaptic noise effect the propagation of asynchronous spikes in a network, we used coding fraction (CF) that represents how good the instantaneous firing rate of the second layer tracks that of the first layer. CF is calculated as follows.

$$CF = 1 - \frac{\left\| Firing(Layer\ 2) - Firing(Layer\ 1) \right\|_2}{\left\| Firing(Layer\ 1) \right\|_2} \tag{3}$$

where *Firing (Layer 1)* and *Firing (Layer 2)* denote the instantaneous firing rate of the 1st and the 2nd layers, respectively. And, $||.||_2$ indicates the norm 2. CF lies within [−1, 1], where 1 represents perfect transmission.

Besides, in order to validate whether CF is equivalent to an information-theoretic measure, we calculated Kullback–Leibler (KL) divergence measures between the instantaneous firing rate of the second and the first layers of the FFN. The Kullback–Leibler (KL) divergence (Timme and Lapish, 2018) quantifies the similarity between the probability distributions of the instantaneous firing rates of the first (Fr_1) and the second (Fr_2) layers. The KL divergence is defined as (Perez-Cruz, 2008):

$$D_{KL}\left(P_{Fr_2}(r_2), P_{Fr_1}(r_1)\right) = \sum_{r_1 \in Fr_1,\, r_2 \in Fr_2} P_{Fr_2}(r_2) * \log\left(\frac{P_{Fr_2}(r_2)}{P_{Fr_1}(r_1)}\right) \tag{4}$$

where P_{Fr_1} and P_{Fr_2} are the probability distributions of the firing rates, namely, Fr_1 and Fr_2, respectively. To calculate these probability distributions, we utilized a non-parametric estimation method to approximate them using normal kernel smoothing (Bowman and Azzalini, 1997). The estimated density function, $\hat{f}_h(x)$, for each layer, can be written as follows.

$$\hat{f}_h(x) = \frac{1}{Nh}\sum_{i=1}^{N} K\left(\frac{x - x_i}{h}\right);\ -\infty < x < \infty \tag{5}$$

where N is the sample size, $K(.)$ is the kernel function, and h is the bandwidth. D_{KL} is non-negative (≥ 0) and non-symmetric in $P(r_2)$ and $P(r_1)$. D_{KL} is equal to zero if $P(r_2)$ and $P(r_1)$ match exactly, and can potentially equal infinity if there is no similarity between the two distributions. In other words, the KL measure is equal to zero if the firing rates of both layers have the

same statistical characteristics, implying a perfect information propagation (across two layers).

We calculated CF and KL divergence for different levels of background synaptic noise (for $N = 400$ and $N = 500$). **Figure S3** shows that these two measures are inversely correlated, i.e., a higher CF is equivalent to a lower KL divergence, which represents more similarities between the probability distributions of the firing rates, and thus, a better propagation of the asynchronous firing rates.

Synchrony Measure

Similar to Lankarany et al. (2019), we used a threshold method to detect synchronous events. Synchrony threshold, for each layer of an FFN with specific network size, noise level, and optimal synaptic weights, was chosen such that the superimposed spikes during a 10 msec interval generate >20 mV post-synaptic potential. As the optimal synaptic weights depend on the network size and noise level, the synchrony threshold can be interpreted as a minimum number of spikes required for post-synaptic neurons to fire in a short interval.

RESULTS

Network Size and Level of Background Synaptic Noise Are Substanital Factors That Enable Reliable Propagation of Slowly Time-Varying Asynchronous Spikes in FFNs

To explore the effects of network size and the level of background synaptic noise in the propagation of slowly time-varying asynchronous spikes, we vary these parameters in an FFN consisting of two layers and calculate the coding fraction (CF). In addition, the number of synchronous events (see Methods) is quantified to provide more numerical characterizations of the roles of the above parameters in signal propagation.

A recent study (Barral et al., 2019) has shown that channel capacity and decoding accuracy decreased after the first layer but remained above the chance (and almost unchanged). However, the rationale behind using two layers is that the propagation of information across layers of an FFN is more likely to maintain robust if the information is preserved within the first two layers.

Figure 1A shows CF and synchrony measure of spikes as a function of network size and std of the background synaptic noise. Three values underlying each parameter are chosen to classify the range of network size and level of noise into three categories, namely, small (low), medium (moderate), and large (high). Note that spikes in the second layer of an FFN (for each pair of parameters) are produced by optimal synaptic weights (see Methods and Distribution of Optimal Synaptic Weights Depends on Network Size and Level of Synaptic Noise).

CF increases for medium and large network sizes with moderate to high levels of background noise. However, CF is relatively low outside these ranges; it decreases for small network sizes regardless of the noise level. More precise characterization of CF against different values of network size and noise level (see **Figure S4**) indicates that CF has the highest values for a specific range of parameters, i.e., network sizes in a range of {400–500} and std of the noise in a range of $\{20–30\}^{pA}$.

The number of synchronous events increases for medium- and large-size FFNs with low levels of noise (also note that the signal is not well represented in small-size FFNs with low noise). Thus, one can predict that more synchronous events would be generated in the subsequent layers for low levels of background synaptic noise. **Figure 1A** (right) also shows that synchronous events occur in large-size FFNs with moderate level of noise, suggesting that an improper ratio of network size and noise level causes synchrony propagation in the subsequent layers.

To visually inspect the effects of network size and noise std, the instantaneous firing rates of the first and the second layers are shown in **Figure 1B**. CF is relatively high for (i) large network sizes ($\geq$750) with a high level of synaptic noise ($\sigma =$ (35–40) pA) and (ii) small network sizes (<300) with a high level of synaptic noise ($\sigma = 40$ pA). In (i), signal propagation in deeper layers is not stable due to the high level of noise: the rate code remains asynchronous but tends to the average firing rate (see **Figure S5B**). As will be discussed in section Distribution of Optimal Synaptic Weights Depends on Network Size and Level of Synaptic Noise, the corresponding synaptic weights are very weak and biologically unrealistic ($N = 750$, $\sigma = 40$ pA). The high level of noise in (ii) makes the network spontaneously active, so it is not possible to discern any signal from background synaptic activity. As shown in **Figure 1B**, a large network size ($N = 750$) with a small level of background activity ($\sigma = 10$ pA) increases the shared input activity and amplifies the signal in the second layer. In this case, synchrony increases across the layers (see also **Figure S5C** for medium-size network with low noise). In medium-size FFNs ($N = 400$), small (10 pA) and large (40 pA) synaptic noise amplifies and attenuates signal propagation, respectively.

To better understand the effects of network size and noise level on the transmission of asynchronous spikes within the first two layers of an FFN, CF is plotted in **Figure 2** for different values of each factor while the other factor is kept constant. **Figure 2 (Left)** shows CF for different network sizes when the noise std, $\sigma =$ 30 pA. CF is maximized for $N = 400$ and $N = 500$, meaning that information of asynchronous spikes is best transmitted to the 2nd layer of an FFN for particular network sizes (given a fixed level of synaptic noise). Similarly, **Figure 2 (Right)** shows CF as a function of background synaptic noise for FFNs consisting of 400 and 500 neurons. CF decreases for small and large (CF = 0.56 for $\sigma = 50$ pA, data is not shown in the figure) levels of background synaptic noise which implies that information is best transmitted to the 2nd layer of a medium-size FFN ($N = 400$, 500) given a moderate level of synaptic noise, i.e., $\sigma = 30$ pA.

Simulating an FFN with the optimal network size and level of synaptic noise enhances the propagation of asynchronous spikes in deeper layers. To evaluate the performance of an FFN with the optimal parameters in the propagation of asynchronous spikes in deeper layers, we show the raster plot and the instantaneous firing rate of an FFN up to five layers in **Figure 3**. Stable propagation of asynchronous spikes can be seen in this figure for $N = 400$, $\sigma = 30$ pA (as well as in **Figure S5A**). The instantaneous firing rates represent signal propagation across layers (bottom right curves in **Figure 3**).

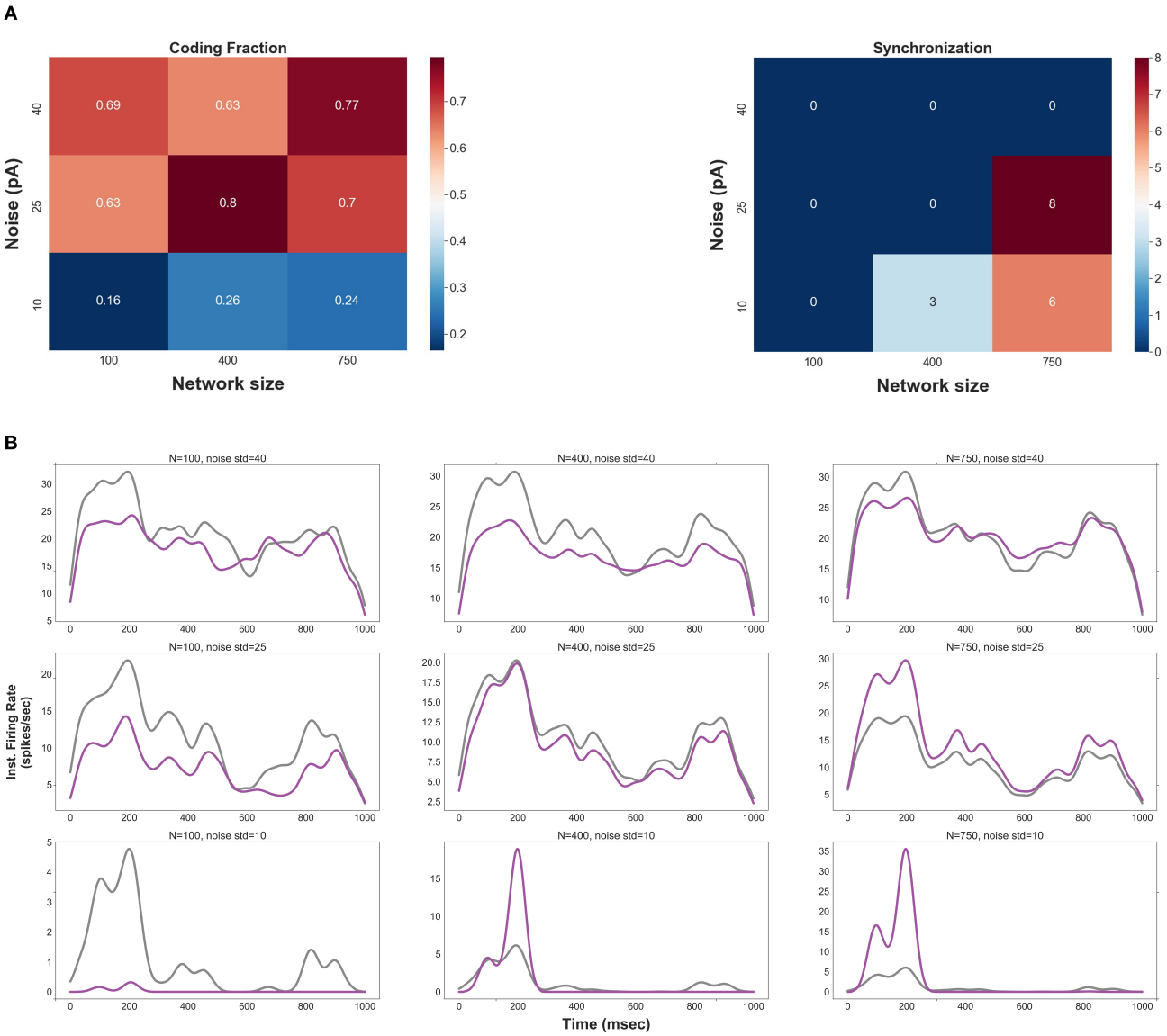

FIGURE 1 | (A) Coding fraction (CF), CF is plotted for different values of network size and the level of background synaptic noise (in a two layers FFN). **(B)** Instantaneous firing rate of the 2nd layer (purple) vs. that of the 1st layer (gray) for different pairs of network size and noise level. Note that the delay between layers is compensated in these curves.

Previously, it was shown that a homogeneous random network could not transmit a completely asynchronous population activity due to the existence of the residual correlations (Kumar et al., 2008). Such correlations are mainly originated from the shared connectivity and could be reduced by increasing the network size (Kumar et al., 2008). Unlike homogeneous random networks, in an FFN with all-to-all connectivity, i.e., a dense network, a larger network size results in a larger number of shared pre-synaptic inputs. Although background synaptic noise induces some heterogeneities in the network, the propagation of asynchronous spikes in deeper layers is not enhanced for network sizes $>$ 750. It is to be noted that increasing the noise level in large network sizes remains spikes propagated asynchronously; however, at the expense of losing information of time-varying firing rates (**Figure S5B**). Moreover, decreasing the level of background synaptic noise causes synchronous spikes built up in the subsequent layers of a medium-size FFN (**Figure S5C**).

Distribution of Optimal Synaptic Weights Depends on Network Size and Level of Synaptic Noise

In addition to synaptic noise that provides heterogeneities in an FFN [or a locally connected random network (Mehring

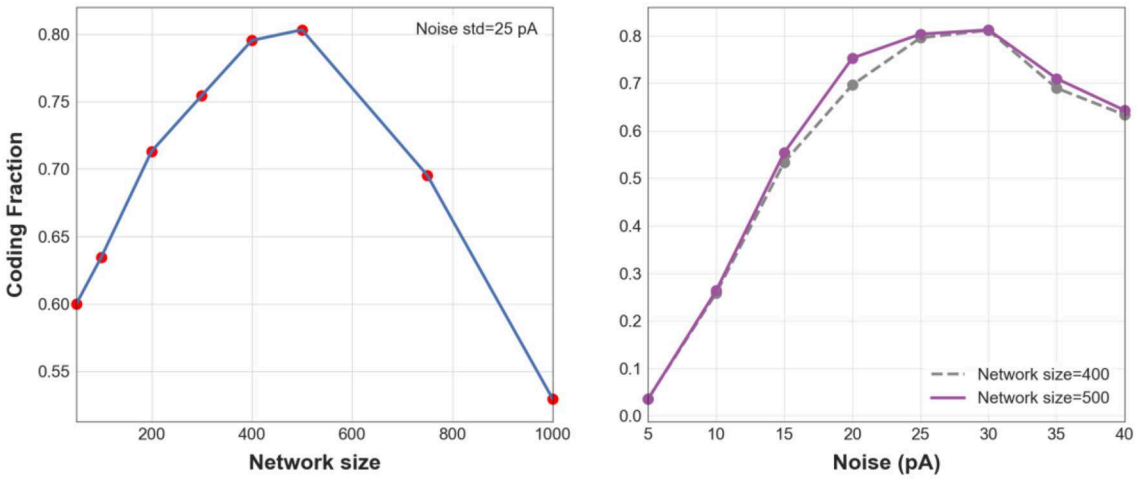

FIGURE 2 | 1-D CF. CF vs. different network sizes when noise std is equal to 30 pA (**Left**). CF vs. std of synaptic noise (**Right**) for FFNs with 400 (blue dashed line) and 500 neurons (purple).

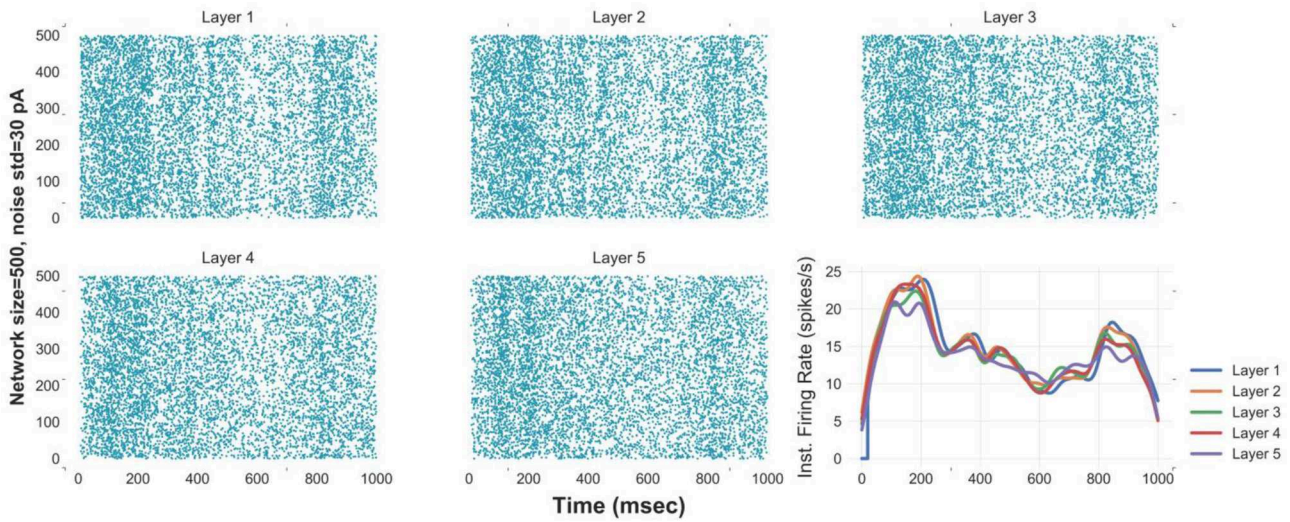

FIGURE 3 | Reliable propagation of slowly-time varying asynchronous spikes in 5 layers [raster plot & instantaneous firing rate (Gaussian kernel of width 25 msec)] of FFNs with optimal values of network size (500), level of synaptic noise (30 pA), and synaptic strength. Note: the delay between layers is compensated in these curves.

et al., 2003)], previous studies considered that a distribution of passive properties of neurons that rises to a distribution of the synaptic weights (Kumar et al., 2008) induces heterogeneities across neurons. In this study, synaptic weights are estimated from the firing activities of the first layer, and the underlying mean and standard deviation (with an assumption of Gaussian distribution) are used to draw weights for subsequent layers. Thus, the mean of the estimated weights reflects the strength of the synapses, and the std of the weights indicates heterogeneities across synapses.

To test whether the estimated weights are biologically realistic, we validate if the corresponding postsynaptic potential (PSP) lies within a biologically-reasonable range. We show in **Figure 4,** a heatmap of PSPs—corresponding to the mean of the estimated weights—as a function of network size and level of background synaptic noise. Synaptic weights generating extremely weak ($<$ 0.05 mV) or extremely strong ($>$1.5 mV) PSPs per presynaptic spike are considered as unrealistic weights. As can be seen in this figure, the synaptic weights for large network sizes ($N > 750$) and moderate to high background noise are weak, generating postsynaptic potentials less than 0.1 mV (per single presynaptic spikes). In contrast to the large network sizes with weak synapses, the estimated synaptic weights of FFNs with small network sizes generate larger postsynaptic potentials, specifically for the lower levels of synaptic noise. Therefore, one can conclude that only a specific range of synaptic weights is biologically realistic.

Improper selection of synaptic weights for an FFN results in unstable propagation of asynchronous spikes. Systematically varying synaptic strengths from moderately weak to moderately

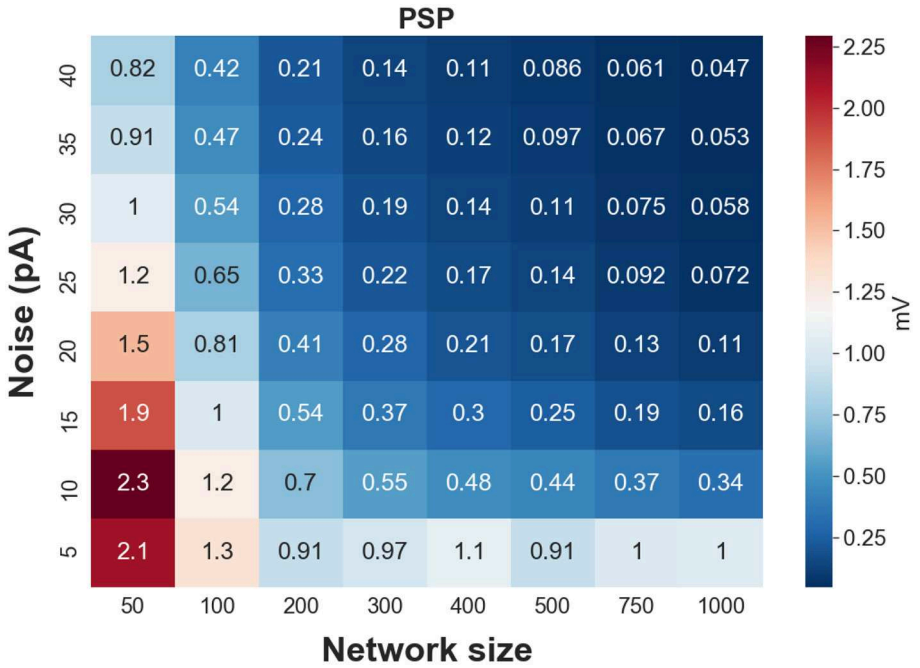

FIGURE 4 | Heat map of postsynaptic potentials (corresponding to synaptic weights) for different network sizes and levels of synaptic noise.

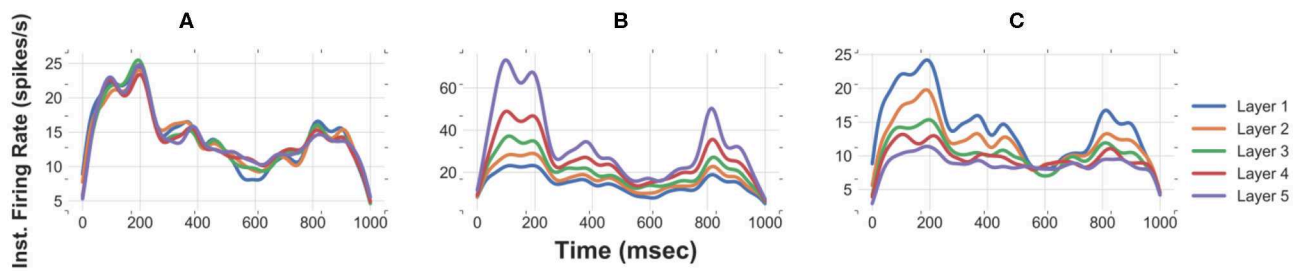

FIGURE 5 | Improper synaptic weights result in either attenuation or amplification of the slow signal in the subsequent layers. Instantaneous firing rates of an FFN with $N = 500$ and noise std = 30 pA for **(A)** proper synaptic weights, **(B)** strong synapses (corresponding to $N = 400$ and noise std = 30 pA), and **(C)** weak synapses (corresponding to $N = 600$ and noise std = 30 pA).

strong synapses, we find a transition between attenuation mode, where transmission of time-varying rates failed, to an amplification mode where the average firing rate increased at the subsequent layers (**Figure 5**). We repeat running an FFN of size 500 with the level of noise of 30 pA with synaptic weights (mean and std) obtained from an FFN with the same level of synaptic noise but different network sizes, i.e., relatively stronger synapses (0.1 pA/mV corresponding to $N = 400$) and weaker synapses (< 0.05 pA/mV corresponding to $N = 600$). **Figure 5** demonstrates that the propagation of a slow signal is either amplified (**Figure 5B**) or attenuated (**Figure 5C**) when the weights are stronger or weaker, respectively.

It is to be noted that the time constants of excitatory synapses and neurons' membrane potential influence the estimated synaptic weights. Chan et al. (2016) showed that different combinations of these time constants alter the pairwise correlation and synchrony of spikes, e.g., burst firing occurs when the time constant the neuron's membrane is small and that for excitatory synapses is large. These time constants are fixed in our estimation method (see Methods), and no burst firing was observed in medium-size FFNs ($N = 400$, 500) with a moderate level of synaptic noise (20, 25, and 30 pA).

Variability of Synaptic Delays Has a Slight Influence on the Propagation of Asynchronous Spikes

In biologically-realistic scenarios, presynaptic inputs are delivered from neurons with different distances to the target (postsynaptic) neuron. In the context of signal propagation,

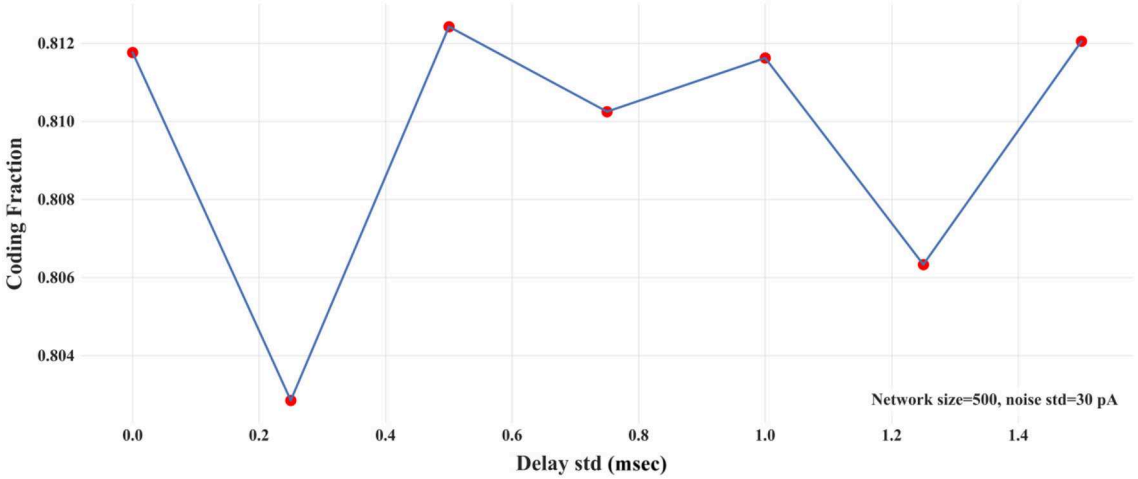

FIGURE 6 | CF vs. std of synaptic delays in a two-layer FFN with $N = 500$ and noise std = 30 pA.

for example, Joglekar et al. (2018) have introduced inter-areal delays by considering the corresponding inter-areal wiring distances (Markov et al., 2014) and assuming a constant axonal conduction velocity. To explore how different distances between the pre- and post-synaptic neurons affect signal propagation, we introduce variability in the synaptic delays. Synaptic delays are drawn from a Gaussian distribution of a fixed mean (3 msec) and variable std. We calculate CF for different values of the std of synaptic delays in a two-layer FFN ($N = 500$ and $\sigma = 30$ pA). **Figure 6** shows that CF is nearly independent of the variability of synaptic delays. In addition, we show the propagation of asynchronous spikes in an FFN with two levels of variabilities of synaptic delays in **Figure S6**. Similar to **Figure 3**, the FFN comprises 500 neurons, each receiving 30 pA synaptic noise. As expected, the variability of synaptic delays does not change signal propagation significantly compared to that with no variability.

What Range of Firing Rates Can Be Reliably Tranmistted Across Multiple Layers?

We construct logistic maps of the firing rates in the first and the 5th layers of an FFN ($N = 500$ and $\sigma = 30$ pA) to determine the range of frequencies that can be reliably transmitted. **Figure 7** shows that the information of asynchronous spikes is preserved across layers for frequencies in the range of [5–25] Hz. The curve in this figure is obtained by superimposing the average instantaneous firing rate of 5 networks, each of which receiving different slow stimuli (with the same time constants). For frequencies > 10 Hz, although the logistic curve becomes slightly sublinear, the firing rate of the 5th layer still tracks that of the first layer, indicating that the firing rate is uniquely represented in deep layers of the FFNs. However, for frequencies > 25 Hz, the logistic curve is supralinear that capitalizes on the tendency of the transmitted firing rates to the synchrony mode.

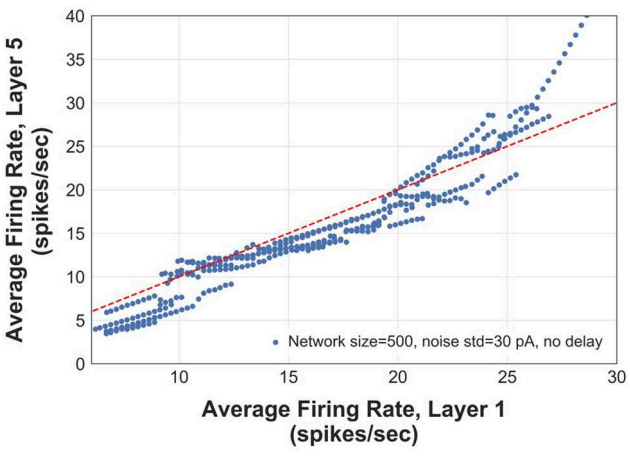

FIGURE 7 | Logistic map showing output rate (firing rate in layer 5) vs. input rate (firing rate in layer 1). The red dashed line has a slope of 1.

DISCUSSION

Necessary conditions for reliable propagation of time-varying asynchronous spikes in FFNs were investigated in this paper. Previous studies have addressed those conditions for transmission of synchronous spikes as well as the mean of firing rate of asynchronous spikes. However, these conditions barely remain valid for transmission of time-varying asynchronous spikes whose instantaneous rate of changes represents information of slow fluctuations of the stimulus. In this paper, we investigated the necessary conditions for reliable transmission of asynchronous spikes in an FFN. Specifically, we explored the effect of network size, level of background synaptic noise, and the variability of synaptic delays in an FFN with all-to-all connectivity. We demonstrated that network size and the level of background synaptic noise, together with optimal synaptic

weights, were substantial factors that cooperatively enable the propagation of asynchronous spikes in deep layers of an FFN. Nevertheless, the variability of synaptic delays had a minor effect on signal propagation. Varying these factors systematically, we found that an FFN of network size of {400–500} with the level of synaptic noise in a range of {20–30} pA and proper synaptic weights transmits time-varying asynchronous spikes reliably.

Reliable propagation of a rate code is challenging: even weak pairwise correlations in the spike timing can notably deteriorate the fidelity of the rate code (Kumar et al., 2010). Thus, the robust transmission of asynchronous spikes cannot be simply achieved from a general theory based on the balance of (common) signal and (background) noise. We showed that the same level of (moderate) noise provides very different signal propagations in networks with different numbers of neurons; it degrades the firing rate of an FFN of size = 100 whereas it can maintain a reliable signal propagation for a network size of 400 (the synaptic weights are optimized in both networks). One can interpret that the network size and noise level are competing to maintain a constant signal to noise ratio. Nevertheless, this interpretation is not necessarily correct in the context of signal propagation through an FFN with all-to-all connectivity. For example, the performance of signal representation in the **second layer** of an FFN with $N = 400$ & $\sigma = 25$ pA is almost equivalent with that of $N = 750$ & $\sigma = 40$ pA (see **Figure 1B**). However, despite the similarity within the two layers, the firing rate of the latter degrades in the subsequent layers. Therefore, a specific range of parameters—network size, synaptic strength, and level of synaptic noise in an FFN—should be identified to cooperatively maintain a consistent representation of the common (slow) signal across multiple layers.

Signal Propoagation in Neural Networks With Recurrent Connections

Despite the focus of this paper on the use of networks with feed-forward connections, recent studies demonstrated the feasibility of reliable signal propagation in recurrent networks (Joglekar et al., 2018; Barral et al., 2019). Of note in these studies is that lateral or feedback connectivity are biologically plausible architectures to transmit information between cortical layers (Stroud and Vogels, 2018). Joglekar et al. (2018) showed that reliable signal propagation could be achieved in large-scale recurrent network models of the macaque cortex. It is worth mentioning that the activity dynamics of recurrent networks can be compatible with those of FFNs (Kumar et al., 2010). Kumar et al. (2010) used FFNs as a part of a recurrent network to study reliable signal propagation in a biologically plausible scenario. In the context of visual perception, a recent study showed that a recurrent neural network with a few layers can be unfolded as very deep FFNs (Liao and Poggio, 2016; Rajaei et al., 2019). In addition, other studies illustrated that recurrent random networks may behave similar to an FFN (Ganguli et al., 2008; Goldman, 2009; Murphy and Miller, 2009; Kumar et al., 2010).

Thus, studying conditions of reliable signal propagation in an FFN helps better understanding the underlying mechanisms of information propagation in more biologically realistic scenarios in which the dynamics of an FFN interact with that of the embedding recurrent network.

Impact of Heterogeneous Synaptic Delays on Signal Transmission in Recurrent Neural Network

We showed that the impact of heterogeneous synaptic delays on the propagation of asynchronous spikes in an FFN was not significant. However, the heterogeneity in synaptic delays can alter the dynamics of a recurrent neural network and have a substantial effect on signal transmission. Here, we discuss three scenarios that capitalize on the substantial but ambiguous impact of the variability of synaptic delays in signal propagation. First, the heterogeneity in synaptic delays might have contrasting effects if applied to excitatory and inhibitory neurons. In a strongly recurrent network with excitatory synapses, this heterogeneity can reduce the likelihood of synchronization. In contrast, it can alter the tight balance of excitatory and inhibitory inputs—necessary for a reliable signal propagation—in a recurrent network with lateral inhibition (Joglekar et al., 2018; Stroud and Vogels, 2018). Second, the variability of synaptic delays might be compromised by that caused by synaptic transmissions. A recent study demonstrated that a network with non-instantaneous synaptic transmission and fixed spike delivery delay is equivalent to a network with a proper distribution of spike delays and instantaneous synaptic transmission (Mattia et al., 2019). Thus, a network with various types of synapses is differently influenced by the heterogeneity of synaptic delays. For example, the impact of the variability of synaptic delays (e.g., std $\sim$ 1 msec) can be more significant in a recurrent network with AMPA receptors (with a short time constant) compared to that with NMDA receptors (with a large time constant). Third, the variability of synaptic delays can be regulated by short- and long-term forms of synaptic plasticity. It was experimentally demonstrated that the synaptic latency at monosynaptically connected pairs of L5 and CA3 pyramidal neurons is inversely correlated with the amplitude of the postsynaptic current and sensitive to manipulations of the presynaptic release probability (Boudkkazi et al., 2007). Therefore, incorporating models of synaptic plasticity in the simulation of neural networks can induce another source of heterogeneity in synaptic delays, which in turn, alters signal propagation.

Optimal Level of Synaptic Noise and Stochastic Resonance

Stochastic resonance in the neural systems was first observed in the response of the neuronal network to a weak periodic signal (Gluckman et al., 1996) [see (Bulsara et al., 1991; Longtin, 1993) for stochastic resonance in neuron models]. In stochastic resonance, a proper amount of noise can boost signal representation. Noise in our model recreated the effect of the background synaptic noise that exists *in vivo* (Destexhe et al., 2001). Previous studies, e.g., (van Rossum et al., 2002; Kumar et al., 2010), demonstrated that background synaptic noise may balance the synchronizing effect of shared connectivity

in an FFN, thereby enabling the propagation of asynchronous firing rates. Unlike other studies that the average firing rate of asynchronous spikes is the sole information that must remain unchanged across layers, we used a time-varying slow signal as an input to an FFN to explore to what extent information of slow fluctuations can be transmitted across layers. In this study, although the input was not fully periodic, we found that a certain range of noise (see Results) enables maximum information transfer across layers. Interestingly, this range of noise (std of membrane fluctuations (no stimulus) $\approx$ 4 mV for std of noise = 25 pA) is in agreement with the level of background synaptic noise observed *in-vivo* high-conductance state (Destexhe et al., 2003). Thus, from a system level perspective and in the context of consistent information transfer, one can conclude that stochastic resonance might occur in an FFN with the optimal level of background synaptic noise (see also the effect of network topologies on stochastic resonance in FFNs Zhao et al., 2018).

Optimal Network Size and System-Size Coherent Resonance

Similar to noise-induced resonance, an optimal number of elements in a biological system can maximize the regularity in the emitted signal (in the presence of optimal level of noise), i.e., system-size coherence resonance (Toral et al., 2003). Toral et al. (2003) demonstrated that there is a coherence resonance effect, in the sense of maximum regularity in the signal generated by an ensemble of globally coupled Fitzhugh-Nagumo models, as a function of the number of coupled neurons, namely, N, in the presence of noise. It was shown that, for a specific set of parameters, the maximum regularity occurs for N $\approx$160 (Toral et al., 2003). A coherence resonance with respect to the number of neurons may exist in an FFN. Both the stochastic resonance and coherence resonance (CR) were investigated in the triple-neuron feed-forward-loop network motifs (Guo and Li, 2009). It was demonstrated that noise could enrich the stochastic dynamics of those motifs. In this study, we showed that for specific distribution of synaptic strength and optimal level of background synaptic noise, a reliable propagation of the time-varying rate of asynchronous spikes occurs for medium size FFNs (in the rang of {400–500}). Despite differences in the types of inputs, the architecture of an FFN vs. excitatory/inhibitory coupled neurons, propagation vs. representation, and other factors that resulted in two different values underlying optimum network size [compared to Toral et al. (2003)], our study capitalized on the significance of system-size coherence resonance in neuronal dynamics.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author. Codes are available at https://github.com/nsbspl/Necessary-Conditions-for-Reliable-Propagation.

AUTHOR CONTRIBUTIONS

NH: Conceptualization and analysis, programming, and writing the manuscript. MR and SF: Analysis and programming and contributing in the writing of the revised manuscript. MP: Supervising the study and revising the manuscript. ML: Conceptualization and analysis, design the study, programming, and writing the manuscript.

FUNDING

The present study was supported by Dr. Lankarany's start-up Grant and NSERC Discovery Grant (RGPIN-2020-05868).

ACKNOWLEDGMENTS

The authors declare that some preliminary results of this paper have been released as a pre-print (Lankarany, 2019).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncom.2020.00064/full#supplementary-material

REFERENCES

Abeles, M., Prut, Y., Bergman, H., and Vaadia, E. (1994). "Synchronization in neuronal transmission and its importance for information processing," in *Progress in Brain Research, vol. 102*, eds J. Van Pelt, M. A. Corner, H. B. M. Uylings, and F. H. Lopes Da Silva (Elsevier), 395–404. doi: 10.1016/S0079-6123(08)60555-5

Barral, J., Wang, X.-J., and Reyes, A., D. (2019). Propagation of temporal and rate signals in cultured multilayer networks. *Nat. Commun.* 10:3969. doi: 10.1038/s41467-019-11851-0

Boudkkazi, S., Carlier, E., Ankri, N., Caillard, O., Giraud, P., Fronzaroli-Molinieres, L., et al. (2007). Release-dependent variations in synaptic latency: a putative code for short- and long-term synaptic dynamics. *Neuron* 56, 1048–1060. doi: 10.1016/j.neuron.2007.10.037

Bowman, A. W., and Azzalini, A. (1997). Azzalini, *Applied Smoothing Techniques for Data Analysis: The Kernel Approach With S-Plus Illustrations*. Oxford : Clarendon Press, Oxford University Press.

Bulsara, A., Jacobs, E. W., Zhou, T., Moss, F., and Kiss, L. (1991). Stochastic resonance in a single neuron model: theory and analog simulation. *J. Theor. Biol.* 152, 531–555. doi: 10.1016/S0022-5193(05)80396-0

Chan, H. K., Yang, D.-P., Zhou, C., and Nowotny, T. (2016). Burst firing enhances neural output correlation. *Front. Comput. Neurosci.* 10:42. doi: 10.3389/fncom.2016.00042

Cortes, N., and van Vreeswijk, C. (2015). Pulvinar thalamic nucleus allows for asynchronous spike propagation through the cortex. *Front. Comput. Neurosci.* 9:60. doi: 10.3389/fncom.2015.00060

Destexhe, A., Rudolph, M., Fellous, J.-M., and Sejnowski, T. J. (2001). Fluctuating synaptic conductances recreate *in vivo*-like activity in neocortical neurons. *Neuroscience* 107, 13–24. doi: 10.1016/S0306-4522(01)00344-X

Destexhe, A., Rudolph, M., and Paré, D. (2003). The high-conductance state of neocortical neurons *in vivo*. *Nat. Rev. Neurosci.* 4, 739–751. doi: 10.1038/nrn1198

Diesmann, M., Gewaltig, M.-O., and Aertsen, A. (1999). Aertsen. Stable propagation of synchronous spiking in cortical neural networks. *Nature* 402, 529–533. doi: 10.1038/990101

Faraz, S., Mellal, I., and Lankarany, M. (2020). Impact of synaptic strength on propagation of asynchronous spikes in biologically realistic feed-forward neural network. *IEEE J. Sel. Top. Signal Process.* 99. doi: 10.1109/JSTSP.2020.2983607

Ganguli, S., Huh, D., and Sompolinsky, H. (2008). Memory traces in dynamical systems. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18970–18975. doi: 10.1073/pnas.0804451105

Gluckman, B. J., Netoff, T. I., Neel, E. J., Ditto, W. L., Spano, M. L., and Schiff, S. J. (1996). Stochastic resonance in a neuronal network from mammalian brain. *Phys. Rev. Lett.* 77, 4098–4101. doi: 10.1103/PhysRevLett.77.4098

Goldman, M. S. (2009). Memory without feedback in a neural network. *Neuron* 61, 621–634. doi: 10.1016/j.neuron.2008.12.012

Guo, D., and Li, C. (2009). Stochastic and coherence resonance in feed-forward-loop neuronal network motifs. *Phys. Rev. E* 79:051921. doi: 10.1103/PhysRevE.79.051921

Han, D., De Schutter, E., and Hong, S. (2019). Lamina-specific neuronal properties promote robust, stable signal propagation in feedforward networks. *bioRxiv [preprint]*. bioRxiv 596676. doi: 10.1101/596676

Joglekar, M. R., Mejias, J. F., Yang, G. R., and Wang, X. J. (2018). Inter-areal balanced amplification enhances signal propagation in a large-scale circuit model of the primate cortex. *Neuron* 98, 222–234.e8. doi: 10.1016/j.neuron.2018.02.031

Kremkow, J., Aertsen, A., and Kumar, A. (2010). Gating of signal propagation in spiking neural networks by balanced and correlated excitation and inhibition. *J. Neurosci.* 30, 15760–15768. doi: 10.1523/JNEUROSCI.3874-10.2010

Kumar, A., Rotter, S., and Aertsen, A. (2008). Conditions for propagating synchronous spiking and asynchronous firing rates in a cortical network model. *J. Neurosci.* 28, 5268–5280. doi: 10.1523/JNEUROSCI.2542-07.2008

Kumar, A., Rotter, S., and Aertsen, A. (2010). Spiking activity propagation in neuronal networks: reconciling different perspectives on neural coding. *Nat. Rev. Neurosci.* 11, 615–627. doi: 10.1038/nrn2886

Lankarany, M. (2019). Necessary conditions for reliable representation of asynchronous spikes through a single-layered feedforward network. *bioRxiv [preprint]*. bioRxiv 538736. doi: 10.1101/538736

Lankarany, M., Al-Basha, D., Ratté, S., and Prescott, S. A. (2019). Differentially synchronized spiking enables multiplexed neural coding. *Proc. Natl. Acad. Sci. U.S.A.* 116, 10097–10102. doi: 10.1073/pnas.1812171116

Liao, Q., and Poggio, T. (2016). *Bridging the Gaps Between Residual Learning, Recurrent Neural Networks and Visual Cortex*. ArXiv160403640 Cs. Available Online at: http://arxiv.org/abs/1604.03640 (accessed May 01, 2020).

Litvak, V., Sompolinsky, H., Segev, I., and Abeles, M. (2003). On the transmission of rate code in long feedforward networks with excitatory–inhibitory balance. *J. Neurosci.* 23, 3006–3015. doi: 10.1523/JNEUROSCI.23-07-03006.2003

London, M., Roth, A., Beeren, L., Häusser, M., and Latham, P. E. (2010). Sensitivity to perturbations *in vivo* implies high noise and suggests rate coding in cortex. *Nature* 466:123–127. doi: 10.1038/nature09086

Longtin, A. (1993). Stochastic resonance in neuron models. *J. Stat. Phys.* 70, 309–327. doi: 10.1007/BF01053970

Markov, N. T., Ercsey-Ravasz, M. M., Ribeiro Gomes, A. R., Lamy, C., Magrou, L., Vezoli, J., et al. (2014). A weighted and directed interareal connectivity matrix for macaque cerebral cortex. *Cereb. Cortex* 24, 17–36. doi: 10.1093/cercor/bhs270

Mattia, M., Biggio, M., Galluzzi, A., and Storace, M. (2019). Dimensional reduction in networks of non-Markovian spiking neurons: equivalence of synaptic filtering and heterogeneous propagation delays. *PLoS Comput. Biol.* 15:e1007404. doi: 10.1371/journal.pcbi.1007404

Mehring, C., Hehl, U., Kubo, M., Diesmann, M., and Aertsen, A. (2003). Activity dynamics and propagation of synchronous spiking in locally connected random networks. *Biol. Cybern* 88, 395–408. doi: 10.1007/s00422-002-0384-4

Montemurro, M. A., Panzeri, S., Maravall, M., Alenda, A., Bale, M., Peterson, R. S. et al. (2007). Role of precise spike timing in coding of dynamic vibrissa stimuli in somatosensory thalamus. *J. Neurophysiol.* 98, 1871–1882. doi: 10.1152/jn.00593.2007

Murphy, B. K., and Miller, K. D. (2009). Balanced amplification: a new mechanism of selective amplification of neural activity patterns. *Neuron* 61, 635–648. doi: 10.1016/j.neuron.2009.02.005

Noble, B. (2019). *Spikes: Exploring the Neural Code / Edition 1*. Barnes & Noble. Available Online at: https://www.barnesandnoble.com/w/spikes-fred-rieke/1117355212 (accessed Feburary 01, 2019).

Panzeri, S., Harvey, C. D., Piasini, E., Latham, P. E., and Fellin, T. (2017). Cracking the neural code for sensory perception by combining statistics, intervention, and behavior. *Neuron* 93, 491–507. doi: 10.1016/j.neuron.2016.12.036

Panzeri, S., Petersen, R. S., Schultz, S. R., Lebedev, M., and Diamond, M. E. (2001). The role of spike timing in the coding of stimulus location in rat somatosensory cortex. *Neuron* 29, 769–777. doi: 10.1016/S0896-6273(01)00251-3

Perez-Cruz, F. (2008). "Kullback-Leibler divergence estimation of continuous distributions," in *2008 IEEE International Symposium on Information Theory* (IEEE: Toronto, ON), 1666–1670. doi: 10.1109/ISIT.2008.4595271

Rajaei, K., Mohsenzadeh, Y., Ebrahimpour, R., and Khaligh-Razavi, S.-M. (2019). Beyond core object recognition: Recurrent processes account for object recognition under occlusion. *PLoS Comput. Biol.* 15:e1007001. doi: 10.1371/journal.pcbi.1007001

Runyan, C. A., Piasini, E., Panzeri, S., and Harvey, C. D. (2017). Distinct timescales of population coding across cortex. *Nature* 548, 92–96. doi: 10.1038/nature23020

Shadlen, M. N., and Newsome, W. T. (1994). Noise, neural codes and cortical organization. *Curr. Opin. Neurobiol.* 4, 569–579. doi: 10.1016/0959-4388(94)90059-0

Shimazaki, H., and Shinomoto, S. (2010). Kernel bandwidth optimization in spike rate estimation. *J. Comput. Neurosci.* 29, 171–182. doi: 10.1007/s10827-009-0180-4

Stroud, J. P., and Vogels, T. P. (2018). Cortical signal propagation: balance, amplify, transmit. *Neuron* 98, 8–9. doi: 10.1016/j.neuron.2018.03.028

Timme, N. M., and Lapish, C. (2018). A tutorial for information theory in neuroscience. *Eneuro* 5, 0052–18. doi: 10.1523/ENEURO.0052-18.2018

Toral, R., Mirasso, C. R., and Gunton, J. D. (2003). System size coherence resonance in coupled fitzhugh-nagumo models. *Europhys. Lett.* 61, 162–167. doi: 10.1209/epl/i2003-00207-5

van Rossum, M. C. W., Turrigiano, G. G., and Nelson, S. B. (2002). Fast propagation of firing rates through layered networks of noisy neurons. *J. Neurosci.* 22, 1956–1966. doi: 10.1523/JNEUROSCI.22-05-01956.2002

Wang, S., Wang, W., and Liu, F. (2006). Propagation of firing rate in a feed-forward neuronal network. *Phys. Rev. Lett.* 96:018103. doi: 10.1103/PhysRevLett.96.018103

Zhao, J., Qin, Y.-M., and Che, Y.-Q. (2018). Effects of topologies on signal propagation in feedforward networks. *Chaos Interdiscip. J. Nonlinear Sci.* 28:013117. doi: 10.1063/1.4999996

Zuo, Y., Safaai, H., Notaro, G., Mazzoni, A., Panzeri, S., and Diamond, M. E. (2015). Complementary contributions of spike timing and spike rate to perceptual decisions in rat S1 and S2 cortex. *Curr. Biol.* 25, 357–363. doi: 10.1016/j.cub.2014.11.065

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.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 07 August 2020
doi: 10.3389/fncom.2020.00062

Optimal Self-Induced Stochastic Resonance in Multiplex Neural Networks: Electrical vs. Chemical Synapses

Marius E. Yamakou[1,2]*, *Poul G. Hjorth*[2] *and Erik A. Martens*[2,3,4]

[1] *Max-Planck-Institut für Mathematik in den Naturwissenschaften, Leipzig, Germany,* [2] *Department of Applied Mathematics and Computer Science, Technical University of Denmark, Lyngby, Denmark,* [3] *Department of Biomedical Science, University of Copenhagen, Copenhagen, Denmark,* [4] *Centre for Translational Neuromedicine, University of Copenhagen, Copenhagen, Denmark*

Electrical and chemical synapses shape the dynamics of neural networks, and their functional roles in information processing have been a longstanding question in neurobiology. In this paper, we investigate the role of synapses on the optimization of the phenomenon of self-induced stochastic resonance in a delayed multiplex neural network by using analytical and numerical methods. We consider a two-layer multiplex network in which, at the intra-layer level, neurons are coupled either by electrical synapses or by inhibitory chemical synapses. For each isolated layer, computations indicate that weaker electrical and chemical synaptic couplings are better optimizers of self-induced stochastic resonance. In addition, regardless of the synaptic strengths, shorter electrical synaptic delays are found to be better optimizers of the phenomenon than shorter chemical synaptic delays, while longer chemical synaptic delays are better optimizers than longer electrical synaptic delays; in both cases, the poorer optimizers are, in fact, worst. It is found that electrical, inhibitory, or excitatory chemical multiplexing of the two layers having only electrical synapses at the intra-layer levels can each optimize the phenomenon. Additionally, only excitatory chemical multiplexing of the two layers having only inhibitory chemical synapses at the intra-layer levels can optimize the phenomenon. These results may guide experiments aimed at establishing or confirming to the mechanism of self-induced stochastic resonance in networks of artificial neural circuits as well as in real biological neural networks.

Keywords: optimization, self-induced stochastic resonance, synapses, multiplex neural network, community structure

OPEN ACCESS

Edited by:
Tatyana Sharpee,
Salk Institute for Biological Studies, United States

Reviewed by:
Ergin Yilmaz,
Bulent Ecevit University, Turkey
Dibakar Ghosh,
Indian Statistical Institute, India

***Correspondence:**
Marius E. Yamakou
yamakoumarius@gmail.com

Received: *01 April 2020*
Accepted: *28 May 2020*
Published: *07 August 2020*

Citation:
Yamakou ME, Hjorth PG and Martens EA (2020) Optimal Self-Induced Stochastic Resonance in Multiplex Neural Networks: Electrical vs. Chemical Synapses. Front. Comput. Neurosci. 14:62. doi: 10.3389/fncom.2020.00062

1. INTRODUCTION

Noise is an inherent part of neuronal dynamics, and its effects can be observed experimentally in neuronal activity at different spatiotemporal scales, e.g., at the level of ion channels, neuronal membrane potentials, local field potentials, and electroencephalographic or magnetoencephalographic measurements (Guo et al., 2018). While noise is mostly undesirable in many systems, it is now widely accepted that its presence is crucial to the proper functioning of neurons in terms of their information processing capabilities.

Some mechanisms for optimal information processing are provided via the well-known and extensively studied phenomena of stochastic resonance (SR) (Benzi et al., 1981; Longtin, 1993; Gammaitoni et al., 1998; Lindner et al., 2004; Zhang et al., 2015) and coherence resonance (CR) (Hu and MacDonald, 1993; Neiman et al., 1997; Pikovsky and Kurths, 1997; Lindner and Schimansky-Geier, 1999; Lindner et al., 2004; Beato et al., 2007; Hizanidis and Schöll, 2008; Liu et al., 2010; Bing et al., 2011; Gu et al., 2011) or via the lesser-known phenomenon of self-induced stochastic resonance (SISR) (Freidlin, 2001; Muratov et al., 2005; DeVille and Vanden-Eijnden, 2007; DeVille et al., 2007; Yamakou and Jost, 2017, 2018) whose mechanism remains to be confirmed experimentally in real neural systems. Although these noise-induced phenomena may exhibit similar dynamical behaviors, each of them has different dynamical preconditions and emergent mechanisms and may therefore play different functional roles in information processing. For further details behind the mechanisms of SR and CR, we refer the reader to references given above. We also note that the control of SR and CR in neural networks has attracted a lot of attention. In particular, it has been shown that hybrid synapses and autapses (i.e., those characterized by both electrical and chemical coupling) could be effectively used to control SR and CR (Yilmaz et al., 2013, 2016).

In this paper, we focus on self-induced stochastic resonance (SISR). SISR can occur when a multiple timescale excitable dynamical system is driven by vanishingly small noise. During SISR, the escape time of trajectories from one attracting region in phase space to another is distributed exponentially, and the associated transition frequency is governed by an activation energy. Suppose the system describing the neuron is placed out of equilibrium, and its activation energy decreases monotonically as the neuron relaxes slowly to a stable quiescent state (fixed point); then, at a specific instant during the relaxation, the timescale of escape events and the timescale of relaxation match, and the neuron almost surely fires at this point. If this activation brings the neuron back out-of-equilibrium, the relaxation stage can start over again, and the scenario repeats itself indefinitely, leading to a cyclic coherent spiking of the neuron which cannot occur without noise. SISR essentially depends on (i) strong timescale separation between the dynamical variables; (ii) vanishingly small noise amplitude; (iii) a monotonic activation energy barrier; (iv) and, most importantly, the periodic matching of the slow timescale of neuron's dynamics to the timescale characteristic to the noise. Thus, compared to CR and SR, the conditions to be met for observing SISR are more subtle: Like CR, SISR does not require an external periodic signal as in SR. Remarkably, unlike CR, SISR does not require the neuron's parameters be close to the bifurcation thresholds, making it more robust to parameter tuning than CR. Moreover, unlike both SR and CR, SISR requires a strong timescale separation between the neuron's dynamical variables.

The mechanism behind SISR suggests that, in an excitable neuron, the level of noise embedded in the neuron's synaptic input may be decoded into a (quasi-) deterministic and coherent signal. To exemplify, in a network of neurons in a quiescent state (without any activity), the action of a sufficiently weak synaptic noise amplitude could occasionally generate a spike in each neuron. These spikes will have random phases so that their total input on each individual neuron may average to a stationary random signal of low intensity. If the noise amplitude suddenly increases due to a change in the synaptic input, the neurons may switch to the noise-assisted oscillatory mode. This can further increase the effective noise amplitude so that the oscillatory mode may persist even after the disturbance is removed and the entire neural network in a dormant state may wake up from the outside rattle. The phenomenon of SISR in neural networks could therefore play important functional roles in the regulation of the Sleep-wake transition (Patriarca et al., 2012; Booth and Behn, 2014; Pereda, 2014).

Communication between neurons occurs through synaptic interactions. Two main types of synapses may be identified in neural networks, electrical synapses and chemical synapses (Pereda, 2014). The corresponding functional form of the bidirectional interaction mediated by the electrical synapses is defined as the difference between the membrane potentials of two adjacent neurons, thereby making the coupling mediated by electrical synapses to be local. Meanwhile, chemical synaptic interaction always takes place unidirectionally, with the signal conveyed chemically via neurotransmitter molecules through the synapses, thereby making chemical synaptic couplings nonlocal. The functional form of the chemical synaptic interaction is considered as a nonlinear sigmoidal input-output function (Greengard, 2001). Moreover, chemical synapses can be inhibitory or excitatory. When an inhibitory pre-synaptic neuron spikes, the post-synapses neuron connected to it is prevented from spiking. When an excitatory neuron spikes, it induces the post-synaptic neuron to spike. In real biological neurons, the distance between pre- and post-synaptic ends is approximately 3.5 nm in electrical synapses, and it is comparatively large, nearly 20–40 nm (Hormuzdi et al., 2004), in chemical synapses. Distances between pre- and post-synaptic ends induce time delays in neural networks with the time delays of electrical synapses being generally shorter than those of chemical synapses.

It is well-known from magnetic resonance imaging that neural networks may exhibit several types of coupling schemes: neurons coupled via electrical synapses only; neurons coupled via chemical synapses only; and neurons coupled by both electrical and chemical synapses—so-called hybrid synapses (Galarreta and Hestrin, 1999, 2001; Gibson et al., 1999; Connors and Long, 2004; Hestrin and Galarreta, 2005; Yilmaz et al., 2013; Bera et al., 2019; Majhi et al., 2019). Moreover, multiplex networks of neurons can be formed from different network layers depending on their connectivity through a chemical link or by an ionic channel. In brain networks, different regions can be seen connected by functional and structural neural networks (Pisarchik et al., 2014; De Domenico, 2017; Andreev et al., 2018). In a multiplex network, each type of interaction between the nodes is described by a single layer network and the different layers of networks describe the different modes of interaction. Multilayer networks (Pisarchik et al., 2014) open up new possibilities of optimization, allowing to regulate neural information processing by means of the interplay between the neurons' dynamics and multiplexing (Crofts et al., 2016; Battiston

et al., 2017). Optimization based on multiplexing could have many advantages. In particular, the coherent spiking activity of one layer (induced for example by SISR) can be optimized by adjusting the parameters of another layer. This is important from the point of view of engineering and brain surgery since it is not always possible to directly access the desired layer, though the network with which this layer is multiplexed may be accessible and adaptable.

Several studies have shown that multiplex networks can generate patterns with significant differences from those observed in single-layer networks (Kouvaris et al., 2015; Majhi et al., 2016, 2017; Berner et al., 2020). Their use in the optimization and control of dynamical behaviors have therefore attracted much attention recently. The multiplexing of networks has been shown to control many dynamical behaviors in neural networks, including synchronization (Gambuzza et al., 2015; Singh et al., 2015; Andrzejak et al., 2017; Leyva et al., 2017; Zhang et al., 2017), pattern formation (Kouvaris et al., 2015; Ghosh and Jalan, 2016; Ghosh et al., 2016, 2018; Maksimenko et al., 2016; Bera et al., 2017; Bukh et al., 2017), solitary waves (Mikhaylenko et al., 2019), and chimera states (Panaggio and Abrams, 2015; Schöll, 2016; Ghosh et al., 2018, 2019; Omelchenko et al., 2018; Sawicki et al., 2019). Chimera states are synchronization patterns occurring in symmetric networks (on average), characterized by the coexistence of varying synchronization levels side-by-side. They have been shown to exist in mechanical and chemical experiments (Tinsley et al., 2012; Martens et al., 2013; Totz et al., 2017) and are thought play an important role in neural systems (Andrzejak et al., 2016; Bera et al., 2019; Majhi et al., 2019). In particular, synchronization patterns such as chimera states occur in networks with community structure where connections are all-to-all, but coupling strengths are modulated so that the inter-coupling between communities (layers) are weak/sparse compared to their intra-coupling (Abrams et al., 2008; Martens et al., 2016a,b; Bick et al., 2020)—a configuration that bears strong similarity with the multilayer structure. Chimera states in such networks are of interest as they are multistable (Martens, 2010) and thus configurable; they can in principle be employed to solve functional tasks such as computations (Bick and Martens, 2015) and routing of information (Deschle et al., 2019) in the brain. Moreover, community networks of QIF neurons exhibit synchronization patterns that have been demonstrated viable for memory storage and recall (Schmidt et al., 2018). However, the optimization of noise-induced resonance mechanisms in neural networks based on the multiplexing approach have only very recently attracted attention. The few research works investigating the optimization of CR in neural networks are those of Semenova and Zakharova (2018) and Yamakou and Jost (2019).

In Semenova and Zakharova (2018), it is shown that connecting a one-layer network exhibiting CR in a multiplex way to another one-layer network, i.e., multiplexing, allows us to control CR in the latter layer network. In particular, it is found that multiplexing induces CR in networks that do not demonstrate this phenomenon in isolation. Moreover, it has been shown that CR can be achieved even for weak multiplexing between the layers. Surprisingly, it has also been shown that the multiplex-induced CR in the layer which is deterministic in isolation can manifest itself even more strongly than the CR in the noisy layer. However, the work in Semenova and Zakharova (2018) considers only instantaneous synaptic connections, while it is well known that synaptic time delays (not negligible in neural networks) exhibit crucial effects in neural information processing.

Yamakou and Jost (2019) considered synaptic time delays and their role in optimizing CR in a layer affected by another layer via multiplexing, which already exhibits optimal CR or SISR. In an isolated layer, it was shown that shorter synaptic time delays combined with weaker synaptic strengths optimize CR. Meanwhile, in the multiplex network configurations, stronger synaptic strengths combined with shorter synaptic time delays between layers induce and optimize CR in the layer where this phenomenon is non-existent in isolation. Moreover, their numerical simulations indicate that, even at very long multiplexing time delays, weak (but not too weak) multiplexing strengths between the layers can induce and optimize CR in the layer where it is non-existent in isolation. Interestingly, it was further shown that, with the occurrence of a different resonance phenomenon (i.e., SISR) in one layer, weak multiplexing, even at very short synaptic time delays, completely fails to optimize CR in the other layer where latter phenomenon does not exist in isolation. This behavior further confirms the fact that, even though SISR and CR lead to the occurrence of the same dynamical behavior (i.e., coherent noise-induced spiking activity) in neurons in the excitable regime, they are fundamentally different in their dynamical and emergent nature (DeVille et al., 2005); in particular, SISR and CR also lead to different behaviors in multiplex networks, and they possibly therefore play different functional roles in neural information processing.

The optimization of CR in neural networks based on the multiplexing approach have so far been studied only in Semenova and Zakharova (2018) and Yamakou and Jost (2019). A study on the optimization of SISR in neural networks based on the multiplexing approach is still lacking. Moreover, in Semenova and Zakharova (2018) and Yamakou and Jost (2019), the coupling between the neurons are mediated only by electrical synapses. The role of chemical synapses in the optimization of noise-induced resonance mechanisms should be equally important. Therefore, the aim of this paper is to study the optimization of SISR based on the multiplex approach of neural networks connected through time-delayed electrical and chemical synapses. In particular, we wish to address the following main questions:

(i) Can SISR occurring in one layer of a multiplex network be used to optimize SISR in another layer where the phenomenon non-existent in isolation?
(ii) What combinations of intra- and inter-layer synaptic strengths and time delays best optimize SISR?
(iii) Which type (electrical, inhibitory, or excitatory) of synapse is best optimizer SISR within an isolated layer and in the multiplex configuration?

The rest of the paper is organized as follows: in section 2, we present the mathematical model equations, and we explain and

motivate the different configurations considered. In section 3, we briefly describe the numerical methods used in simulations and analysis. In section 4, we consider an isolated single layer of neurons, coupled either by electrical synapses or chemical synapses. For both types of coupling, we analytically establish the necessary conditions in terms of noise amplitudes and timescale separation parameter that allow us to observe SISR. In section 5, we systematically investigate synaptic parameterizations that best optimize SISR in an isolated layer in which the neurons are coupled either by electrical synapses or by inhibitory chemical synapses. We will then compare the optimization of SISR by electrical and inhibitory chemical synapses. In section 6, we consider multiplexed layer networks using numerical simulations. Having identified which synaptic configurations deteriorate SISR the most in isolated layers, we use the multiplexing between a first layer, where SISR is optimal and a second layer where SISR is non-optimal (very poor or even non-existent), with the goal of optimizing SISR in the second layer. For multiplex networks, we will consider the optimization of SISR in six case scenarios: electrical, inhibitory, and excitatory multiplexing of two layers with electrical synaptic intra-connections and electrical, inhibitory, and excitatory multiplexing of two layers with inhibitory synaptic intra-connections. Finally, we summarize and conclude our findings in section 7.

2. MATHEMATICAL MODEL

We consider a two-layer multiplex neuronal network in the excitable regime in the presence of synaptic noise, as illustrated in **Figure 1**. In our study, we consider one of the simplest network topologies—a ring network topology within layers and a multiplex network between these layers such that they contain the same number of neurons and the interaction between the layers are allowed only for replica neurons. Each layer consists of N identical FitzHugh-Nagumo (FHN) neurons (Hodgkin and Huxley, 1952; FitzHugh, 1961), connected in a ring by either only electrical synapses or inhibitory chemical synapses, while the inter-connections between layers can be either via electrical, inhibitory chemical synapses, or excitatory chemical synapses. It is important to point out that excitatory chemical synapses are found to induce, via time-delayed coupling bifurcations, a self-sustained spiking activity in the network of FHN neurons (each in the excitable regime) even in the complete absence of noise. We want to avoid such regimes—those in which the deterministic network can oscillate due to some time-delayed coupling induced bifurcations—as the coherent oscillations induced by SISR should be due only to the presence of noise and not because of the occurrence of bifurcations. For this reason, the excitatory chemical synapses are used in the optimization of SISR only when they do not induce oscillatory behaviors in the deterministic network, i.e., only in the multiplexing connections with carefully chosen synaptic strengths and time delays.

Real electrical synapses mediate bidirectional interactions and transfer signals only between neighboring neurons; in contrast, chemical synapses convey information unidirectionally between distantly situated neurons. To account for this, the model implements layers with bidirectional electrical coupling with nearest neighbor interactions (**Figure 1A**), while unidirectional chemical coupling is implemented with nonlocal interactions, i.e., also including connections other than nearest neighbor interactions (**Figure 1B**). These coupling topologies and interaction modes are biologically relevant and will also allow us to compare the functional role played by chemical and electrical synaptic interactions in processing information generated during SISR.

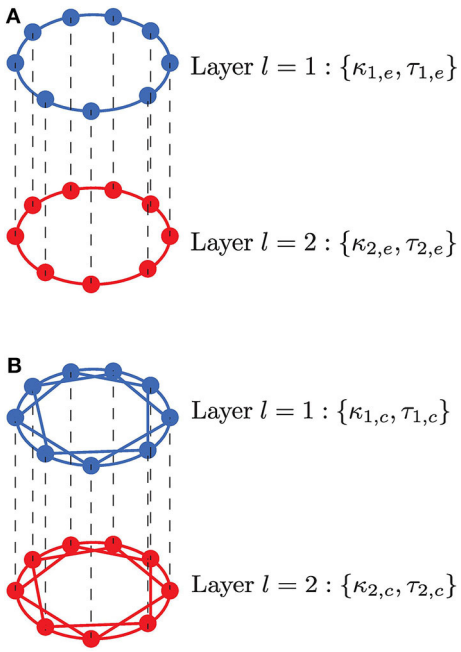

FIGURE 1 | Neurons are connected in a multiplex network with two layers $l = 1, 2$. Neurons within a layer are coupled in a ring, while neurons in adjacent layers are connected only to their adjacent neurons. **(A)** Neurons within each layer are coupled in a ring and interact only with their nearest neighbors via electrical synapses ($\kappa_{1,e}, \tau_{1,e}, \kappa_{2,e}, \tau_{2,e}$); multiplexing between these layers, represented by the black vertical dashed lines, may occur via electrical ($\kappa_{m,e}, \tau_{m,e}$) or (inhibitory or excitatory) chemical synapses ($\kappa_{m,c}, \tau_{m,c}$). **(B)** Neurons within each layer are coupled in a ring and interact with $n_{l,c}$ nearest neighbors via inhibitory chemical synapses ($\kappa_{1,c}, \tau_{1,c}, \kappa_{2,c}, \tau_{2,c}$); multiplexing between these layers, represented by the black vertical dashed lines, may occur via electrical ($\kappa_{m,e}, \tau_{m,e}$) or (inhibitory or excitatory) chemical ($\kappa_{m,c}, \tau_{m,c}$) synapses. In both scenarios, an enhanced SISR in layer $l = 1$ is used to optimize a poor or non-existent SISR in layer $l = 2$ by variation of the time-delay coupling parameters within a population ($\kappa_{1,e}, \tau_{1,e}, \kappa_{2,e}, \tau_{2,e}, \kappa_{1,c}, \tau_{1,c}, \kappa_{2,c}, \tau_{2,c}$) and between populations ($\kappa_{m,e}, \tau_{m,e}, \kappa_{m,c}, \tau_{m,c}$).

The stochastic differential equations resulting from this two-layer FHN neural network are given by

$$\begin{cases} \mathrm{d}v_{l,i} = \left(v_{l,i} - \dfrac{v_{l,i}^3}{3} - w_{l,i} + E_{l,i} + M_{l,i}^e - C_{l,i} - M_{l,i}^c\right)\mathrm{d}t \\ \qquad\qquad + \sigma_l \,\mathrm{d}W_{l,i}, \\ \mathrm{d}w_{l,i} = \varepsilon(v_{l,i} + \alpha - \beta w_{l,i})\,\mathrm{d}t, \end{cases} \tag{1}$$

where each neuron is represented by a node $i = 1, \ldots, N$ in the multiplex network with layers $l = 1, 2$, and the functional dependencies are given by,

$$\begin{cases} E_{l,i} &= \dfrac{\kappa_{l,e}}{2n_{l,e}} \displaystyle\sum_{j=i-n_{l,e}}^{i+n_{l,e}} \left(v_{l,j}(t-\tau_{l,e}) - v_{l,i}(t)\right), \\ C_{l,i} &= \dfrac{\kappa_{l,c}}{2n_{l,c}}(v_{l,i}(t) - V_{\text{syn}}) \displaystyle\sum_{j=i-n_{l,c}}^{i+n_{l,c}} \\ & \quad \left\{1 + \exp\left[-\lambda\left(v_{l,j}(t-\tau_{l,c}) - \Theta_{\text{syn}}\right)\right]\right\}^{-1}, \\ M^{e}_{1,i} &= \kappa_{m,e}\left(v_{2,i}(t-\tau_{m,e}) - v_{1,i}(t)\right), \\ M^{e}_{2,i} &= \kappa_{m,e}\left(v_{1,i}(t-\tau_{m,e}) - v_{2,i}(t)\right), \\ M^{c}_{1,i} &= \kappa_{m,c}\left(v_{1,i}(t) - V_{\text{syn}}\right) \\ & \quad \left\{1 + \exp\left[-\lambda\left(v_{2,i}(t-\tau_{m,c}) - \Theta_{\text{syn}}\right)\right]\right\}^{-1}, \\ M^{c}_{2,i} &= \kappa_{m,c}\left(v_{2,i}(t) - V_{\text{syn}}\right) \\ & \quad \left\{1 + \exp\left[-\lambda\left(v_{1,i}(t-\tau_{m,c}) - \Theta_{\text{syn}}\right)\right]\right\}^{-1}. \end{cases} \tag{2}$$

We fixed the number of neurons per layer to $N = 25$ throughout this study. The membrane potential and the recovery current variables of neuron i in layer l are given by $v_{l,i} \in \mathbb{R}$ and $w_{l,i} \in \mathbb{R}$, respectively, and $0 < \varepsilon \ll 1$ sets the timescale separation between the fast membrane potential and the slow recovery current variables. The excitability threshold $\beta > 0$ of the neurons is a codimension-one Hopf bifurcation parameter. $\alpha \in (0, 1)$ is a constant parameter. The additive noise term $\mathrm{d}W_{l,i}$ represents mean-centered Gaussian noise with $\langle \mathrm{d}W_{l,i}(t)\, \mathrm{d}W_{l,i}(t')\rangle_t = \delta(t - t')$ and variance (strength) σ_l, and it models the synaptic fluctuations observed in neural networks.

$E_{l,i}$ represent the electrical synaptic interactions between neurons coupled within a ring layer network with strength $\kappa_{l,e}$ and time delay τ_{le}, respectively, and an interaction range set to $n_{l,e} = 1$ since electrical synapses interact only locally. The coupling mediated by electrical synapses is of diffusive type, i.e., the electrical coupling term (intra- or inter-layer) vanishes if $v_{1,i}$ and $v_{1,j}$ (resp. $v_{2,i}$ and $v_{2,j}$) or $v_{1,i}$ and $v_{2,i}$ are equal.

$M^{e}_{1,i}$ and $M^{e}_{2,i}$ represent the coupling between layers via electrical synapses (i.e., electrical multiplexing of layers) with strength $\kappa_{m,e}$ and delay $\tau_{m,e}$, respectively.

$C_{l,i}$ represent chemical synaptic interactions between neurons coupled within a layer with ring topology, with strength $\kappa_{l,c}$ and time delay $\tau_{l,c}$ and where $1 < n_{l,c} < (N-1)/2$ represents interaction range on the ring network layer; we fix $n_{l,e} = 8$ all through this paper. The chemical synaptic function is modeled by a sigmoidal input-output function, $\Gamma(v_i) = \dfrac{1}{1 + e^{-\lambda(v_i - \Theta_{\text{syn}})}}$ (see Equation 2 in Greengard, 2001), where parameter $\lambda = 10.0$ determines the slope of the function and $\Theta_{\text{syn}} = -0.25$ the synaptic firing threshold.

$M^{c}_{1,i}$ and $M^{c}_{2,i}$ represent the coupling between layers mediated by chemical synapses (i.e., chemical multiplexing of layers) with $\kappa_{m,c}$ and $\tau_{m,c}$ representing the strength and time delay, respectively. V_{syn} represents the synaptic reversal potential. For $V_{\text{syn}} < v_{l,i}(t)$, the chemical synaptic interaction has a depolarizing effect that makes the synapse inhibitory; for $V_{\text{syn}} > v_{l,i}(t)$, the synaptic interaction has a hyper-polarizing effect, making the synapse excitatory. For the version of the FHN neuron model used in this study, the membrane potentials $|v_{l,i}(t)| \leq 2.0$ ($l = 1, 2; i = 1, 2, \ldots, N$) for all time t. For the choice of fixed $V_{\text{syn}} = -3.0$ (maintained throughout our computations), the term $(v_{l,i}(t) - V_{\text{syn}})$ in Equation (2) is always positive. So, the inhibitory and excitatory natures of chemical synapses will depend only on the sign in front of the synaptic coupling strengths $\kappa_{l,c}$ and $\kappa_{m,c}$. To make the chemical synapse inhibitory, we chose a negative sign i.e., when the pre-synaptic neuron spikes, it prevents the post-synaptic neuron from spiking and, conversely, a positive sign for excitatory chemical synapses.

3. NUMERICAL METHODS

In our numerical simulations, we used the fourth-order Runge-Kutta algorithm for stochastic processes (Kasdin, 1995) to integrate over a very long time interval ($T = 600,000$ time units) to average time series over time with seven realizations for each noise amplitude. In the numerical simulations, this long time interval permitted us to collect with a small noise amplitude at least 125 interspike intervals with $\varepsilon = 0.0005 \ll 1$. Each network layer had $N = 25$ neurons.

To measure how pronounced SISR is, we used the coefficient of variation (R_T), which is an important statistical measure based on the time intervals between spikes. It measures the regularity of noise induced spiking and therefore a measure of how pronounced SISR can be at a particular noise amplitude. R_T exploits the inter-spike interval (ISI) where the mth interval is defined as the difference between two consecutive spike times t_i^m and t_i^{m+1} of neuron i in a network, namely $\text{ISI}_i = t_i^{m+1} - t_i^m > 0$. For the ith neuron, the ratio between the standard deviation and the mean defines the coefficient of variation of the ISIs over a time interval $[0, T]$ as (Pikovsky and Kurths, 1997):

$$R_{T_i} = \frac{\sqrt{\langle ISI_i^2\rangle - \langle ISI_i\rangle^2}}{\langle ISI_i\rangle}, \tag{3}$$

where $\langle ISI_i\rangle$ and $\langle ISI_i^2\rangle$ represent the mean and the mean squared inter-spike intervals of the ith neuron, respectively. The above definition of R_T is limited to characterizing SISR in an isolated neuron. For a network of coupled neurons, SISR can be measured by redefining R_T as follows (Masoliver et al., 2017):

$$R_T = \frac{\sqrt{\langle \overline{ISI^2}\rangle - \langle \overline{ISI}\rangle^2}}{\langle \overline{ISI}\rangle}, \tag{4}$$

with

$$\begin{cases} \langle \overline{ISI}\rangle = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N}\langle ISI_i\rangle, \\ \langle \overline{ISI^2}\rangle = \dfrac{1}{N}\displaystyle\sum_{i=1}^{N}\langle ISI_i^2\rangle, \end{cases} \tag{5}$$

where the extra bar indicates the additional average over the total number of neurons N in the layer.

Of course, other statistical measures exist such as the correlation time, the power spectral density, and the signal-to-noise ratio which are commonly used measures to quantify the coherence of noise induced spiking activity. However, from a neurobiological point of view, R_T is more important than the other measures because it is related to the timing precision of the information processing in neural systems (Pei et al., 1996). Because of R_T's importance in neural information processing, we shall use it to characterize the regularity of the noise-induced oscillations generated by SISR in our neural network. For a Poissonian spike train (rare and incoherent spiking), $R_T = 1$. If $R_T < 1$, the sequence becomes more coherent, and R_T vanishes for a periodic deterministic spike train. R_T values greater than 1 correspond to a point process that is more variable than a Poisson process (Kurrer and Schulten, 1995; Yamakou and Jost, 2018).

4. CONDITIONS FOR SISR IN ISOLATED LAYERS IN THE EXCITABLE REGIME

We first consider the case of isolated layers of the multiplex networks in **Figures 1A,B**. Thus, neurons in such an isolated layer are connected either only via electrical synapses or via chemical synapses. In particular, here we will establish the analytic conditions necessary for the emergence of the SISR in these isolated network layers of FHN neurons in the excitable regime. From these conditions, we will furthermore obtain the minimum and maximum noise amplitudes required for SISR to occur in an isolated layer.

For SISR to occur, it is necessary to be in the excitable parameter regime. The isolated FHN neuron has a unique and stable fixed point in this regime. Choosing an initial condition in the basin of attraction of this fixed point will result in at most one large non-monotonic excursion into the phase space after which the trajectory asymptotically approaches the fixed point and stays there until initial conditions are changed again (Izhikevich, 2000; Yamakou and Jost, 2018).

Considering the multiplex networks in **Figure 1** with disconnected layers ($\kappa_{m,e} = \kappa_{m,c} = 0$), we may place an isolated neuron ($\kappa_{l,e} = 0$ or $\kappa_{l,c} = 0$) into an excitable regime by fixing parameter $\alpha = 0.5$. The bifurcation parameter β is chosen such that $\beta > \beta_h(\varepsilon)$, where $\beta_h(\varepsilon)$ is defined as the Hopf bifurcation value of an isolated neuron. Fixing the timescale separation parameter value to $\varepsilon = 0.0005$, we calculate the Hopf bifurcation value to be $\beta_h(\varepsilon) = 0.7497$. It is important to note That, for $\beta \leq \beta_h(\varepsilon)$, an isolated neuron is in the oscillatory regime—a regime that we want to avoid since the coherent oscillations generated by SISR are due only to the presence of noise rather than to the occurrence of a Hopf bifurcation (Yamakou and Jost, 2018).

Moreover, we have to ensure that the network of coupled neurons as a whole stays in the excitable regime rather than just single neurons in isolation. Indeed, certain time-delayed couplings may induce self-sustained oscillations in a network layer even though the isolated neurons remain inside the excitable regime. In layers with excitatory chemical synapses, a saddle-node bifurcation onto a limit cycle may generate self-sustained oscillations induced via time-delayed couplings (Schöll et al., 2009). On the other hand, when used for the multiplexing of layers, some values of time delays and coupling strengths of the excitatory chemical synapses cannot provoke this saddle-node bifurcation. Therefore, we did not consider excitatory chemical synapses for the coupling of neurons within layers but rather only for the coupling between layers. We therefore need to make sure that neurons connected in each network layer stay outside the parameter regime where oscillations are induced by time-delayed coupling. First, we need to determine if such a regime exists and identify it.

Taking the limit $\varepsilon \to 0$ in the isolated layer $l = 1, 2$ ($\kappa_{m,e} = \kappa_{m,c} = 0$) for either electrical ($\kappa_{l,c} = 0$) or chemical synapses ($\kappa_{l,e} = 0$) only, the equations for each neuron in this layer reduces to coupled Langevin equations of the form,

$$\mathrm{d}v_{l,i} = -\frac{\partial U_i^{e,c}(v_{l,i}, w_{l,i})}{\partial v_{l,i}}\mathrm{d}t + \sigma_l \,\mathrm{d}W_{l,i}, \tag{6}$$

where the electrical $U_i^e(v_{l,i}, w_{l,i})$ and chemical $U_i^c(v_{l,i}, w_{l,i})$ interaction potentials ($i = 1, \ldots, N$) are double-well potentials given by Equation (7) and may be viewed as functions of $v_{l,i}$ where $w_{l,i}$ is nearly constant. **Figures 2**, **3**, respectively show the modulation of landscapes of electrical and chemical interaction potentials with changing synaptic strength.

$$\begin{cases} U_i^e(v_{l,i}, w_{l,i}) = \dfrac{1}{12}v_{l,i}^4 - \dfrac{1}{2}v_{l,i}^2 + v_{l,i}w_{l,i} - \dfrac{\kappa_{l,e}}{2n_{l,e}} \\ \qquad\qquad \displaystyle\sum_{j=i-n_{l,e}}^{i+n_{l,e}} \Big(v_{l,i}(t)v_{l,j}(t-\tau_{l,e}) - \frac{1}{2}v_{l,i}(t)^2\Big), \\ U_i^c(v_{l,i}, w_{l,i}) = \dfrac{1}{12}v_{l,i}^4 - \dfrac{1}{2}v_{l,i}^2 + v_{l,i}w_{l,i} \\ \qquad + \dfrac{\kappa_{l,c}}{2n_{l,c}} \displaystyle\sum_{j=i-n_{l,c}}^{i+n_{l,c}} \frac{1}{2}v_{l,i}(t)\big(v_{l,i}(t) - 2V_{\mathrm{syn}}\big) \\ \qquad\qquad \Big\{1 + \exp\Big[-\lambda\big(v_{l,j}(t-\tau_{l,c}) - \Theta_{\mathrm{syn}}\big)\Big]\Big\}^{-1}, \end{cases} \tag{7}$$

We observe three different behaviors for the electrical potential interaction $U_i^e(v_{l,i}, w_{l,i})$. (i) When $w_{l,i} < 0$, we find that $U_i^e(v_{l,i}, w_{l,i})$ is asymmetric with the shallower well on the left. The neuron is close to the stable homogeneous fixed point at $(v_{l,i}^*, w_{l,i}^*) = (-1.003975, -0.666651)$, and a spike consists of jumping over the left energy barrier $\triangle U^{le}(w_{l,i})$ into the right well (see **Figure 2A**). (ii) When $w_{l,i} = 0$, then $U_i^e(v_{l,i}, w_{l,i})$ is symmetric with $\triangle U^{le}(w_{l,i}) = \triangle U^{re}(w_{l,i})$, and the neuron is half way between the quiescent state and the spike state (see **Figure 2B**). (iii) When $w_{l,i} > 0$, then $U_i^e(v_{l,i}, w_{l,i})$ is also asymmetric. The neuron has spiked and a return to the quiescent state (the homogeneous fixed point) consists of jumping over the right energy barrier $\triangle U^{re}(w_{l,i})$ into the left well (see **Figure 2C**). The intra-layer electrical synapse $\kappa_{l,e}$ does not change the symmetry (or asymmetry) of the interaction potential $U_i^e(v_{l,i}, w_{l,i})$. It only changes the depth of the energy barriers. The stronger $\kappa_{l,e}$ is,

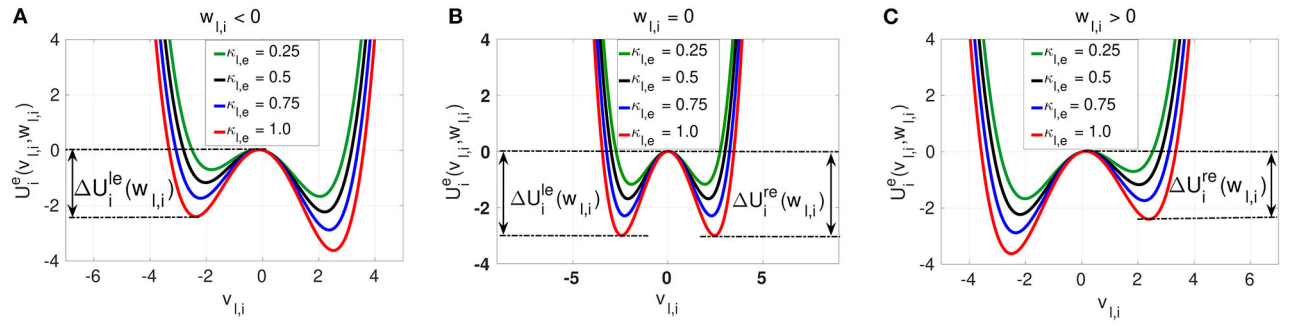

FIGURE 2 | The electrical interaction potential $U_i^e(v_{l,i}, w_{l,i})$ in Equation (7) is shown for a locally coupled ring network topology ($n_{l,e} = 1$) with the energy barriers for the asymmetric cases ($w_{l,i} \neq 0$) **(A,C)** and symmetric ($w_{l,i} = 0$) case **(B)**. The stronger the intra-layer synaptic strength $\kappa_{l,e}$ is, the deeper the energy barrier functions $\triangle U_i^{le}(w_{l,i})$ and $\triangle U_i^{re}(w_{l,i})$ are. The saddle point and the left and right minima of the interaction potential are located at $v_{l,i} = v_m^*(w_{l,i})$, $v_{l,i} = v_l^*(w_{l,i})$, and $v_{l,i} = v_r^*(w_{l,i})$, respectively.

the deeper the energy barrier functions $\triangle U^{le}(w_{l,i})$ and $\triangle U^{re}(w_{l,i})$ defined in Equation (9) are.

The chemical potential interaction $U_i^c(v_{l,i}, w_{l,i})$ shows richer landscape dynamics due to its stronger nonlinearity. We first notice that, just like with intra-layer electrical synaptic strength $\kappa_{l,e}$, intra-layer inhibitory chemical synaptic strength $\kappa_{l,c}$ changes the depth of the energy barriers $\triangle U_i^{lc}(w_{l,i})$ and $\triangle U_i^{rc}(w_{l,i})$. That is, the stronger $\kappa_{l,c}$ is, the deeper the energy barriers $\triangle U_i^{lc}(w_{l,i})$ and $\triangle U_i^{rc}(w_{l,i})$ are. In contrast to the electrical synaptic strength $\kappa_{l,e}$, the inhibitory chemical synaptic strength $\kappa_{l,c}$ is capable of changing the symmetry or (asymmetry) of the chemical potential $U_i^c(v_{l,i}, w_{l,i})$, where we distinguish the following cases: (i) When $w_{l,i} < 0$, then $U_i^c(v_{l,i}, w_{l,i})$ can be symmetric or asymmetric depending on the value of the inhibitory chemical synaptic strength $\kappa_{l,c}$. If $w_{l,i} < 0$ and $\kappa_{l,c} = 0.16$, we see from **Figure 3A** that $U_i^c(v_{l,i}, w_{l,i})$ is symmetric and becomes asymmetric as $\kappa_{l,c}$ changes. (ii) When $w_{l,i} = 0.0$, we do not have any symmetric chemical potential landscape as shown in **Figure 3B**, contrasting our observations for the electrical potential. (iii) For $w_{l,i} > 0$ (see **Figure 3C**), the chemical potential landscape is symmetric for $\kappa_{l,c} = 0.2$ and becomes asymmetric as $\kappa_{l,c}$ changes. Moreover, we notice that, for values of the chemical synaptic strength $\kappa_{l,c}$ for which the chemical interaction potential is symmetric, the energy barriers functions are shallower than in the symmetric case of the electrical potential. The important common feature of the electrical and inhibitory chemical potential is the deepening of the energy barriers $\triangle U_i^{le,c}(w_{l,i})$ and $\triangle U_i^{re,c}(w_{l,i})$ with increase in the intra-layer electrical $\kappa_{l,e}$, and inhibitory chemical $\kappa_{l,c}$ synaptic strengths shall explain why SISR is deteriorated by stronger intra-layer synaptic connections.

We choose parameters of the coupled neurons in Equation (6) such that they satisfy the conditions necessary for the occurrence of SISR. These conditions are adapted from those valid for an isolated FHN neuron (DeVille et al., 2005; Yamakou and Jost, 2018) so that they include (one at a time) the time-delayed electrical and inhibitory chemical synaptic connections between the FHN neurons coupled in a ring network. The resulting conditions are

$$\begin{cases} \lim\limits_{(\varepsilon,\sigma_l)\to(0,0)} \dfrac{\sigma_l^2}{2} \ln(\varepsilon^{-1}) \in \left(\triangle U_i^{le}(w_{l,i}^*), F_e(\kappa_{l,e}, \tau_{l,e}, n_{l,e})\right), \\ \lim\limits_{(\varepsilon,\sigma_l)\to(0,0)} \dfrac{\sigma_l^2}{2} \ln(\varepsilon^{-1}) \in \left(\triangle U_i^{lc}(w_{l,i}^*), F_c(\kappa_{l,c}, \tau_{l,c}, n_{l,c})\right), \\ \lim\limits_{(\varepsilon,\sigma_l)\to(0,0)} \dfrac{\sigma_l^2}{2} \ln(\varepsilon^{-1}) = \mathcal{O}(1), \\ \beta - \beta_h(\varepsilon) > 0, \end{cases} \tag{8}$$

where

$$\begin{cases} F_e(\kappa_{l,e}, \tau_{l,e}, n_{l,e}) : = \left\{(\kappa_{l,e}, \tau_{l,e}, n_{l,e}) : \triangle U_i^{le}(w_{l,i}) = \triangle U_i^{re}(w_{l,i})\right\}, \\ F_c(\kappa_{l,c}, \tau_{l,c}, n_{l,c}) : = \left\{(\kappa_{l,c}, \tau_{l,c}, n_{l,c}) : \triangle U_i^{lc}(w_{l,i}) \text{ or } \triangle U_i^{rc}(w_{l,i}) \right. \\ \qquad\qquad \left. \text{is maximum}\right\}, \\ \triangle U_i^{le,c}(w_{l,i}) : \quad = U_i^{e,c}\left(v_m^*(w_{l,i}), w_{l,i}\right) - U_i^{e,c}\left(v_l^*(w_{l,i}), w_{l,i}\right), \\ \triangle U_i^{re,c}(w_{l,i}) : \quad = U_i^{e,c}\left(v_m^*(w_{l,i}), w_{l,i}\right) - U_i^{e,c}\left(v_r^*(w_{l,i}), w_{l,i}\right), \end{cases} \tag{9}$$

with

$$v_{l,m,r}^*(w_{l,i}) : = \left\{v_{l,i} : v_{l,i} - \frac{v_{l,i}^3}{3} - w_{l,i} + \frac{\kappa_{l,e}}{2n_{l,e}} \sum_{j=i-n_{l,e}}^{i+n_{l,e}} \left(v_{l,j}(t-\tau_{l,e}) - v_{l,i}(t)\right) = 0\right\}, \tag{10}$$

for electrical synapses and

$$v_{l,m,r}^*(w_{l,i}) : = \left\{v_{l,i} : v_{l,i} - \frac{v_{l,i}^3}{3} - w_{l,i} - \frac{\kappa_{l,c}}{2n_{l,c}} \sum_{j=i-n_{l,c}}^{i+n_{l,c}} (v_{l,i} - V_{\text{syn}}) \right. \\ \left. \left\{1 + \exp\left[-\lambda\left(v_{l,j}(t-\tau_{m,c}) - \Theta_{\text{syn}}\right)\right]\right\}^{-1} = 0\right\}, \tag{11}$$

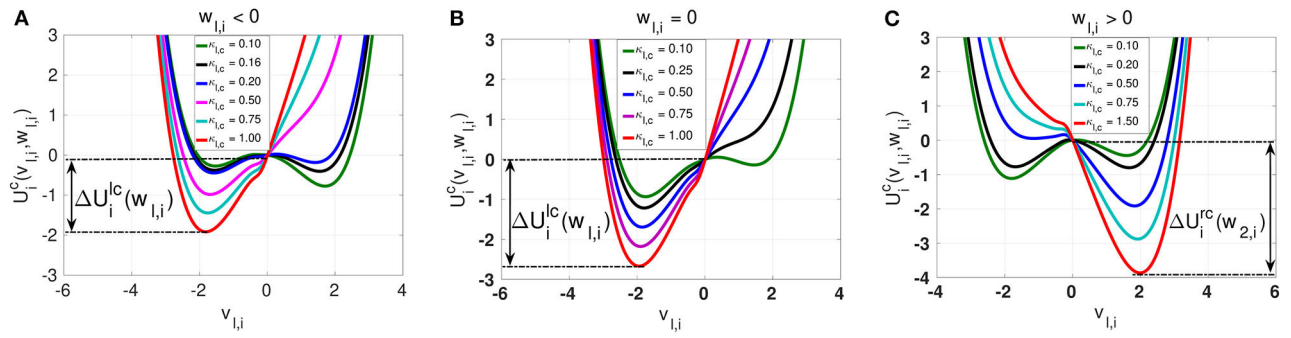

FIGURE 3 | Landscapes of the inhibitory chemical interaction potential $U_i^c(v_{l,i}, w_{l,i})$ in Eq. (7) for a non-locally ($n_{l,c} = 8$) coupled ring network topology. The symmetry of the potential is governed not by the slow variable $w_{l,i}$ as in the case of the electrical interaction potential, but by the chemical synaptic strength $\kappa_{l,c}$. In **(A)** with $w_{l,i} < 0$, the asymmetric potential can become symmetric when the synaptic strength is $\kappa_{l,c} = 0.16$. In **(B)** with $w_{l,i} = 0$, the potential remains asymmetric for all values of the synaptic strength $\kappa_{l,c}$. In **(C)** with $w_{l,i} > 0$, the asymmetric potential can again become symmetric when the synaptic strength is $\kappa_{l,c} = 0.20$. Similarly to the electrical synaptic strength, the stronger the intra-layer chemical synaptic strength $\kappa_{l,c}$ is, the deeper the energy barrier functions $\triangle U_i^{lc}(w_{l,i})$ and $\triangle U_i^{rc}(w_{l,i})$ are.

for chemical synapses. Furthermore, the solution sets of Equations (10) and (11) are such that $v_l^*(w_{l,i}) < v_m^*(w_{l,i}) < v_r^*(w_{l,i})$ define the left stable, middle unstable, and right stable branches of the cubic nullcline of each FHN neuron.

The energy barrier functions $\triangle U_i^{le,c}(w_{l,i})$ and $\triangle U_i^{re,c}(w_{l,i})$ can be obtained from the electrical interaction potential $U_i^e(v_{l,i}, w_{l,i})$ and the inhibitory chemical interaction potential $U_i^c(v_{l,i}, w_{l,i})$ by taking the difference between the potential function value at the saddle point $v_m^*(w_{l,i})$ and at the local minima $v_{l,r}^*(w_{l,i})$ of these interaction potentials (Yamakou and Jost, 2018). The energy barriers $\triangle U_i^{le}(w_{l,i}^*)$ or $\triangle U_i^{lc}(w_{l,i}^*)$ (which has to be crossed to induce a spike) is the value of the left energy barrier function at the $w_{l,i}$-coordinate of the stable homogeneous steady state $\left[v_{l,i}^*(\kappa_{l,e}, \tau_{l,e}, n_{l,e}), w_{l,i}^*(\kappa_{l,e}, \tau_{l,e}, n_{l,e})\right]$ or $\left[v_{l,i}^*(\kappa_{l,c}, \tau_{l,c}, n_{l,c}), w_{l,i}^*(\kappa_{l,c}, \tau_{l,c}, n_{l,c})\right]$, respectively. This is where the electrical $\triangle U_i^{le}\left[w_{l,i}^*(\kappa_{l,e}, \tau_{l,e}, n_{l,e})\right]$ and chemical $\triangle U_i^{lc}\left[w_{l,i}^*(\kappa_{l,c}, \tau_{l,c}, n_{l,c})\right]$ energy barrier functions get their $\kappa_{l,e}$, $\tau_{l,e}$, $n_{l,e}$ and $\kappa_{l,c}$, $\tau_{l,c}$, $n_{l,c}$ dependence from.

Now from the first two conditions of Equation (8), we obtain the noise amplitude range $[\sigma_l^{\min}, \sigma_l^{\max}]$ within which SISR occurs in the layer network of electrically (chemically) coupled FHN neurons:

$$\begin{cases} \sigma_l^{\min^e} = \sqrt{\dfrac{2\triangle U_i^{le}\left(w_{l,i}^*(\kappa_{l,e}, \tau_{l,e}, n_{l,e})\right)}{\ln(\varepsilon^{-1})}}, \\ \sigma_l^{\max^e} = \sqrt{\dfrac{2F_e(\kappa_{l,e}, \tau_{l,e}, n_{l,e})}{\ln(\varepsilon^{-1})}}. \\ \sigma_l^{\min^c} = \sqrt{\dfrac{2\triangle U_i^{lc}\left(w_{l,i}^*(\kappa_{l,c}, \tau_{l,c}, n_{l,c})\right)}{\ln(\varepsilon^{-1})}}, \\ \sigma_l^{\max^c} = \sqrt{\dfrac{2F_c(\kappa_{l,c}, \tau_{l,c}, n_{l,c})}{\ln(\varepsilon^{-1})}}. \end{cases} \tag{12}$$

We observe that $\sigma_l^{\min^e}$ and $\sigma_l^{\min^c}$ depend on the fixed point coordinate $w_{l,i}^*(\kappa_{l,e}, \tau_{l,e}, n_{l,e})$ and $w_{l,i}^*(\kappa_{l,c}, \tau_{l,c}, n_{l,c})$ which in turn also depends on the synaptic parameters $\kappa_{l,e}$, $\tau_{l,e}$, $n_{l,e}$ and $\kappa_{l,c}$, $\tau_{l,c}$, $n_{l,c}$, respectively. Therefore, changing $(\kappa_{l,e}, \tau_{l,e}, n_{l,e})$ or $(\kappa_{l,c}, \tau_{l,c}, n_{l,c})$ will change the value of $w_{l,i}^*(\kappa_{l,e}, \tau_{l,e}, n_{l,e})$ or $w_{l,i}^*(\kappa_{l,c}, \tau_{l,c}, n_{l,c})$, which will in turn change the value of $\sigma_l^{\min^e}$ or $\sigma_l^{\min^c}$ via the energy barrier function $\triangle U_i^{le}(w_{l,i}^*)$ or $\triangle U_i^{lc}(w_{l,i}^*)$, respectively. However, because of the local nature electrical synapses and non-locality of the chemical synapses, we fixed $n_{l,e} = 1$ and $n_{l,c} = 8$ throughout our numerical computations. Hence, the two control parameters used are the synaptic time-delayed couplings $(\tau_{l,e}, \kappa_{l,e})$ and $(\tau_{l,c}, \kappa_{l,c})$. On the other boundary, $\sigma_l^{\max^e}$ and $\sigma_l^{\max^c}$ do not depend on the coordinates of the stable homogeneous fixed point but on the complicated functions $F_e(\kappa_{l,e}, \tau_{l,e}, n_{l,e})$ and $F_c(\kappa_{l,c}, \tau_{l,c}, n_{l,c})$, completely defined in Equations (9), (10), and (11). Knowing the minimum and maximum range of the noise amplitude within which SISR occurs will be very useful in discussing the numerical results in the following sections.

5. SISR IN ISOLATED LAYERS

5.1. SISR in Isolated Layers With Electrical Synapses Only

We begin our numerical study with the dynamics of layer l in isolation, where neurons are connected only via local electrical synapses in a ring network topology, i.e., we consider Equation (1) with $n_{l,e} = 1$, $\kappa_{l,e} \neq 0$ and $\kappa_{m,e} = \kappa_{m,c} = \kappa_{l,c} = 0$. **Figure 4** shows the variation of R_T against the noise amplitude σ_l for this layer. In the numerical computations, we choose $\varepsilon = 0.0005 \ll 1$ because SISR can only occur in the singular limit, $\varepsilon \to 0$, and the weak noise limit, $\sigma_l \to 0$, imposed by Equation (8).

In **Figure 4A**, a weak electrical synaptic strength is considered fixed, $\kappa_{l,e} = 0.1$. All the flat-bottom R_T-curves obtained with different time delays ($\tau_{l,e} = 0.0, \tau_{l,e} = 5.0, \tau_{l,e} = 10.0, \tau_{l,e} = 15.0, \tau_{l,e} = 20.0$) show a deep and broad minimum, indicating that the spike train has a high degree of coherence due to SISR for a wide range of the noise amplitude. We notice that even though the minimum (and low) values of R_T stays constant for various time delays, the left branch of the R_T-curve is significantly being shifted to the right as the time delay increases. This means that with weak electrical synapses, the coherence of the spiking

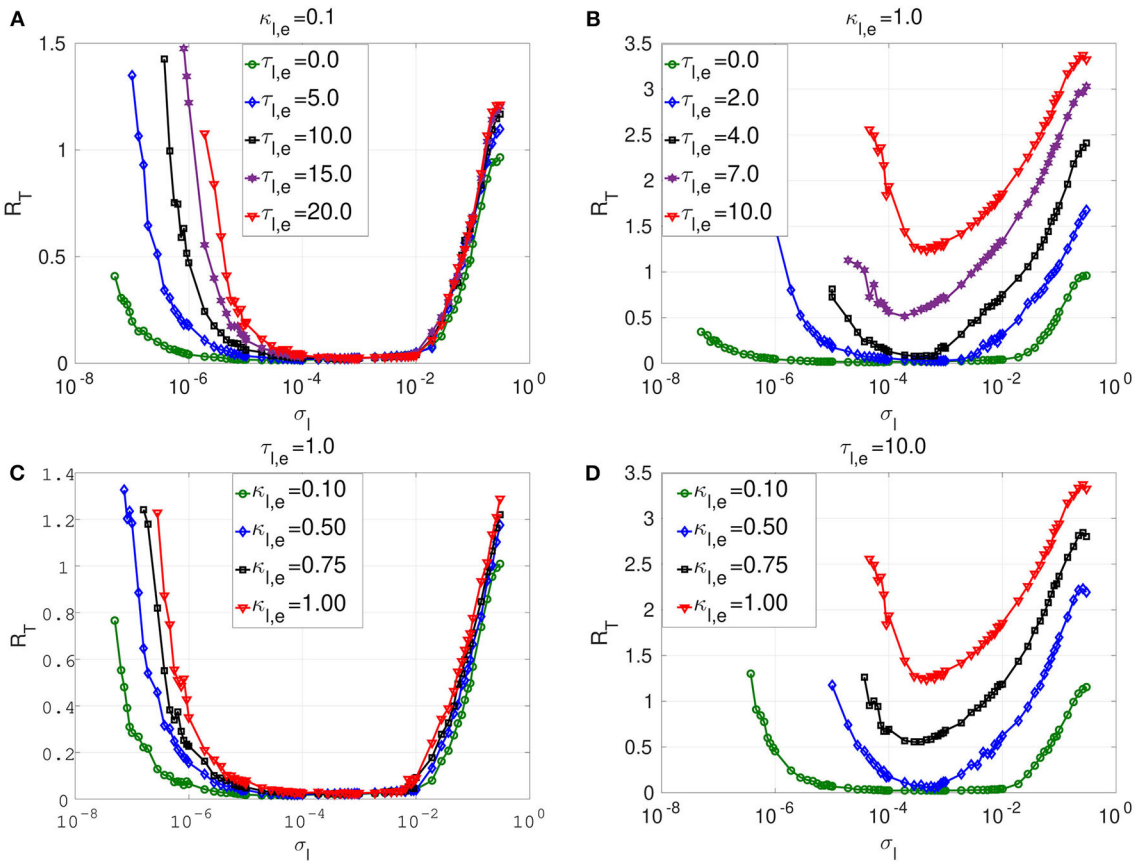

FIGURE 4 | Coefficient of variation R_T against noise amplitude σ_l of layer l in isolation. In **(A)** and **(B)**, we have the R_T curves of weak and strong electrical synaptic strengths $\kappa_{l,e}$, respectively, for short, intermediate and long synaptic time delays $\tau_{l,e}$. In **(C)** and **(D)**, we have the R_T curves of short and relatively long synaptic time delays $\tau_{l,e}$, respectively, for weak, intermediate and strong synaptic strengths $\kappa_{l,e}$. Increasing (decreasing) the electrical synaptic strength $\kappa_{l,e}$ or the length of its time delay $\tau_{l,e}$, deteriorates (enhances) SISR by increasing (decreasing) the values of R_T and by shrinking (extending) the interval of the noise amplitude in which R_T can achieve very low values. For example, in **(D)**, for $\kappa_{l,e} = 1.0$ and $\tau_{l,e} = 10.0$, the red R_T-curve lies entirely above the line $R_T = 1.0$ with a lowest value of $R_{T_{\min}} = 1.24$ occurring at just one point $\sigma_l = 4.6 \times 10^{-4}$, indicating the non-existence of SISR. Parameters of layer l: $N = 25$, $n_{l,e} = 1$, $\beta = 0.75$, $\varepsilon = 0.0005$, $\alpha = 0.5$.

activity due to SISR is not affected as the time delay becomes longer, but the coherence is achieved only at relatively larger noise amplitudes σ_l. Thus, we can obtain the same degree of SISR with longer time delays provided we increase the noise amplitude (within the interval given in Equation 12) as the time delay increases. In **Figure 4A**, we have approximately the same minimum value of $R_{T_{\min}} \approx 0.015$ for: $\tau_{l,e} = 0.0$ with $\sigma_l \in (3.7 \times 10^{-7}, 1.9 \times 10^{-2})$; $\tau_{l,e} = 5.0$ with $\sigma_l \in (2.8 \times 10^{-6}, 1.9 \times 10^{-2})$; $\tau_{l,e} = 10.0$ with $\sigma_l \in (5.5 \times 10^{-6}, 1.0 \times 10^{-2})$; $\tau_{l,e} = 15.0$ with $\sigma_l \in (1.9 \times 10^{-5}, 1.0 \times 10^{-2})$; and $\tau_{l,e} = 20.0$ with $\sigma_l \in (2.8 \times 10^{-5}, 1.0 \times 10^{-2})$. We note that the lower bound of the noise intervals increases as the time delay increases while the upper bounds are almost fixed.

In **Figure 4B**, we consider a strong electrical synapse ($\kappa_{l,e} = 1.0$). We observe that, in contrast to **Figure 4A** with a weak electrical synapse, increasing the time delay squeezes the left and right branches of the R_T-curves into a smaller noise interval, while shifting the curves to higher values, thus deteriorating SISR. In **Figure 4B**, we have different noise intervals for different minima of R_T: $R_{T_{\min}} = 0.015$ at $\tau_{l,e} = 0.0$ for $\sigma_l \in (2.8 \times 10^{-7}, 2.9 \times 10^{-2})$; $R_{T_{\min}} = 0.029$ at $\tau_{l,e} = 2.0$ for $\sigma_l \in (2.8 \times 10^{-5}, 2.8 \times 10^{-3})$; $R_{T_{\min}} = 0.078$ at $\tau_{l,e} = 4.0$ for $\sigma_l \in (1.9 \times 10^{-4}, 6.4 \times 10^{-4})$. In the last two cases, the noise intervals, in which we have the most deteriorated SISR, have shrunk to points with $R_{T_{\min}} = 0.51$ at $\sigma_l = 1.9 \times 10^{-4}$ for $\tau_{l,e} = 7.0$; and $R_{T_{\min}} = 1.24$ at $\sigma_l = 4.6 \times 10^{-4}$ for $\tau_{l,e} = 10.0$. We thus see that, with strong electrical synapses, the effect of the time delay on SISR becomes significant, unlike when the electrical synapse is weak as in **Figure 4A**. In **Figure 4B**, we observe that even though the R_T-curves for $\tau_{l,e} = 7.0$ and $\tau_{l,e} = 10.0$ are non-monotonic (characteristic of the existence of an optimal noise value for coherence), the minimum values of these curves are high (0.51 and 1.24, respectively). Here, at only $\tau_{l,e} = 10.0$, $R_{T_{\min}}$ is already above 1.0 (indicating a stochastic spiking activity that is more variable than the Poisson process), whereas with weak electrical synapses in **Figure 4A**, even at $\tau_{l,e} = 20.0$, we still have $R_{T_{\min}} \approx 0.015$.

In **Figures 4C,D**, we vary the electrical synaptic strength while the synaptic time is fixed at a short ($\tau_{l,e} = 1.0$) and a long ($\tau_{l,e} = 10.0$) delay, respectively. A similar behavior as in **Figures 4A,B** is observed, with weak and strong electrical synaptic strengths, respectively. That is, at short synaptic time

delays (see **Figure 4C**), the R_T-curves show a deep and broad minimum, indicating a high degree of coherence due to SISR for a wide range of the noise amplitude when the electrical synaptic strength $\kappa_{l,e}$ is varied. Here, as $\kappa_{l,e}$ increases, and only the left branches of the R_T-curves are shifted to the right, while the right branch of the R_T-curves are fixed, thereby fixing the upper bound of the noise amplitude σ_l below which SISR is optimal. This means that at short electrical time delays, the coherence of the spiking activity due to SISR is not affected as the electrical synaptic strength becomes stronger, but the coherence is achieved only at relatively larger noise amplitudes σ_l. In **Figure 4D**, where electrical synaptic time delays are longer, increasing the electrical synaptic strength not only increases the minimum value of the R_T-curves (thereby deteriorating SISR), but also shrinks the size of the noise interval in which SISR is optimized on both ends.

The response of SISR to changes in the synaptic strength $\kappa_{l,e}$ and time delay $\tau_{l,e}$ in **Figure 4** can be explained in terms of the electrical interaction potential $U_i^e(v_{l,i}, w_{l,i})$ given in Equation (7) and represented in **Figure 2**. We observe in **Figure 2** that, for a fixed ($n_{l,e} = 1$) ring network topology and time delay $\tau_{l,e}$, as the synaptic strength $\kappa_{l,e}$ increases, the energy barriers $\triangle U_i^{le}(w_{l,i})$ and $\triangle U_i^{re}(w_{l,i})$ become deeper. In particular, when $w_{l,i} < 0$, the trajectory is in the left potential well and as $\kappa_{l,e}$ becomes stronger $(0.25, 0.5, 0.75, 1.0)$, the left energy barrier $\triangle U_i^{le}(w_{l,i})$ becomes deeper (hence the trajectory at the bottom of the well get closer to the homogeneous stable fixed point at $w_{l,i}^* = -0.666651$). The deeper the left energy barrier $\triangle U_i^{le}(w_{l,i})$ is (in other words, the stronger the electrical synaptic strength $\kappa_{l,e}$ is), therefore, the closer the trajectory to the stable fixed point is and the further away the neural system from the oscillatory regime is. For the trajectory to jump over a high energy barrier $\triangle U_i^{le}(w_{l,i})$, a stronger noise amplitude σ_l is of course needed. This is why in **Figure 4** as $\kappa_{l,e}$ increases, the left branch of the R_T-curve is shifted to the right, meaning that stronger noise amplitudes are required to induce frequent spiking (i.e., frequent escaping from the deep left energy barrier). But, as the noise amplitude becomes bigger, the condition in Equation (8) requiring $\sigma_l \to 0$ for the occurrence of SISR is violated. Hence, SISR disappears with increasing synaptic strength.

We can also see from **Figure 4D** that at longer time delay $\tau_{l,e}$, this effect (the shifting of the left branch of the R_T-curve to the right) is more pronounced than in **Figure 4C** with a shorter time delay. This is because, in Equation (7), the longer the time delay is ($\tau_{l,e} \gg 0$), the further away is the quantity $\left[v_{l,i}(t)v_{l,j}(t-\tau_{l,e}) - v_{l,i}(t)^2\right]$ from zero (since neurons are identical); hence, the stronger is the effect of the synaptic strength $\kappa_{l,e}$ on the electrical interaction potential, the energy barrier functions, and, consequently, on the R_T-curves. Otherwise, if $\tau_{l,e} \to 0$, then because the neurons are identical, $\left[v_{l,i}(t)v_{l,j}(t-\tau_{l,e}) - v_{l,i}(t)^2\right] \to 0$, and $\kappa_{l,e}$ will have little effect on the electrical interaction potential, the energy barriers functions, and consequently on the R_T-curves. This is why the synaptic strength $\kappa_{l,e}$ has a stronger effect on SISR only when $\tau_{l,e}$ gets longer, and vice versa. This theoretical explanation will also support the behavior of the time-delayed chemical synapses in the optimization of SISR as we shall see further below.

Secondly, at weak electrical synaptic strengths and short time delays (**Figures 4A,C**), the upper bound of the noise interval for which the R_T-curves achieve their minima is almost constant. Here, only the lower bound of the noise intervals is shifted to the right. Whereas, at strong electrical synaptic strengths and long time delays (**Figures 4B,D**), both the lower and upper bounds of the noise intervals are shifted to the right and to t4he left as $\tau_{l,e}$ and $\kappa_{l,e}$ increase, respectively. This has the overall effect of shrinking the noise interval in which the R_T-curves achieve their minima to a single value of σ_l. This behavior can be explained in terms of the minimum and maximum noise amplitudes between which SISR occurs obtained in Equation (12).

We observe from Equation (12) that $\sigma_l^{\min^e}$ depends on the fixed point coordinate $w_{l,i}^*(\kappa_{l,e}, \tau_{l,e}, n_{l,e})$, which, in turn, also depends on $\kappa_{l,e}, \tau_{l,e}$, and $n_{l,e} = 1$. Therefore, changing $\kappa_{l,e}$ and $\tau_{l,e}$ will change the value of $w_{l,i}^*(\kappa_{l,e}, \tau_{l,e}, n_{l,e})$, which will, in turn, change the value of $\sigma_l^{\min^e}$ via the energy barrier function $\triangle U_i^{le}(w_{l,i}^*)$. Numerical computations indicate that $\sigma_l^{\min^e}$ increases as $\kappa_{l,e}$ and $\tau_{l,e}$ increase (see **Figure 4**). On the other boundary, $\sigma_l^{\max^e}$ does not depend on the coordinates of the homogeneous stable fixed point, but on the complicated function $F_e(\kappa_{l,e}, \tau_{l,e}, n_{2,e})$, fully determined by Equations (9) and (10). In **Figures 4A,C** (i.e., in the regimes of weak electrical synaptic strength and short time delays, respectively), we notice that $\sigma_l^{\max^e} \approx 10^{-2}$ is nearly constant for all values of the time delay and electrical synaptic strength used. In Yamakou and Jost (2018), where a single isolate FHN neuron is considered, such fixation of the upper bound of the noise interval in which SISR occurs was already observed. In the case of a single isolated FHN neuron, the function F_e in Equation (9) takes a simple constant value $F_e = \frac{3}{4}$. This implies (for a fixed $\varepsilon = 0.0005$) a fixed value for $\sigma_l^{\max^e} = \left[3/2 \cdot \log_e(\varepsilon^{-1})\right]^{1/2}$.

In the case where a network of coupled FHN neurons is considered, the fixation of the upper bound of the noise interval for which SISR occurs can only be observed if $F_e(\kappa_{l,e}, \tau_{l,e}, n_{2,e}) \to C$, where C is a constant. In particular, in a weak electrical synaptic regime ($\kappa_{l,e} \to 0$) and short time delay ($\tau_{l,e} \to 0$) regime (or more precisely, $[v_{l,i}v_{l,j}(t-\tau_{l,e}) - v_{l,i}^2(t)] \to 0$ as $\tau_{l,e} \to 0$, because all the neurons are identical), $F(\kappa_{l,e}, \tau_{l,e}, n_{l,e}) \to \frac{3}{4}$. In these regimes (see **Figures 4A,C**), we observe that $\sigma_l^{\max^e} \approx 10^{-2}$, corresponding to the value obtained in Yamakou and Jost (2018) for the case of a single isolated FHN neuron ($\kappa_{l,e} = 0$). In the regimes of strong coupling ($\kappa_{l,e} \gg 0$) and of long time delays ($\tau_{l,e} \gg 0 \Rightarrow [v_{l,i}v_{l,j}(t-\tau_{l,e}) - v_{l,i}^2(t)] \neq 0$) shown in **Figures 4B,D**, the function F_e in Equation (9) is now strongly modified by the large values of $\kappa_{l,e}$ and $\tau_{l,e}$. This is why in these regimes, the upper bound $\sigma_l^{\max^e}$ of the noise interval, for which SISR occurs, is not fixed any longer but shifted to the left as $\tau_{l,e}$ and $\kappa_{l,e}$ take on larger values. In the case of chemical synapses, as we shall see later, the same theoretical explanation holds for the shrinking, on both ends, of the interval of the noise amplitude in which SISR is optimized. Later, we shall focus on layer $l = 2$ with a non-existent SISR when it is in isolation ($\kappa_{2,e} = 1.0$ and $\tau_{2,e} = 10.0$; see the red

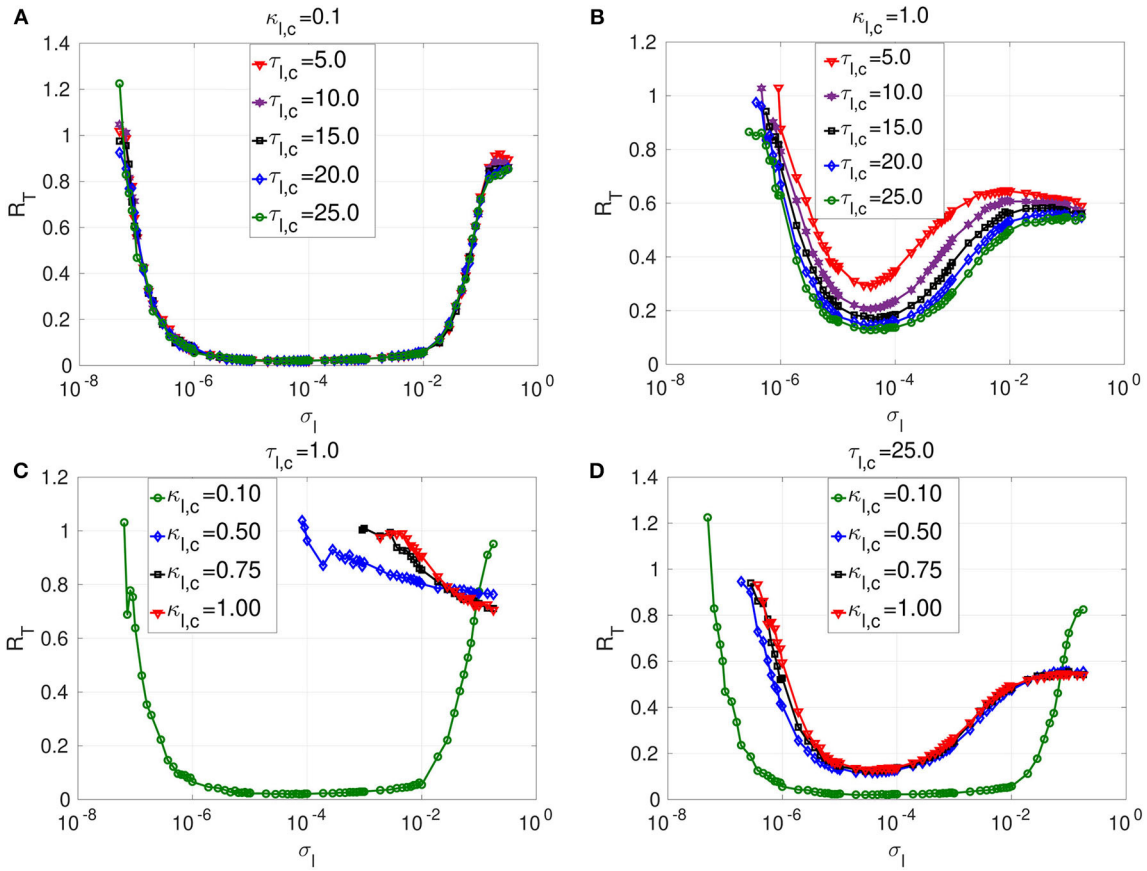

FIGURE 5 | Coefficient of variation R_T vs noise amplitude σ_l of layer l in isolation. In **(A)** and **(B)**, we have the R_T curves of weak and strong chemical synaptic strengths $\kappa_{l,c}$, respectively, for short, intermediate and long synaptic time delays $\tau_{l,c}$. In **(C)** and **(D)**, we have the R_T curves of short and relatively long synaptic time delays $\tau_{l,c}$, respectively, for weak, intermediate and strong synaptic strengths $\kappa_{l,c}$. Increasing (decreasing) the inhibitory chemical synaptic strength $\kappa_{l,c}$ deteriorates (enhances) SISR by increasing (decreasing) the values of R_T and by shrinking (extending) the interval of the noise amplitude in which R_T can achieve very low values. Thus, inhibitory chemical synaptic strength qualitatively behaves as the electrical synaptic strength in optimizing SISR. However, electrical synaptic and inhibitory chemical synaptic time delays show opposite behaviors in the enhancement of SISR. Decreasing (increasing) the length of inhibitory chemical time delays $\tau_{l,c}$, deteriorates (enhances) SISR by increasing (decreasing) the values of R_T and by shrinking (extending) the interval of the noise amplitude in which R_T can achieve very low values. This effect is particularly pronounced when the chemical synaptic strength is strong. For example, in **(C)**, for $\kappa_{l,c} = 1.0$ and $\tau_{l,e} = 1.0$, the red R_T-curve achieves relatively high minimum value $R_{T_{\min}} = 0.71$ occurring at a relatively large noise amplitude $\sigma_l = 4.6 \times 10^{-4}$, indicating a very poor SISR. Parameters for layer l are: $N = 25$, $n_{l,c} = 8$, $\beta = 0.75$, $\varepsilon = 0.0005$, $\alpha = 0.5$.

curve in **Figure 4D** with $R_{T_{\min}} > 1$) and then investigate which multiplexing configuration can best optimize SISR in this layer when it is multiplexed with layer $l = 1$ when it already exhibits pronounced SISR.

5.2. SISR in Isolated Layers With Inhibitory Chemical Synapses Only

We investigated the dynamics of layer l in isolation, where neurons are connected only via (non-local) inhibitory chemical synapses in a ring network topology. Specifically, we consider Equation (1) with $n_{l,e} = 8$, $\kappa_{l,c} \neq 0$ and $\kappa_{m,e} = \kappa_{m,c} = \kappa_{l,e} = 0$. **Figure 5** shows the variation of R_T against the noise amplitude σ_l for this layer. We also fixed $\varepsilon = 0.0005 \ll 1$ so that Equation (8) can be satisfied in a weak noise limit $\sigma_l \to 0$, leading to the occurrence of SISR. We shall now mainly compare the enhancement of SISR in layer l for two situations, i.e., when the neurons are locally connected via time-delayed electrical synapses (see **Figure 4**) and when the neurons are non-locally connected via time-delayed inhibitory chemical synapses (see **Figure 5**).

The first observation is that longer inhibitory time delays enhance SISR, while longer electrical time delays deteriorate SISR. However, similarly to electric time delays, chemical time delays ($\tau_{l,c}$) have a strong effect on SISR only for stronger chemical synaptic strength ($\kappa_{l,c}$). In **Figure 4A**, the electrical synaptic strength is weak ($\kappa_{l,e} = 0.1$). Even though the interval of the noise amplitude, for which a pronounced SISR occurs (as indicated by the very low values of R_T), shrinks on the left bound with increasing time delay, the low values of R_T within that interval remain unchanged (≈ 0.015). Similarly, results in **Figure 5A** with the same weak inhibitory synaptic strength ($\kappa_{l,c} = 0.1$) show that changing the chemical time delays does also

not affect the low and constant values of $R_T \approx 0.014$ (indicating an optimized SISR). In contrast, however, the lower bound of the noise interval with optimal SISR remains independent of varying levels of time delay. Thus, for weak synaptic strength and for increasing synaptic time delays, inhibitory chemical synapses outperform electrical synapses in optimizing SISR, in the sense that the former allow for a wider range of noise amplitudes for which R_T remains low.

In **Figure 5B**, where a large inhibitory chemical synaptic strength ($\kappa_{l,c} = 1.0$) is considered, time delays can have significant effect on SISR, and this is in contrast to **Figure 5A**, where $\kappa_{l,c}$ is weak. In **Figure 5B**, increasing the chemical time delay enhances SISR by lowering the minimum value of R_T. In comparison to **Figure 5A**, the noise interval for which SISR is optimal has shrunk on both sides. The reason for this shrinking on both ends of the optimal noise interval is essentially the same as for the case where the strength of electrical synapses is varied (see **Figures 4B,D**). In **Figure 5B**, we have $R_{T_{\min}} = 0.29$ at $\sigma_l = 3.7 \times 10^{-5}$ for $\tau_{l,c} = 5.0$; $R_{T_{\min}} = 0.21$ at $\sigma_l = 3.7 \times 10^{-5}$ for $\tau_{l,c} = 10.0$; $R_{T_{\min}} = 0.17$ at $\sigma_l = 4.6 \times 10^{-5}$ for $\tau_{l,c} = 15.0$; $R_{T_{\min}} = 0.15$ at $\sigma_l = 4.6 \times 10^{-5}$ for $\tau_{l,c} = 20.0$; $R_{T_{\min}} = 0.12$ at $\sigma_l = 3.7 \times 10^{-5}$ for $\tau_{l,c} = 25.0$. However, the deteriorating effects of electrical time delays on SISR is more pronounced than those of the chemical time delays at the same synaptic strength ($\kappa_{l,e} = \kappa_{l,c} = 1.0$) (see **Figures 4B**, **5B**. This confirms that chemical synapses are better at optimizing SISR than electrical synapses, not only because they allow for a wider range of noise amplitude in which optimal SISR may occur but also for the occurrence of a more enhanced SISR, as indicated by the relatively lower values of R_T at long time delays.

In **Figures 5C,D**, we investigate the effects of chemical synaptic strength in a short and long time delay regime. Irrespective of the time delay regime, the stronger the chemical synaptic strength is, the more deteriorated SISR is. The reason behind this behavior is the same as the one given for the case of electrical synapses. That is, as the chemical synaptic strength $\kappa_{l,c}$ becomes larger, the energy barriers $\triangle U_i^{lc}(w_{l,i})$ and $\triangle U_i^{rc}(w_{l,i})$ become deeper (see **Figure 3**). However, now a stronger noise amplitude is required to jump over the deep energy barriers and induce spiking, and this strong noise amplitude destroys the coherence of the spiking (by violating the conditions in Equation (8) requiring $\sigma_l \to 0$) and hence deteriorates SISR.

The deterioration of SISR by stronger chemical synaptic strengths is also observed with stronger electrical synaptic strength. However, for short synaptic time delay regimes ($\tau_{l,e} = 1.0 = \tau_{l,c}$, see **Figures 4C**, **5C**), we notice the following difference for both synaptic types: When the time delay is relatively short, electrical synapses optimize SISR compared to chemical synapses as the synaptic strength is weakened. We see in **Figure 5C** with $\tau_{l,c} = 1.0$ that SISR is destroyed as the $\kappa_{l,c}$ increases, whereas in **Figure 4C** with $\tau_{l,e} = 1.0$, SISR remains enhanced as $\kappa_{l,e}$ increases. This means that an electrical synapse is a better means than a chemical synapse in optimizing SISR at very short time delays, irrespective of the synaptic strengths, while a chemical synapse is better than an electrical synapse at very long time delays, irrespective of the synaptic strengths.

In **Figure 5**, the reason for the deterioration of SISR with decreasing time delays could be inferred from the reason given for the deterioration of SISR with increasing synaptic strength. That is, shortening the chemical synaptic time delays increases the depth of the chemical energy barrier functions given in Equation (9). This will in turn demand larger noise amplitude to jump over deep energies barriers to induce spiking with no coherence and hence very poor SISR, as seen, for example, from the red curve in **Figure 5C**. Here, we see that rare spiking can be induced only when the noise $\sigma_l \geq 10^{-4}$ as R_T stays high with increasing noise amplitude σ_l. However, from conditions in Equation (8), SISR requires $\sigma_l \to 0$, which implies that increasing the noise would not improve SISR. We can see from the red curve in **Figure 5C** that a minimum value of $R_{T_{\min}} \approx 0.71$, already indicating a very poor SISR, occurs at a relatively large noise amplitude of $\sigma_l = 0.18$. Below, we shall focus on the enhancement of this very poor SISR in layer $l = 2$ (for $\kappa_{2,c} = \tau_{2,c} = 1.0$; see also red curve in **Figure 5C**) by using various multiplexing configurations, where layer $l = 1$ already exhibits enhanced SISR in isolation.

6. MULTIPLEXING AND OPTIMIZATION OF SISR

We now address the following questions: (1) Is an optimization of SISR based on the multiplexing of layers possible? (2) Which synaptic multiplexing configuration is the best optimizer of SISR? To answer these two questions, we configure the synaptic strength and time delay of one layer (say layer 1) such that SISR is optimal in this layer. The corresponding parameters of layer 2 are configured such that SISR is non-existent in this layer. Then, we connect these two layers in a multiplex fashion (see **Figure 1**) in six different multiplexing configurations.

In the first three configurations, the two layers of the multiplex network each consist of neurons that are intra-connected by only electrical synapses ($\kappa_{1,e}, \tau_{1,e}$) and ($\kappa_{2,e}, \tau_{2,e}$), and are inter-connected (multiplexed) by (i) electrical synapses ($\kappa_{m,e}, \tau_{m,e}$), (ii) inhibitory chemical synapses ($\kappa_{m,c}, \tau_{m,c}$), and (iii) excitatory chemical synapses ($\kappa_{m,c}, \tau_{m,c}$).

In the next three configurations, we use the same three synaptic multiplexing configurations of two layers, each consisting of neurons that are intra-connected by only inhibitory chemical synapses. We do not consider excitatory chemical synapses for intra-connectivity because this type of synapse induces coherent spiking activities even in the absence of noise, which is not a requirement for SISR. On the other hand, we use excitatory chemical synapses in the inter-layer connections (multiplexing) Because, for some synaptic strengths and time delays, the multiplex network remains in the excitable regime in the absence of noise—a requirement for observing SISR. However, we also have situations in which the multiplexing excitatory chemical synapses strengthen the excitable regime of the network by making the homogeneous fixed point more stable, thereby requiring very large noise amplitudes to have a chance of inducing a spike. Large noise amplitudes, however, violate the conditions necessary for the occurrence of SISR.

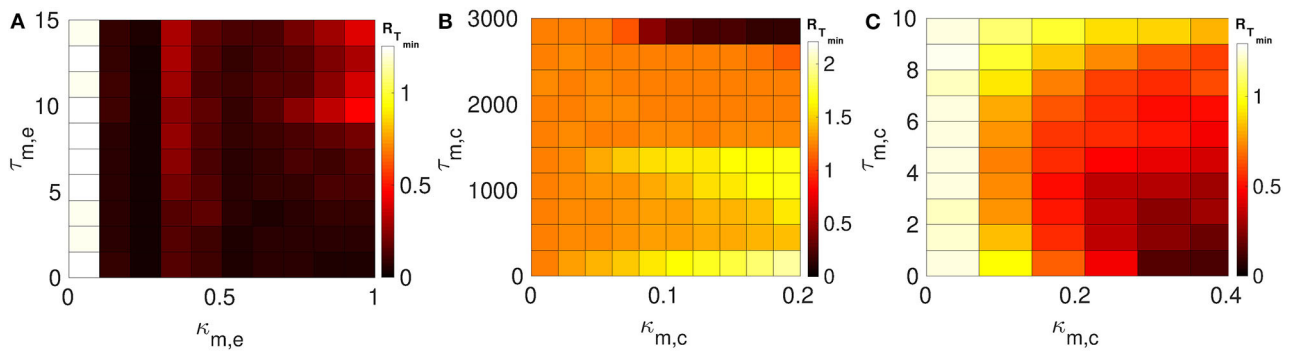

FIGURE 6 | Color-coded minimum coefficient of variation ($R_{T_{min}}$) of layer 2 as a function of the multiplexing parameters. Both layers 1 and 2 are intra-connected by electrical synapses. **(A)** Shows the enhancement performances of the electrical multiplexing ($\kappa_{m,e}, \tau_{m,e}$). For optimized SISR in layer 2, we need either a shorter $\tau_{m,e}$ and stronger $\kappa_{m,e}$; or a longer $\tau_{m,e}$ and a weaker $\kappa_{m,e}$. **(B)** Shows the enhancement performances of inhibitory chemical multiplexing ($\kappa_{m,c}, \tau_{m,c}$). For optimized SISR in layer 2, we need a stronger $\kappa_{m,c}$ and a very long $\tau_{m,c}$. **(C)** Shows the enhancement performances of excitatory chemical multiplexing ($\kappa_{m,c}, \tau_{m,c}$). For optimized SISR in layer 2 we need very short $\tau_{m,c}$ and stronger $\kappa_{m,c}$. Parameters for layer 1, we need $N = 25$, $n_{1,e} = 1$, $\beta = 0.75$, $\varepsilon = 0.0005$, $\alpha = 0.5$, $\kappa_{1,e} = 0.1$, and $\tau_{1,e} = 1.0$. Parameters for layer 2, we need $N = 25$, $n_{2,e} = 1$, $\beta = 0.75$, $\varepsilon = 0.0005$, $\alpha = 0.5$, $\kappa_{2,e} = 1.0$, and $\tau_{2,e} = 10.0$.

We shall therefore also avoid such regimes and stay in the excitable regimes where vanishingly small noise amplitudes have a non-zero probability of inducing at least a spike in the large time interval we considered in our simulations.

In the following numerical simulations, we have ensured that the multiplex network stays in the excitable regime by checking that all the synaptic strengths and time delays of the multiplexing synapses are such that no self-sustained spiking activity occurs in the absence of noise. It is worth mentioning here that the optimization of SISR based on the multiplexing approach appears not to be feasible in a network of mixed layer type, i.e., consisting of an electrical layer multiplexed to a chemical layer. We investigated all the possible configurations of mixed layered networks, i.e., those with electrical, inhibitory, and excitatory chemical multiplexing. Extensive numerical simulations (not shown) clearly indicated that the optimization of SISR, for the ranges of the multiplexing synaptic strengths and time delays considered, was not possible in mixed layered networks. For this reason, we only discuss the layered networks that display the capability of optimizing SISR and compare the optimization abilities of various multiplexing connections between two electrical layers and then between two inhibitory chemical layers.

6.1. Multiplexing of Electrical Layers

We consider layer 1 and layer 2 in which neurons are electrically coupled only. We choose the synaptic parameters $(\kappa_{1,e}, \tau_{1,e})$ of layer 1 such that SISR is pronounced, i.e., we choose a weak synaptic strength $\kappa_{1,e} = 0.1$ and a short synaptic time delay $\tau_{1,e} = 1.0$ (green R_T-curve **Figure 4C** with $R_{T_{\min}} = 0.015$). For layer 2, we set the synaptic parameters such that SISR is non-existent, i.e., we choose a strong synaptic strength $\kappa_{1,e} = 1.0$ and a long synaptic time delay $\tau_{2,e} = 10.0$ (red R_T-curve **Figure 4D** with $R_{T_{\min}} = 1.24$). These two layers are then coupled in a multiplex network as shown in **Figure 1A**. The multiplexing introduces two other parameters—the multiplexing synaptic strengths $\{\kappa_{m,e}, \kappa_{m,c}\}$ and their corresponding time delays $\{\tau_{m,e}, \tau_{m,c}\}$. **Figure 6** shows the color-coded minimum values of the coefficient of variation $R_{T_{\min}}$ of layer 2 as a function of the multiplexing parameters for the three multiplexing configurations considered.

In **Figure 6A**, the multiplexing between the two layers is mediated by electrical synapses with parameters $(\kappa_{m,e}, \tau_{m,e})$. We can clearly see that even very weak multiplexing $\kappa_{m,e} \geq 0.1$, particularly at short time delays $\tau_{m,e} \leq 9.5$, can induce a very pronounced SISR in layer 2 (where SISR was non-existent in isolation) as indicated by the dark red color corresponding to very low values of $R_{T_{\min}}$. In the region $\tau_{m,e} \leq 9.5$, stronger multiplexing strengths push $R_{T_{\min}}$ to even lower values as indicated by the darker red color, thus optimizing SISR in layer 2. However, as the multiplexing time delay becomes longer $\tau_{m,e} > 9.5$, this time delay starts to dominate the control of SISR. As the time delay $\tau_{m,e} > 9.5$ increases, SISR progressively deteriorates and the effect of strong multiplexing is reversed, i.e., the stronger $\kappa_{m,e}$ is, the more SISR deteriorates, as indicated by the change of color of $R_{T_{\min}}$ from dark red to light red. While in this same region, i.e., $\tau_{m,e} > 9.5$, weaker multiplexing optimize SISR better than strong ones, as seen in the region bounded by $\tau_{m,e} \in [9.5, 15.0]$ and $\kappa_{m,e} \in [0.1, 1.0]$ with a dark red color.

In **Figure 6B**, the multiplexing between the two electrical layers is mediated by inhibitory chemical synapses with parameters $(\kappa_{m,c}, \tau_{m,c})$. We notice, in contrast to **Figure 6A**, that the multiplexing inhibitory chemical synaptic strength takes a maximum value of $\kappa_{m,c} = 0.2$, i.e., it stays in the weak multiplexing regime and the time delay goes up to the very large value of $\tau_{m,c} = 3,000$. As already pointed out, we always want the network to stay in the excitable regime such that self-sustained oscillations do not arise due to bifurcations. In this multiplexing configuration, values of the inhibitory chemical synaptic strength greater than 0.2 induces oscillations in the absence of noise—for SISR, the system should oscillate coherently due only to the presence of noise and not due to a bifurcation. As we can see in **Figure 6B**, weak multiplexing inhibitory chemical synaptic strength $\kappa_{m,c} \in [0.0, 0.2]$ cannot induce SISR in layer 2 as the

values of $R_{T_{\min}}$ stay very high above 1.0, except at very long multiplexing delays $\tau_{m,c} \geq 2,750$. It can be observed that for time delays $\tau_{m,c} \leq 1,500$, stronger multiplexing values, $\kappa_{2,c} \gtrsim 0.1$, deteriorate SISR (yellow region) to a larger extent than the weaker values, $\kappa_{2,e} \lesssim 0.1$ (orange region). But when the multiplexing time delay becomes very long, e.g., at $\tau_{m,c} = 3,000$, stronger multiplexing ($\kappa_{2,c} \gtrsim 0.1$) induces an optimized (dark red color of $R_{T_{\min}}$) SISR in layer 2, while weaker multiplexing ($\kappa_{2,c} \lesssim 0.1$) cannot optimize SISR, as indicated by the orange color of $R_{T_{\min}}$. This means that multiplexing with inhibitory chemical synapses has the opposite effect compared to multiplexing with electrical synapses, in terms of SISR in layer 2. To sum up, stronger $\kappa_{m,c}$ means poorer SISR at shorter $\tau_{m,c}$ but better SISR at longer $\tau_{m,c}$; the opposite is also true, as stronger $\kappa_{m,e}$ means better SISR at shorter $\tau_{m,e}$, but poorer SISR at longer $\tau_{m,e}$.

In **Figure 6C**, the multiplexing between the two electrical layers is mediated by excitatory chemical synapses with parameters $(\kappa_{m,c}, \tau_{m,c})$. First, we notice the range of the synaptic strength and the time delay. For $\kappa_{m,c} > 0.4$, the excitability of the network becomes so strong that even large noise amplitudes (SISR requires vanishingly small noise) are not longer capable of inducing a spike in the large time interval considered. In contrast to weak multiplexing electrical synapses in **Figure 6A**, weak multiplexing excitatory chemical synapses are incapable of inducing SISR in layer 2. In **Figure 6C**, for a weak multiplexing $\kappa_{m,c} \in [0.0, 0.28]$, $R_{T_{\min}}$ remains high with the lowest value above 0.5 (as indicated by the white, yellow, orange, and light red colors of $R_{T_{\min}}$) for all of the time delays considered. This inability of weak excitatory chemical multiplexing to optimize SISR in layer 2 is similar to that of weak inhibitory chemical multiplexing in **Figure 6B**. However, for an intermediate excitatory chemical multiplexing, i.e., $\kappa_{m,c} \in]0.28, 0.4]$, an optimized SISR is induced in layer 2 (just like with intermediate electrical multiplexing in **Figure 6A**), but only at very short time delays $\tau_{m,c} \in [0.0, 2.0]$, where $R_{T_{\min}}$ assumes low values corresponding to the dark red colors of $R_{T_{\min}}$. And the main difference between inhibitory chemical multiplexing and excitatory chemical multiplexing is in terms of their time delays. While inhibitory chemical multiplexing requires extremely long time delay ($\tau_{m,c} \geq 2,750$) to optimize SISR in layer 2, excitatory chemical multiplexing requires extremely short time delays ($\tau_{m,c} \in [0.0, 2.0]$) for the optimization.

6.2. Multiplexing of Inhibitory Chemical Layers

Here we consider layers 1 and 2 in which neurons are coupled only via inhibitory chemical synapses. We set the synaptic parameters $(\kappa_{1,c}, \tau_{1,c})$ of layer 1 such that SISR is pronounced, i.e., we choose a weak synaptic strength $\kappa_{1,c} = 0.1$ and a long synaptic time delay $\tau_{1,c} = 25.0$, see the green R_T-curve **Figure 5D** with $R_{T_{\min}} = 0.015$. For layer 2, we set the synaptic parameters such that SISR is very poor, i.e., we choose a strong synaptic strength $\kappa_{1,c} = 1.0$ and a short synaptic time delay $\tau_{2,c} = 1.0$, see the red R_T-curve **Figure 5C** with $R_{T_{\min}} = 0.71$. These two layers are then coupled in a multiplex network as shown in **Figure 1B**. The multiplexing introduces two other parameters—the multiplexing synaptic strengths $\{\kappa_{m,e}, \kappa_{m,c}\}$ and their corresponding time delays $\{\tau_{m,e}, \tau_{m,c}\}$. **Figure 7** shows the color-coded minimum values of the coefficient of variation $R_{T_{\min}}$ of layer 2 as a function of the multiplexing parameters for the three multiplexing configurations considered.

In **Figure 7A**, the multiplexing between the two inhibitory chemical layers is mediated by electrical synapses with parameters $(\kappa_{m,e}, \tau_{m,e})$. It is observed that electrical multiplexing cannot at all optimize SISR in layer 2 as indicated by the very high values of $R_{T_{\min}}$ in the entire $\kappa_{m,e} - \tau_{m,e}$ parameter space. Comparing **Figure 6A** and **Figure 7A**, we can conclude that electrical multiplexing become good optimizers of SISR only when the multiplexed layers are both intra-connected by electrical synapses. In particular, we observe that, while a strong multiplexing $\kappa_{m,e}$ with a short delay $\tau_{m,e}$ of layers intra-connected by electrical synapses optimizes SISR in layer 2, (see **Figure 6A**), a strong multiplexing $\kappa_{m,e}$ with a short delay $\tau_{m,e}$ of layers intra-connected by inhibitory chemical synapses makes SISR rather worse ($R_{T_{\min}} \geq 1.0$, see **Figure 7A**) in layer 2 than when this layer is in isolation ($R_{T_{\min}} = 0.71$).

In **Figure 7B**, the multiplexing between the two inhibitory chemical layers is mediated by inhibitory chemical synapses with parameters $(\kappa_{m,c}, \tau_{m,c})$. In this multiplexing configuration, an optimization of SISR in layer 2 is impossible as well, especially at intermediate multiplexing strengths and short time delays, where the $R_{T_{\min}}$ assumes an even larger value ($R_{T_{\min}} \approx 0.9$) than layer 2 in isolation ($R_{T_{\min}} = 0.71$). Moreover, even very long multiplexing time delays, as in the case of electrical layers multiplexed by inhibitory chemical synapses (see **Figure 6B**), cannot optimize SISR in layer 2 irrespective of the multiplexing strength. Thus, we conclude that multiplexing inhibitory chemical synapses is generally a bad optimizer of SISR in layers intra-connected by either chemical synapses or electrical synapses. However, recall that inhibitory chemical synapses can be very good optimizers of SISR within a layer (see **Figure 5**).

In **Figure 7C**, the multiplexing between the two inhibitory chemical layers is mediated by excitatory chemical synapses with parameters $(\kappa_{m,c}, \tau_{m,c})$. Note that the range of multiplexing time delay we considered is very short, i.e., $\tau_{m,c} \in [0.0, 2.0]$. For $\tau_{m,c} > 2.0$, excitatory chemical multiplexing induces self-sustained oscillations in the absence of noise—a regime not required for SISR. In contrast to electrical and inhibitory chemical multiplexing of layers intra-connected with chemical synapses (see **Figures 7A,B**), excitatory chemical multiplexing of such layers can perform extremely well at optimizing SISR. However, this capability for very strong optimization is only possible at strong excitatory chemical multiplexing ($\kappa_{m,c} \geq 0.8$) and very short time delays ($\tau_{2,c} \in [0.0, 1.2]$) with $R_{T_{\min}} \approx 0.03$. This implies that excitatory chemical synapses, as a multiplexing synapse, could play more important functional roles (than electrical and inhibitory chemical synapses) in neural information processing due to SISR in multiplexed layers intra-connected by inhibitory chemical synapses.

7. CONCLUSION

We have investigated the effects of electrical and chemical synaptic couplings on the noise-induced phenomenon of SISR in isolated layers as well as in multiplexed layer networks of the FHN neuron model in the excitable regime. We have presented

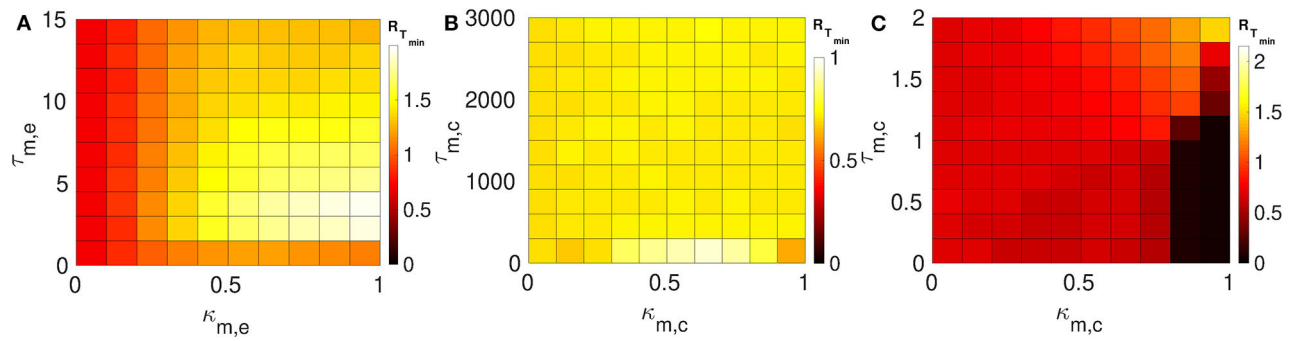

FIGURE 7 | Color-coded minimum coefficient of variation ($R_{T_{min}}$) of layer 2 against multiplexing parameters. Each of layer 1 and layer 2 is intra-connected by inhibitory chemical synapses. **(A)** Shows the enhancement performances of the electrical multiplexing ($\kappa_{m,e}$, $\tau_{m,e}$). Here, we observe that electrical multiplexing cannot optimize SISR in layer 2, especially at stronger multiplexing. **(B)** Shows the enhancement performances of the inhibitory chemical multiplexing ($\kappa_{m,c}$, $\tau_{m,c}$). In this case, the enhancement of SISR is even worse than in **(A)**, especially at intermediate multiplexing strengths and short time delays. **(C)** Shows the enhancement performances of the excitatory chemical multiplexing ($\kappa_{m,c}$, $\tau_{m,c}$). Here, an enhancement is possible. An optimized SISR ($R_{T_{min}} \approx 0.03$) emerging at strong excitatory chemical synapses ($\kappa_{m,c} \geq 0.8$) with short time delays ($\tau_{m,c} \leq 1.2$). Parameters of layer 1: $N = 25$, $n_{1,c} = 8$, $\beta = 0.75$, $\varepsilon = 0.0005$, $\alpha = 0.5$, $\kappa_{1c} = 0.1$, $\tau_{1,c} = 25.0$. Parameters of layer 2: $N = 25$, $n_{l,c} = 8$, $\beta = 0.75$, $\varepsilon = 0.0005$, $\alpha = 0.5$, $\kappa_{l,c} = 1.0$, $\tau_{l,c} = 1.0$.

the analytic conditions necessary for SISR to occur in isolated layers with neurons connected either via electrical or inhibitory chemical synapses. From these analytic conditions, we have also obtained the minimum and maximum synaptic noise amplitude required for the occurrence of SISR in isolated layers.

Numerical computations indicate that, in an isolated layer, the weaker the electrical synaptic strength and the shorter the corresponding synaptic time delay are, the more enhanced SISR is. However, the deteriorating effect of stronger electrical synaptic couplings is significant only at longer time delays and vice versa. On the other hand, in an isolated layer with inhibitory chemical synapses, weaker inhibitory chemical synaptic couplings just like their weaker electrical counterparts enhance SISR. Moreover, the longer the synaptic time delay is, the more enhanced SISR is—in contrast to isolated layers with electrical synapses. The enhancing effect of the longer synaptic time delays in isolated layers with inhibitory chemical synapses becomes significant only at stronger synaptic strengths. Furthermore, it is also found that at very short time delays and irrespective of the synaptic strengths, electrical synapses are better optimizers of SISR than chemical synapses. Meanwhile, at very long time delays, and irrespective of the synaptic strengths, chemical synapses are a better optimizers of SISR than electrical synapses. The expressions of electrical and chemical interaction potentials together with the minimum and maximum values of the noise amplitude within which an optimized SISR can occur are used to provide a theoretical explanation of the above effects.

After identifying the electrical and chemical synaptic strengths and time delays that destroy (or optimize) SISR in an isolated layer, we proceeded with identifying multiplexing configurations between the two layers that would optimize SISR in the second layer where SISR would be very poor or non-existent in isolation. For this identification, the synaptic parameters of one layer is configured such that SISR is optimal and this layer is multiplexed with a second layer where synaptic parameters are such that SISR is very poor or even non-existent. We then investigated which multiplexing connection (i,e., electrical, inhibitory chemical, or excitatory chemical synapses) is a better optimizer of SISR in the second layer.

In the first optimization configuration, we were interested in optimizing SISR in an electrically coupled layer (i.e., a layer where neurons are coupled only via electrical synapses) by multiplexing this layer with another electrically coupled layer. We found that even weak multiplexing with electrical synaptic connections may optimize SISR in the layer where SISR was even absent in isolation. However, the longer the multiplexing electrical synaptic time delay is, the less efficient this configuration becomes in optimizing SISR. In a second scenario, the multiplexing connection was mediated by inhibitory chemical synapses between these electrical layers. Here, we found that only very long multiplexing inhibitory chemical synaptic time delays at weak (but not too weak) synaptic strength may optimize SISR in the layer where it was non-existent in isolation. And in the third scenario, the multiplexing connection was mediated by excitatory chemical synapses between these electrical layers. It is found that only very short multiplexing excitatory chemical time delays at intermediate synaptic strengths can optimize SISR in the layer where the phenomenon is non-existent in isolation.

In the second optimization configuration, we were interested in optimizing SISR in an (inhibitory) chemically coupled layer (i.e., a layer where neurons are coupled only via inhibitory chemical synapses) by multiplexing this layer with another (inhibitory) chemically coupled layer. Here it is found that the optimization of SISR based multiplexing between chemical layers does work equally well as in the case of the multiplexing between electrical layers. Multiplexing of the chemical layers by electrical synapses and inhibitory chemical synapses cannot optimize SISR at all in the chemical layer, where in isolation SIRS is otherwise very poor. We found that only multiplexing excitatory chemical synapses (using a strong synaptic coupling and short time delay regime) can optimize SISR in the chemical layer, whereas SISR in isolation is very poor.

Comparing the first and the second optimization configurations of SISR, we conclude that the optimization of SISR is generally better in layers with electrically coupled neurons rather than with chemically coupled neurons, provided that multiplexing connections between the layers are either electrical or inhibitory chemical synapses. Vice versa, optimization of SISR is generally better in layers with (inhibitory) chemically coupled neurons than with electrically coupled neurons, when multiplexing connections between the layers are excitatory chemical synapses.

The manipulation of chemical and electrical patterns in the brain has become more accessible, either via drugs that cross the blood brain barrier, via electrical stimulation delivered through electrodes implanted in the brain, or via light delivered through optical fibers selectively exciting genetically manipulated neurons (Runnova et al., 2016; De Domenico, 2017; Andreev et al., 2018); however, the manipulation of the functional connectivity seems to be a more difficult goal to achieve. Our approach of modeling multi-layer networks in combination with stochastic dynamics offers a novel perspective on the modeling of the brain's structural and functional connectivity. We therefore expect that our findings could provide promising applications in controlling synaptic connections to optimize neural information (generated by noise-induced phenomena like SISR and CR) processing in experiments, surgery involving brain networks stimulation, and in designing networks of artificial neural circuits to optimize information processing via SISR (Eberhardt et al., 1989; Moopenn et al., 1989, 1990).

Interesting future research directions on the topic would be to investigate the optimization performances of electrical and chemical synapses in other intra-layer topologies like small-world network, scale-free network, and random network; and other inter-layer topologies like the multiplex topology in which neurons in one layer are connected (randomly) to more than one neuron in the other layer.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

MY designed the study, conducted the analysis and simulations, and wrote the manuscript. PH conducted the analysis and wrote the manuscript. EM conducted the analysis and wrote the manuscript. All authors contributed to the article and approved the submitted version.

ACKNOWLEDGMENTS

MY would like to acknowledge the warm hospitality at Department of Applied Mathematics and Computer Science of the Technical University of Denmark and for financial support.

REFERENCES

Abrams, D. M., Mirollo, R., Strogatz, S. H., and Wiley, D. A. (2008). Solvable model for chimera states of coupled oscillators. *Phys. Rev. Lett*. 101:84103. doi: 10.1103/PhysRevLett.101.084103

Andreev, A. V., Makarov, V. V., Runnova, A. E., Pisarchik, A. N., and Hramov, A. E. (2018). Coherence resonance in stimulated neuronal network. *Chaos Solit. Fract*. 106, 80–85. doi: 10.1016/j.chaos.2017.11.017

Andrzejak, R. G., Rummel, C., Mormann, F., and Schindler, K. (2016). All together now: analogies between chimera state collapses and epileptic seizures. *Sci. Rep*. 6:23000. doi: 10.1038/srep23000

Andrzejak, R. G., Ruzzene, G., and Malvestio, I. (2017). Generalized synchronization between chimera states. *Chaos* 27:053114. doi: 10.1063/1.4983841

Battiston, F., Nicosia, V., Chavez, M., and Latora, V. (2017). Multilayer motif analysis of brain networks. *Chaos* 27:047404. doi: 10.1063/1.4979282

Beato, V., Sendina-Nadal, I., Gerdes, I., and Engel, H. (2007). Coherence resonance in a chemical excitable system driven by coloured noise. *Philos. Trans. R. Soc. A* 366, 381–395. doi: 10.1098/rsta.2007.2096

Benzi, R., Sutera, A., and Vulpiani, A. (1981). The mechanism of stochastic resonance. *J. Phys. A* 14:L453. doi: 10.1088/0305-4470/14/11/006

Bera, B. K., Majhi, S., Ghosh, D., and Perc, M. (2017). Chimera states: effects of different coupling topologies. *Europhys. Lett*. 118:10001. doi: 10.1209/0295-5075/118/10001

Bera, B. K., Rakshit, S., Ghosh, D., and Kurths, J. (2019). Spike chimera states and firing regularities in neuronal hypernetworks. *Chaos* 29:053115. doi: 10.1063/1.5088833

Berner, R., Sawicki, J., and Schöll, E. (2020). Birth and stabilization of phase clusters by multiplexing of adaptive networks. *Phys. Rev. Lett*. 124:088301. doi: 10.1103/PhysRevLett.124.088301

Bick, C., Goodfellow, M., Laing, C. R., and Martens, E. A. (2020). Understanding the dynamics of biological and neural oscillator networks through exact mean-field reductions: a review. *J. Math. Neurosc.* 10, 1–43. doi: 10.1186/s13408-020-00086-9

Bick, C., and Martens, E. A. (2015). Controlling chimeras. *New J. Phys*. 17:033030. doi: 10.1088/1367-2630/17/3/033030

Bing, J., Hua-Guang, G., and Yu-Ye, L. (2011). Coherence-resonance-induced neuronal firing near a saddle-node and homoclinic bifurcation corresponding to type-i excitability. *Chin. Phys. Lett*. 28:090507. doi: 10.1088/0256-307X/28/9/090507

Booth, V., and Behn, C. G. D. (2014). Physiologically-based modeling of sleep-wake regulatory networks. *Math. Biosci*. 250, 54–68. doi: 10.1016/j.mbs.2014.01.012

Bukh, A., Rybalova, E., Semenova, N., Strelkova, G., and Anishchenko, V. (2017). New type of chimera and mutual synchronization of spatiotemporal structures in two coupled ensembles of nonlocally interacting chaotic maps. *Chaos* 27:111102. doi: 10.1063/1.5009375

Connors, B. W., and Long, M. A. (2004). Electrical synapses in the mammalian brain. *Annu. Rev. Neurosci*. 27, 393–418. doi: 10.1146/annurev.neuro.26.041002.131128

Crofts, J. J., Forrester, M., and O'Dea, R. D. (2016). Structure-function clustering in multiplex brain networks. *Europhys. Lett*. 116:18003. doi: 10.1209/0295-5075/116/18003

De Domenico, M. (2017). Multilayer modeling and analysis of human brain networks. *Giga Sci*. 6:gix004. doi: 10.1093/gigascience/gix004

Deschle, N., Daffertshofer, A., Battaglia, D., and Martens, E. A. (2019). Directed flow of information in chimera states. *Front. Appl. Math. Stat*. 5, 1–8. doi: 10.3389/fams.2019.00028

DeVille, R. L., and Vanden-Eijnden, E. (2007). A nontrivial scaling limit for multiscale markov chains. *J. Stat. Phys*. 126, 75–94. doi: 10.1007/s10955-006-9237-1

DeVille, R. L., Vanden-Eijnden, E., et al. (2007). Self-induced stochastic resonance for brownian ratchets under load. *Commun. Math. Sci*. 5, 431–466. doi: 10.4310/CMS.2007.v5.n2.a10

DeVille, R. L., Vanden-Eijnden, E., and Muratov, C. B. (2005). Two distinct mechanisms of coherence in randomly perturbed dynamical systems. *Phys. Rev. E* 72:031105. doi: 10.1103/PhysRevE.72.031105

Eberhardt, S. P., Duong, T., and Thakoor, A. (1989). "A VLSI building block chip for hardware neural network implementations," in *Proceedings Third Annual Parallel Processing Symposium, Vol. 1* (Fullerton, CA), 257–267.

FitzHugh, R. (1961). Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.* 1, 445–466. doi: 10.1016/S0006-3495(61)86902-6

Freidlin, M. I. (2001). On stable oscillations and equilibriums induced by small noise. *J. Stat. Phys*. 103, 283–300. doi: 10.1023/A:1004827921214

Galarreta, M., and Hestrin, S. (1999). A network of fast-spiking cells in the neocortex connected by electrical synapses. *Nature* 402:72. doi: 10.1038/47029

Galarreta, M., and Hestrin, S. (2001). Electrical synapses between gaba-releasing interneurons. *Nat. Rev. Neurosci.* 2:425. doi: 10.1038/35077566

Gambuzza, L. V., Frasca, M., and Gomez-Garde nes, J. (2015). Intra-layer synchronization in multiplex networks. *Europhys. Lett*. 110:20010. doi: 10.1209/0295-5075/110/20010

Gammaitoni, L., Hänggi, P., Jung, P., and Marchesoni, F. (1998). Stochastic resonance. *Rev. Modern Phys*. 70:223. doi: 10.1103/RevModPhys.70.223

Ghosh, S., and Jalan, S. (2016). Emergence of chimera in multiplex network. *Int. J. Bifurcat. Chaos* 26:1650120. doi: 10.1142/S0218127416501200

Ghosh, S., Kumar, A., Zakharova, A., and Jalan, S. (2016). Birth and death of chimera: interplay of delay and multiplexing. *Europhys. Lett*. 115:60005. doi: 10.1209/0295-5075/115/60005

Ghosh, S., Schülen, L., Kachhvah, A. D., Zakharova, A., and Jalan, S. (2019). Taming chimeras in networks through multiplexing delays. *Europhys. Lett*. 127:30002. doi: 10.1209/0295-5075/127/30002

Ghosh, S., Zakharova, A., and Jalan, S. (2018). Non-identical multiplexing promotes chimera states. *Chaos Solit. Fract*. 106, 56–60. doi: 10.1016/j.chaos.2017.11.010

Gibson, J. R., Beierlein, M., and Connors, B. W. (1999). Two networks of electrically coupled inhibitory neurons in neocortex. *Nature* 402:75. doi: 10.1038/47035

Greengard, P. (2001). The neurobiology of slow synaptic transmission. *Science* 294, 1024–1030. doi: 10.1126/science.294.5544.1024

Gu, H., Zhang, H., Wei, C., Yang, M., Liu, Z., and Ren, W. (2011). Coherence resonance-induced stochastic neural firing at a saddle-node bifurcation. *Int. J. Modern Phys. B* 25, 3977–3986. doi: 10.1142/S0217979211101673

Guo, D., Perc, M., Liu, T., and Yao, D. (2018). Functional importance of noise in neuronal information processing. *Europhys. Lett*. 124:50001. doi: 10.1209/0295-5075/124/50001

Hestrin, S., and Galarreta, M. (2005). Electrical synapses define networks of neocortical gabaergic neurons. *Trends Neurosci*. 28, 304–309. doi: 10.1016/j.tins.2005.04.001

Hizanidis, J., and Schöll, E. (2008). Control of coherence resonance in semiconductor superlattices. *Phys. Rev. E* 78:066205. doi: 10.1103/PhysRevE.78.066205

Hodgkin, A. L., and Huxley, A. F. (1952). A quantitative description of membrane current and its application to conduction and excitation in nerve. *J. Physiol*. 117, 500–544. doi: 10.1113/jphysiol.1952.sp004764

Hormuzdi, S. G., Filippov, M. A., Mitropoulou, G., Monyer, H., and Bruzzone, R. (2004). Electrical synapses: a dynamic signaling system that shapes the activity of neuronal networks. *Biochim. Biophys. Acta* 1662, 113–137. doi: 10.1016/j.bbamem.2003.10.023

Hu, J., and MacDonald, A. H. (1993). Two-dimensional vortex lattice melting. *Phys. Rev. Lett*. 71:432. doi: 10.1103/PhysRevLett.71.432

Izhikevich, E. M. (2000). Neural excitability, spiking and bursting. *Int. J. Bifurcat. Chaos* 10, 1171–1266. doi: 10.1142/S0218127400000840

Kasdin, N. J. (1995). Runge-kutta algorithm for the numerical integration of stochastic differential equations. *J. Guid. Control Dyn*. 18, 114–120. doi: 10.2514/3.56665

Kouvaris, N. E., Hata, S., and Díaz-Guilera, A. (2015). Pattern formation in multiplex networks. *Sci. Rep*. 5:10840. doi: 10.1038/srep10840

Kurrer, C., and Schulten, K. (1995). Noise-induced synchronous neuronal oscillations. *Phys. Rev. E* 51:6213. doi: 10.1103/PhysRevE.51.6213

Leyva, I., Sevilla-Escoboza, R., Sendina-Nadal, I., Gutierrez, R., Buldu, J., and Boccaletti, S. (2017). Inter-layer synchronization in nonidentical multi-layer networks. *Sci. Rep*. 7:45475. doi: 10.1038/srep45475

Lindner, B., García-Ojalvo, J., Neiman, A., and Schimansky-Geier, L. (2004). Effects of noise in excitable systems. *Phys. Rep*. 392, 321–424. doi: 10.1016/j.physrep.2003.10.015

Lindner, B., and Schimansky-Geier, L. (1999). Analytical approach to the stochastic fitzhugh-nagumo system and coherence resonance. *Phys. Rev. E* 60:7270. doi: 10.1103/PhysRevE.60.7270

Liu, Z.-Q., Zhang, H.-M., Li, Y.-Y., Hua, C.-C., Gu, H.-G., and Ren, W. (2010). Multiple spatial coherence resonance induced by the stochastic signal in neuronal networks near a saddle-node bifurcation. *Phys. A* 389, 2642–2653. doi: 10.1016/j.physa.2010.02.029

Longtin, A. (1993). Stochastic resonance in neuron models. *J. Stat. Phys*. 70, 309–327. doi: 10.1007/BF01053970

Majhi, S., Bera, B. K., Ghosh, D., and Perc, M. (2019). Chimera states in neuronal networks: a review. *Phys. Life Rev*. 28, 100–121. doi: 10.1016/j.plrev.2018.09.003

Majhi, S., Perc, M., and Ghosh, D. (2016). Chimera states in uncoupled neurons induced by a multilayer structure. *Sci. Rep*. 6:39033. doi: 10.1038/srep39033

Majhi, S., Perc, M., and Ghosh, D. (2017). Chimera states in a multilayer network of coupled and uncoupled neurons. *Chaos* 27:073109. doi: 10.1063/1.4993836

Maksimenko, V. A., Makarov, V. V., Bera, B. K., Ghosh, D., Dana, S. K., Goremyko, M. V., et al. (2016). Excitation and suppression of chimera states by multiplexing. *Phys. Rev. E* 94:052205. doi: 10.1103/PhysRevE.94.052205

Martens, E. A. (2010). Bistable chimera attractors on a triangular network of oscillator populations. *Phys. Rev. E* 82:016216. doi: 10.1103/PhysRevE.82.016216

Martens, E. A., Bick, C., and Panaggio, M. J. (2016a). Chimera states in two populations with heterogeneous phase-lag. *Chaos* 26:094819. doi: 10.1063/1.4958930

Martens, E. A., Panaggio, M. J., and Abrams, D. M. (2016b). Basins of attraction for chimera states. *New J. Phys. Fast Track Commun*. 18:022002. doi: 10.1088/1367-2630/18/2/022002

Martens, E. A., Thutupalli, S., Fourriére, A., and Hallatschek, O. (2013). Chimera states in mechanical oscillator networks. *Proc. Natl. Acad. Sci. U.S.A*. 110, 10563–10567. doi: 10.1073/pnas.1302880110

Masoliver, M., Malik, N., Schöll, E., and Zakharova, A. (2017). Coherence resonance in a network of FitzHugh-Nagumo systems: interplay of noise, time-delay, and topology. *Chaos* 27:101102. doi: 10.1063/1.5003237

Mikhaylenko, M., Ramlow, L., Jalan, S., and Zakharova, A. (2019). Weak multiplexing in neural networks: switching between chimera and solitary states. *Chaos* 29:023122. doi: 10.1063/1.5057418

Moopenn, A., Duong, T., and Thakoor, A. (1990). "Digital-analog hybrid synapse chips for electronic neural networks," in *Advances in Neural Information Processing Systems* (San Francisco, CA), 769–776.

Moopenn, A., Thakoor, A. P., and Duong, T. (1989). "Digital-analog-hybrid neural simulator: a design-aid for custom-VLSI neurochips," in *High Speed Computing II, Vol. 1058* (Los Angeles, CA: International Society for Optics and Photonics), 147–157. doi: 10.1117/12.951677

Muratov, C. B., Vanden-Eijnden, E., and Weinan, E. (2005). Self-induced stochastic resonance in excitable systems. *Phys. D* 210, 227–240. doi: 10.1016/j.physd.2005.07.014

Neiman, A., Saparin, P. I., and Stone, L. (1997). Coherence resonance at noisy precursors of bifurcations in nonlinear dynamical systems. *Phys. Rev. E* 56:270. doi: 10.1103/PhysRevE.56.270

Omelchenko, I., Hülser, T., Zakharova, A., and Schöll, E. (2018). Control of chimera states in multilayer networks. *Front. Appl. Math. Stat*. 4:67. doi: 10.3389/fams.2018.00067

Panaggio, M. J., and Abrams, D. M. (2015). Chimera states: coexistence of coherence and incoherence in networks of coupled oscillators. *Nonlinearity* 28:R67. doi: 10.1088/0951-7715/28/3/R67

Patriarca, M., Postnova, S., Braun, H. A., Hernández-García, E., and Toral, R. (2012). Diversity and noise effects in a model of homeostatic regulation of the sleep-wake cycle. *PLoS Comput. Biol*. 8:e1002650. doi: 10.1371/journal.pcbi.1002650

Pei, X., Wilkens, L., and Moss, F. (1996). Noise-mediated spike timing precision from aperiodic stimuli in an array of hodgekin-huxley-type neurons. *Phys. Rev. Lett*. 77:4679. doi: 10.1103/PhysRevLett.77.4679

Pereda, A. E. (2014). Electrical synapses and their functional interactions with chemical synapses. *Nat. Rev. Neurosci*. 15:250. doi: 10.1038/nrn3708

Pikovsky, A. S., and Kurths, J. (1997). Coherence resonance in a noise-driven excitable system. *Phys. Rev. Lett*. 78:775. doi: 10.1103/PhysRevLett.78.775

Pisarchik, A. N., Jaimes-Reátegui, R., Magallón-García, C. A., and Castillo-Morales, C. O. (2014). Critical slowing down and noise-induced intermittency in bistable perception: bifurcation analysis. *Biol. Cybern*. 108, 397–404. doi: 10.1007/s00422-014-0607-5

Runnova, A. E., Hramov, A. E., Grubov, V. V., Koronovskii, A. A., Kurovskaya, M. K., and Pisarchik, A. N. (2016). Theoretical background and experimental measurements of human brain noise intensity in perception of ambiguous images. *Chaos Solit. Fract*. 93, 201–206. doi: 10.1016/j.chaos.2016.11.001

Sawicki, J., Ghosh, S., Jalan, S., and Zakharova, A. (2019). Chimeras in multiplex networks: interplay of inter-and intra-layer delays. *Front. Appl. Math. Stat*. 5:19. doi: 10.3389/fams.2019.00019

Schmidt, H., Avitabile, D., Montbrió, E., and Roxin, A. (2018). Network mechanisms underlying the role of oscillations in cognitive tasks. *PLoS Comput. Biol*. 14:e1006430. doi: 10.1371/journal.pcbi.1006430

Schöll, E. (2016). Synchronization patterns and chimera states in complex networks: Interplay of topology and dynamics. *Eur. Phys. J. Spec. Top*. 225, 891–919. doi: 10.1140/epjst/e2016-02646-3

Schöll, E., Hiller, G., Hövel, P., and Dahlem, M. A. (2009). Time-delayed feedback in neurosystems. *Philos. Trans. R. Soc. A* 367, 1079–1096. doi: 10.1098/rsta.2008.0258

Semenova, N., and Zakharova, A. (2018). Weak multiplexing induces coherence resonance. *Chaos* 28:051104. doi: 10.1063/1.5037584

Singh, A., Ghosh, S., Jalan, S., and Kurths, J. (2015). Synchronization in delayed multiplex networks. *Europhys. Lett*. 111:30010. doi: 10.1209/0295-5075/111/30010

Tinsley, M. R., Nkomo, S., and Showalter, K. (2012). Chimera and phase-cluster states in populations of coupled chemical oscillators. *Nat. Phys*. 8, 662–665. doi: 10.1038/nphys2371

Totz, J. F., Rode, J., Tinsley, M. R., Showalter, K., and Engel, H. (2017). Spiral wave chimera states in large populations of coupled chemical oscillators. *Nat. Phys*. 14, 282–285. doi: 10.1038/s41567-017-0005-8

Yamakou, M. E., and Jost, J. (2017). A simple parameter can switch between different weak-noise-induced phenomena in a simple neuron model. *Europhys. Lett*. 120:18002. doi: 10.1209/0295-5075/120/18002

Yamakou, M. E., and Jost, J. (2018). Coherent neural oscillations induced by weak synaptic noise. *Nonlin. Dyn*. 93, 2121–2144. doi: 10.1007/s11071-018-4311-1

Yamakou, M. E., and Jost, J. (2019). Control of coherence resonance by self-induced stochastic resonance in a multiplex neural network. *Phys. Rev. E* 100:022313. doi: 10.1103/PhysRevE.100.022313

Yilmaz, E., Ozer, M., Baysal, V., and Perc, M. (2016). Autapse-induced multiple coherence resonance in single neurons and neuronal networks. *Sci. Rep*. 6:30914. doi: 10.1038/srep30914

Yilmaz, E., Uzuntarla, M., Ozer, M., and Perc, M. (2013). Stochastic resonance in hybrid scale-free neuronal networks. *Phys. A* 392, 5735–5741. doi: 10.1016/j.physa.2013.07.011

Zhang, H., He, Q., and Kong, F. (2015). Stochastic resonance in an underdamped system with pinning potential for weak signal detection. *Sensors* 15, 21169–21195. doi: 10.3390/s150921169

Zhang, L., Motter, A. E., and Nishikawa, T. (2017). Incoherence-mediated remote synchronization. *Phys. Rev. Lett*. 118:174102. doi: 10.1103/PhysRevLett.118.174102

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.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 31 August 2020
doi: 10.3389/fncom.2020.00077

Impact of Physical Obstacles on the Structural and Effective Connectivity of *in silico* Neuronal Circuits

Adriaan-Alexander Ludl[1,2,3]* *and Jordi Soriano*[2,3]*

[1] *Computational Biology Unit, Department of Informatics, University of Bergen, Bergen, Norway,* [2] *Departament de Física de la Matèria Condensada, Universitat de Barcelona, Barcelona, Spain,* [3] *Universitat de Barcelona Institute of Complex Systems (UBICS), Barcelona, Spain*

Scaffolds and patterned substrates are among the most successful strategies to dictate the connectivity between neurons in culture. Here, we used numerical simulations to investigate the capacity of physical obstacles placed on a flat substrate to shape structural connectivity, and in turn collective dynamics and effective connectivity, in biologically-realistic neuronal networks. We considered μ-sized obstacles placed in mm-sized networks. Three main obstacle shapes were explored, namely crosses, circles and triangles of isosceles profile. They occupied either a small area fraction of the substrate or populated it entirely in a periodic manner. From the point of view of structure, all obstacles promoted short length-scale connections, shifted the in- and out-degree distributions toward lower values, and increased the modularity of the networks. The capacity of obstacles to shape distinct structural traits depended on their density and the ratio between axonal length and substrate diameter. For high densities, different features were triggered depending on obstacle shape, with crosses trapping axons in their vicinity and triangles funneling axons along the reverse direction of their tip. From the point of view of dynamics, obstacles reduced the capacity of networks to spontaneously activate, with triangles in turn strongly dictating the direction of activity propagation. Effective connectivity networks, inferred using transfer entropy, exhibited distinct modular traits, indicating that the presence of obstacles facilitated the formation of local effective microcircuits. Our study illustrates the potential of physical constraints to shape structural blueprints and remodel collective activity, and may guide investigations aimed at mimicking organizational traits of biological neuronal circuits.

Keywords: network formation, simulations, patterned networks, structural connectivity, effective connectivity, network bursts, modularity, network measures

OPEN ACCESS

Edited by:
Tatyana Sharpee,
Salk Institute for Biological Studies, United States

Reviewed by:
Ernest Barreto,
George Mason University, United States
Marko Gosak,
University of Maribor, Slovenia

***Correspondence:**
Adriaan-Alexander Ludl
adriaan.ludl@fmc.ub.edu
Jordi Soriano
jordi.soriano@ub.edu

Received: *29 April 2020*
Accepted: *21 July 2020*
Published: *31 August 2020*

Citation:
Ludl A-A and Soriano J (2020) Impact of Physical Obstacles on the Structural and Effective Connectivity of in silico Neuronal Circuits. Front. Comput. Neurosci. 14:77. doi: 10.3389/fncom.2020.00077

1. INTRODUCTION

Naturally formed biological neuronal networks are characterized by an intricate spatial organization that is central to ensure the functionality of the neuronal circuits (Achard and Bullmore, 2007; Bullmore and Sporns, 2012). The brain's cortex for instance is arranged in columns and hyper-columns that shape structural and functional modules that conduct specialized tasks. The abnormal formation of neuronal circuits during development or their damage due to disease are known to substantially alter circuits' activity patterns. It is therefore well-accepted

that the structure of a neuronal circuit shapes its dynamics in great measure. Although a direct relationship between structure and dynamics cannot be established given the intrinsic non-linear nature of neuronal circuits and the coexistence of diverse dynamic physiological mechanisms, there is a wealth of evidence indicating direct correspondences between key structural traits and dynamics (Honey et al., 2010; Sporns, 2011). These traits emerge from general constraints imposed by the spatial embedding of brain circuits (Bullmore and Sporns, 2012; Stiso and Bassett, 2018) together with specific topological characteristics such as high clustering, modularity and the existence of central hub nodes (Sporns, 2011). It has been suggested that these traits and even network motifs can in part be explained from the trade-off between topological integration and the biological cost incurred by nervous systems (Schröter et al., 2017).

In the quest to understand the relationship between structure and dynamics, in particular the importance of structural traits, numerical simulations and *in vitro* studies of neuronal cultures have emerged as invaluable tools. On the one hand, numerical models have been employed to explore various configurations ranging from small-scale circuits (Voges and Perrinet, 2012; Orlandi et al., 2013; Pernice et al., 2013; Faci-Lázaro et al., 2019) to whole-brain dynamics (Honey et al., 2007; Messé et al., 2014; Cabral et al., 2017). Messé et al. for instance used elaborate computational models and anatomical brain data to predict the activity patterns observed in resting-state functional magnetic resonance imaging, and concluded that the backbone of anatomical connectivity strongly shaped overall dynamical traits. Neuronal cultures, on the other hand, have helped elucidate the importance of spatial embedding and imposed metric correlations in shaping spontaneous activity (Orlandi et al., 2013; Hernández-Navarro et al., 2017; Okujeni et al., 2017; Tibau et al., 2020), the impact of modular organization (Shein-Idelson et al., 2011; Tang-Schomer et al., 2014; Yamamoto et al., 2018), the emergence of small-worldness (Downes et al., 2012; de Santos-Sierra et al., 2014), or the role of hubs (Schroeter et al., 2015).

The above studies demonstrated that non-random structural characteristics are central to shape distinct activity patterns and, in turn, specific functional connectivity traits. However, an interesting aspect still to be explored in detail is the impact of definite structural motifs on global network dynamics. This is particularly relevant in the context of *engineered* neuronal cultures (Aebersold et al., 2016), in which the spatial arrangement of neurons and connections is dictated by chemical or physical constraints. Microfabricated structures or *scaffolds* have revolutionized the concept of engineered neuronal cultures by providing both connectivity guidance and structural support to two- and three-dimensional neuronal assemblies (Kunze et al., 2011; Bosi et al., 2015; Severino et al., 2016; Larramendy et al., 2019).

In an effort to help understanding how scaffolds, or specific structural motifs, shape the blueprint, dynamics and effective connectivity of neuronal cultures, we explored numerically small two-dimensional neuronal networks similar to biological *in vitro* ones which incorporated specific scaffold designs in the form of arrays of obstacles. We considered μ-sized scaffolds embedded in a mm-size substrate. Three designs with distinct geometries were explored to examine whether they could imprint specific structural and dynamic features to the networks. The studied obstacles were crosses, circles and isosceles triangles. They were designed to facilitate the trapping or deflection of axons (crosses), to gently modulate connectivity across the network (circles) and to dictate the directionality of connectivity (triangles). We selected these shapes in view of recent experimental studies aimed at guiding neuronal connectivity through microfabrication technology (Crowe et al., 2020). We observed that the obstacles molded structural connectivity at short and long length scales. This induced characteristic features of network dynamics and of effective connectivity. Our study can be extended to tailored designs that mimic specific experimental configurations. Thus, it can improve predictions of the action of scaffolds on living neuronal circuits, for instance to tailor specific dynamic patterns or network functionality.

2. RESULTS

2.1. Impact of Obstacle Shape on Structural Connectivity

We explored *in silico* neuronal networks with spatial constraints by considering different sets of obstacles arranged on a circular area of either 2 or 4 mm in diameter. This size was selected to mimic the characteristic size of small *in vitro* cultures (Orlandi et al., 2013; Tibau et al., 2020). In the simulations, neurons were laid out on the surface in a homogeneous manner and connected to one another following a geometric model as in Orlandi et al. (2013), in which the axons grew as concatenated segments according to a biased random walk (**Figure 1A**) and that is known to mimic well the behavior of individual axons (Feinerman et al., 2008). The presence of obstacles altered axonal growth, an aspect that was modeled by reflecting the axon with the same angle of incidence upon contact with an obstacle (**Figures 1A,B**). This "reflection rule" was inspired by experimental observations in cultures of physically-constrained neurons (Feinerman et al., 2008; Gladkov et al., 2017) and was the simplest way to introduce interaction with obstacles for this biased random walk. More biologically-accurate models, in which axons may attach to the walls or follow the path of previous axons (Simitzi et al., 2017) were disregarded for the sake of simplicity. We considered three characteristic sets of obstacles, namely crosses, circles and triangles of isosceles profile (**Figure 1C**), that either occupied a small fraction of the available area or populated it entirely. **Table 1** and **Figure 2** summarize the different designs chosen and their major characteristics. The density of neurons in the simulations in all configurations was maintained constant at 200 neurons/mm^2, leading to networks with 625 and 2,500 neurons for the 2 and 4 mm diameter sizes, respectively.

The shape of the obstacles had an important effect on the paths followed by the axons and on the capacity of neurons to connect to one another. The bottom panels of **Figure 2** show a detail of the positions of neurons and axons and the

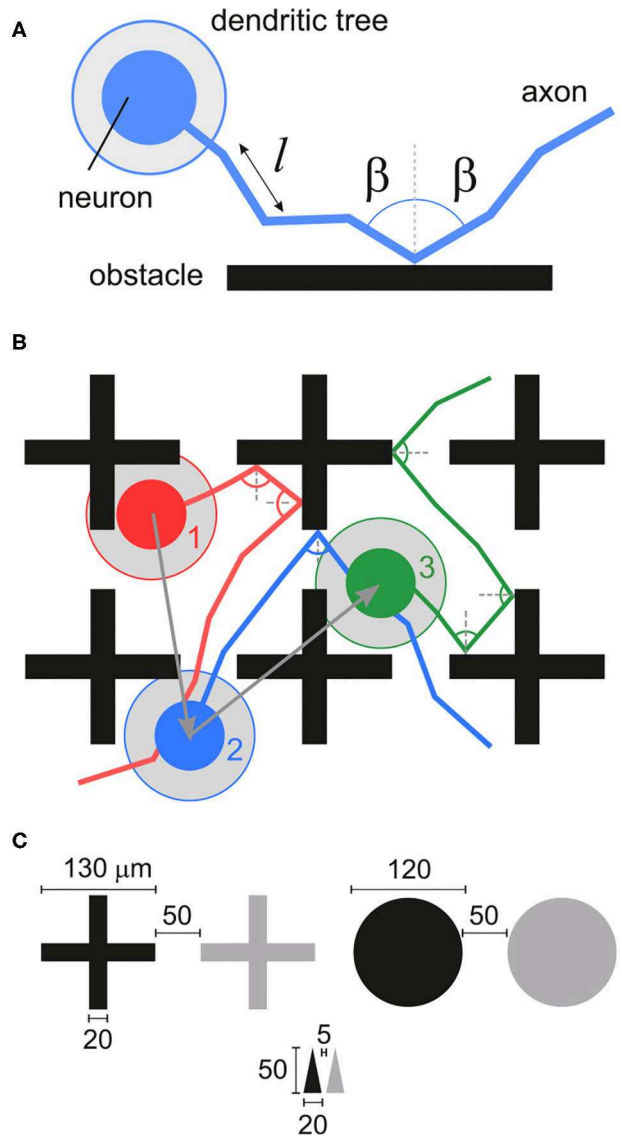

FIGURE 1 | Neuronal network construction. **(A)** Axons are laid out on the substrate by concatenating segments of length *l* that follow a quasistraight path. Whenever the axon hits an obstacle it is deflected with the same angle of incidence. The dendritic tree of the neuron is modeled as a circle (gray area) with a given radius of interaction. **(B)** Neuronal axons in the studied networks interact with obstacles and connect to other neurons. A connection $i \rightarrow j$ is established whenever the axon of a neuron *i* crosses the dendritic tree of another neuron *j*. In the sketch, dark gray arrows indicate the connections and their direction, with neurons connecting as $1 \rightarrow 2$ and $2 \rightarrow 3$. **(C)** Sketch of the obstacle geometries used and their dimensions.

interaction of the latter with the obstacles. Crosses (**Figure 2D**) tended to either deflect axons or to trap them in their vicinity, thus potentially inducing strong local inhomogeneities in the connectivity of the network. Circles (**Figure 2E**) had a milder effect, deflecting the axons toward the neighborhood, but causing alterations in the connectivity due to the relatively large area that they occupied, reducing the probability of spatially close neurons to interconnect. Finally, triangles shaped as arrowheads pointing upwards (**Figure 2F**) promoted a strong anisotropy in the connectivity by funneling the axons reverse in the direction opposite to the triangles' tips. This is because axons had a much higher probability to be deflected at the base of a triangle than at its tip. Effectively, as illustrated in **Figure 2F**, most axons were vertically aligned—although some orthogonal growth remained—and thus neurons tended to connect vertically and downwards.

To quantify the impact of each configuration on network characteristics we analyzed the topological traits of the resulting structural connectivities. **Figure 3A** shows representative structural adjacency matrices of the empty configuration together with the configurations made of crosses, circles, and triangles that fully cover the available area. Neuron indices in the matrices are arranged to highlight the existence of communities along the diagonal. We note that communities already appear in the empty configuration (modularity $Q \simeq 0.37$), a trait that is due to the presence of metric correlations in spatially embedded networks (Hernández-Navarro et al., 2017; Faci-Lázaro et al., 2019) which facilitates the formation of local neuronal microcircuits. The global efficiency is relatively high ($G_{\text{eff}} \simeq 0.54$), indicating that the neurons in the network are well bound together despite spatial effects. The presence of obstacles in the networks in general increased Q and decreased G_{eff}, which reveals a strengthening of metric effects and a reduced capacity for the neurons to connect to one another. The impact of obstacles on structural connectivity depended on their shape. Crosses exhibited the strongest impact, with an increase of Q by 43%, while for the other configurations the increase was by 27% (circles) and 30% (triangles). We argue that the trapping of axons caused by the crosses is the cause of the high increase in Q for this configuration.

The number and size of structural communities was similar across the panel of configurations. This indicates that neurons were still capable of interconnecting to some degree despite the high spatial density of obstacles. In other words, structural microcircuits emerged but they were not fully isolated. This was verified by analyzing the spatial distribution of the observed communities (**Figure 3B**), which were physically compact but interlinked. Crosses and circles showed spatial features that were similar to the empty case, with communities appearing in patches of similar shape and size. The triangles configuration, however, shaped communities distinctively organized as vertical stripes and that revealed the strong capacity of triangles to dictate vertical funneling of axons.

To shed light on the impact of obstacles on neuron-to-neuron connectivity and network structure, we investigated the distributions of Euclidean connection distances d and angles θ of connections (**Figure 3C**). For the empty reference case, the distribution of distances was broad, with most of the neurons connecting in the range $0.1 - 1$ mm, although there was a marked peak at $d \simeq 0.15$ mm, a trait again due to the fact that nearby neurons are more likely to connect in spatially embedded networks. For crosses, however, the distribution was strongly shifted toward small connection distances, clearly indicating the capacity of the crosses to trap axons and boost

TABLE 1 | Network descriptors for 2 and 4 mm configurations.

	a_{obs}/a_{total}(%)	k_{in} μ	k_{in} σ	k_{out} μ	k_{out} σ	d (mm) m	d (mm) s.d.
2 mm							
Triangles	47.2	53.34	22.92	50.21	25.73	0.403	0.306
Circles	44.7	59.08	13.18	54.26	26.26	0.422	0.314
Crosses							
Empty	0.0	66.28	13.83	60.29	33.16	0.533	0.375
1 array	2.4	63.46	15.23	58.21	29.71	0.477	0.346
2 arrays	4.8	60.93	15.06	55.50	28.38	0.446	0.330
Full	13.6	47.98	11.37	45.49	18.07	0.312	0.238
4 mm (crosses)							
Empty	0.0	73.51	14.67	67.35	33.97	0.301	0.220
1 array	0.6	72.75	14.37	66.46	34.44	0.293	0.214
2 arrays	1.2	72.20	15.07	65.72	34.63	0.285	0.209
4 arrays	2.4	70.52	15.82	63.37	33.06	0.271	0.202

For each configuration, we provide the area fraction occupied by the obstacles (a_{obs}/a_{total}) as well as the average value (μ) and standard deviation (σ) obtained for the Gaussian fits to the distributions of in- and out-degrees (k_{in}, k_{out}), and the statistical average value (m) and standard deviation (s.d.) of the distribution of connection distances (d).

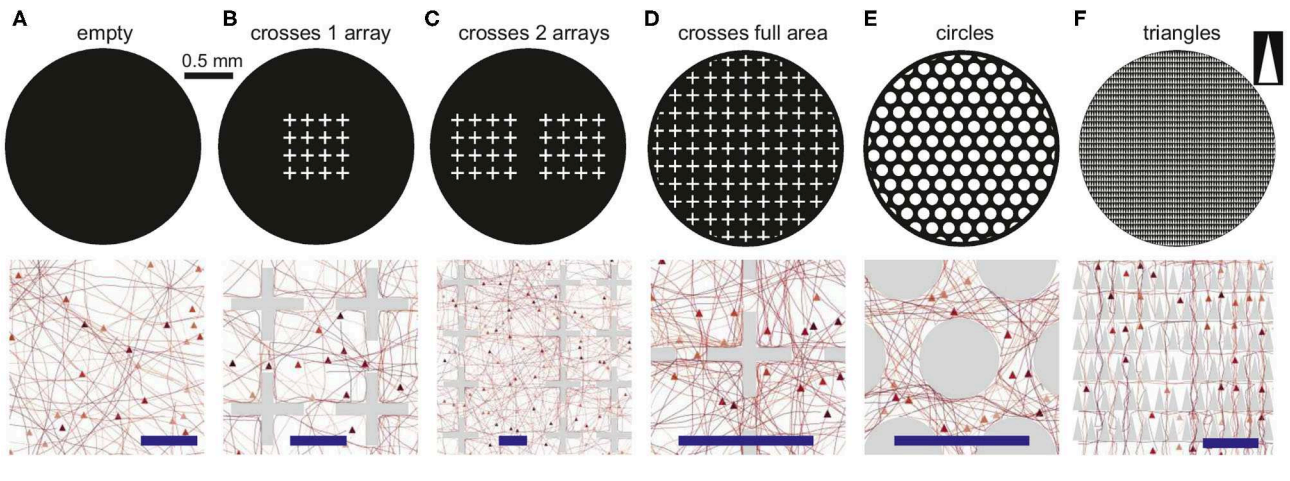

FIGURE 2 | Configurations of obstacles. The top row shows the black and white masks used to set up the simulations, with neurons and connections only placeable in the black areas. The bottom row shows details of the simulated networks, marking the location of neurons (arrowheads) and axons (lines). Obstacles are shown in gray and the blue scale bars are 100 μm. **(A)** Reference empty configuration. **(B,C)** Layouts with 1 array and 2 arrays of crosses. **(D–F)** Layouts of crosses, circles and upwards-pointing isosceles triangles fully covering the substrate.

short-range connectivity. Circle and triangle configurations exhibited a behavior in between the previous cases, with broader distributions than crosses but with characteristic peaks that are associated to the size and inter-spacing of the obstacles. On the other hand, the distribution of angles θ was in general homogeneous and similar across configurations except for triangles, with a characteristic peak at $\theta \simeq 180^\circ$ associated to the guided top-to-bottom connectivity in the network. Additional peaks appeared at $\theta \simeq 90^\circ$ and 270°, which revealed the existence of orthogonal connectivity that facilitated the entire network to be interlinked.

To further analyze the impact of obstacles on connectivity, we inspected the distributions of in-degrees (k_{in}) and out-degrees (k_{out}), and also looked in more detail at the distributions of connection distances d at different length-scales. The distributions shown in **Figure 4** represent averages over 12 replicates for each configuration with the statistical standard deviations shown by the shaded areas. The average value of each distribution and its statistical standard deviation are depicted at the bottom of each graph.

First, we compared the distributions among configurations of circles, triangles and crosses that fully populated the 2 mm substrate (**Figures 4A–C**). The deviations from the empty configuration were pronounced. Circles, on the one hand, showed k_{in} and k_{out} distributions (**Figures 4A,B**) that lay between those for empty and crosses configurations. This moderate impact contrasted with the existence of periodic peaks in the distribution of distances (**Figure 4C**). The first peak

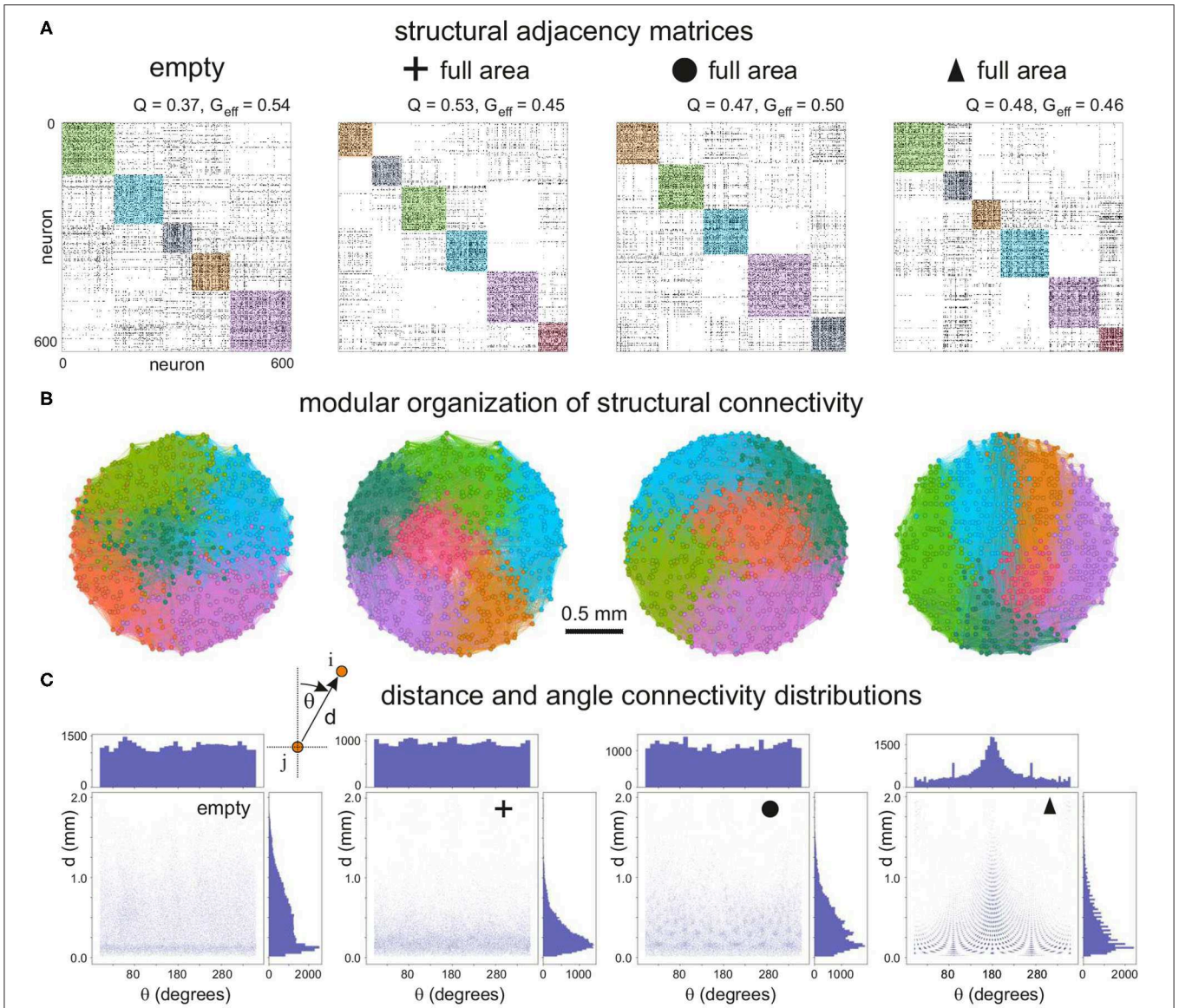

FIGURE 3 | Structural connectivity of 2 mm diameter networks with obstacles. **(A)** Representative connectivity matrices for the "empty" configuration and for networks filled with crosses, circles or triangles. Matrices are arranged to highlight modular structure (colored squares). The values above each matrix indicate the modularity Q and the global efficiency G_{eff}. **(B)** Spatial localization of the identified modules. The color coding matches that of the adjacency matrices. Modularity increases with the presence of obstacles, and the modules are distinctively vertically arranged for triangles. **(C)** Corresponding distribution of connection distances d and angles θ between all pairs of connected neurons in each configuration. For triangles there is a characteristic peak at 180°, indicating that most of the neurons connect downwards, i.e., reversed with respect to the triangles' orientation.

occurred at 170 μm, which is the distance between the centers of neighboring circles (120 μm circle diameter plus 50 μm separation), and the rest of the peaks are multiples of this typical distance. Thus, circles induced characteristic length scales in the network without strongly altering the degree distributions. Triangles, on the other hand, exhibited a shift of k_{in} toward lower values and a marked broadening of the distribution. Their effect on k_{out} was very similar to that of circles. Clearly, the capacity of the triangles to funnel axons along the substrate facilitated long-range connections, whereas the limited orthogonal growth promoted short-range ones. The distribution of distances for triangles (**Figure 4C**) also shows periodic peaks multiples of 50 μm, the triangle height. These peaks are sharper and steeper for triangles than for circles, the values of which lie in between those for triangles. Crosses showed strong effects as well, which we discuss in detail below.

For crosses, we considered the scenario in which they gradually covered a higher area fraction of the substrate, and compared the empty, 1 array, 2 arrays, and full coverage configurations. As shown in **Figures 4D,E**, both k_{in} and k_{out}

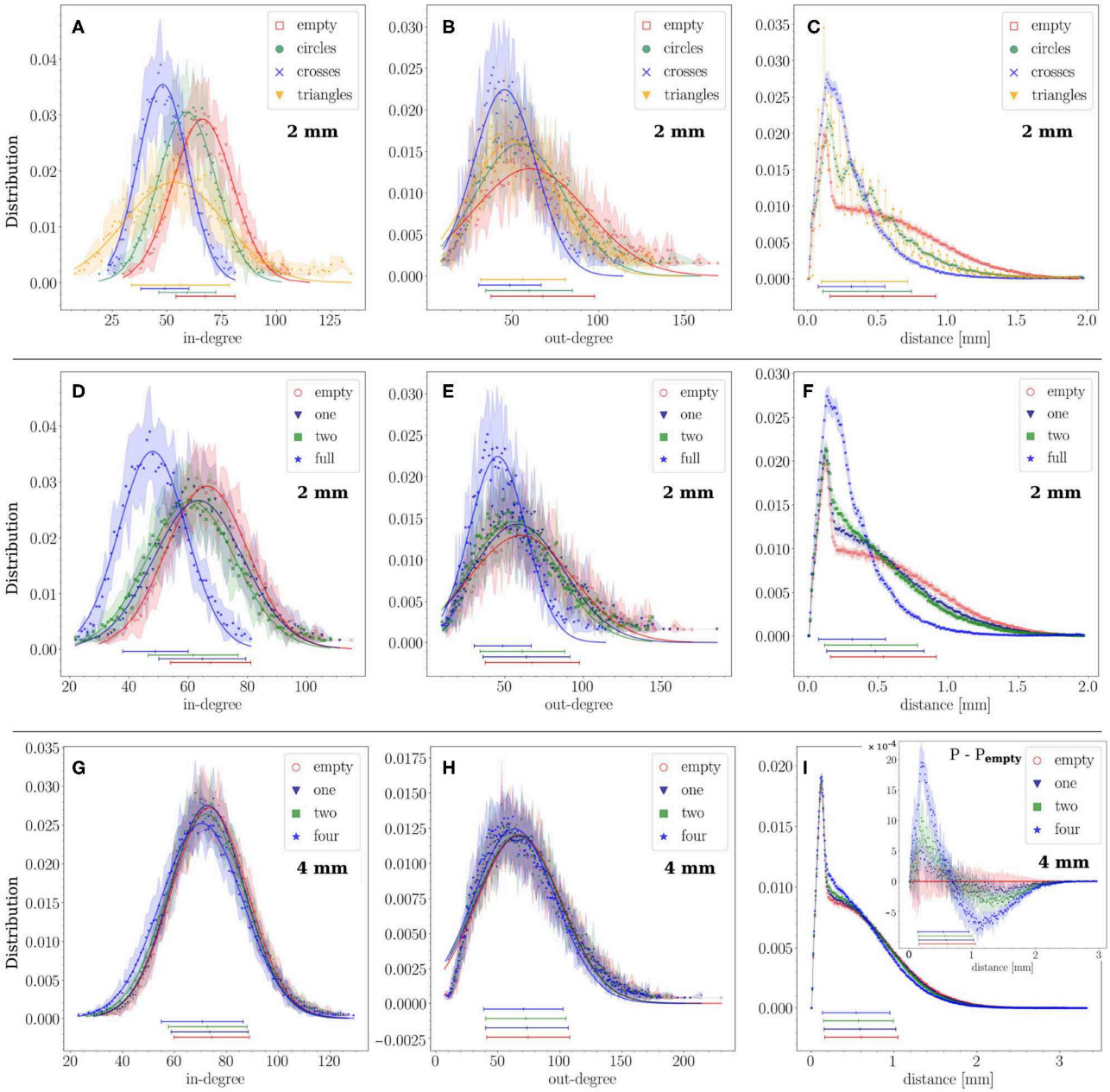

FIGURE 4 | Structural connectivity statistics. **(A–C)** Probability distributions of in- and out-degree (k_{in}, k_{out}) and connection distances d for the empty configuration and crosses, circles and triangles configurations that fully cover the substrate. Crosses and triangles exhibit the strongest effect on k_{in} and k_{out}. **(D–F)** The same distributions for empty, 1 array of crosses, 2 arrays and full coverage in the 2 mm diameter networks. The strongest effect was observed for the configuration in which the crosses fully populated the area, with k_{in} and k_{out} distributions shifting to lower values and the distribution of distances exhibiting a marked peak at small length scales. **(G–I)** The same study in a 4 mm diameter network with crosses partially covering the substrate. Empty, 1, 2, and 4 arrays of crosses are compared. The k_{in} distribution gradually shifts to lower values, and the short distances gain prominence, as the number of arrays increases. The inset of panel **(I)** shows the difference between each distance distribution P and that of the empty configuration P_{empty}. For all distributions and configurations, data is averaged over 12 network replicates. For the distributions of k_{in} and k_{out} lines show a Gaussian fit to the data, their parameters are given in **Table 1**.

distributions gradually shifted toward lower values as the density of occupation increased, although the change was substantial only for full coverage, with the average values of k_{in} and k_{out} decreasing by 30%. The distribution of connection distances d (**Figure 4F**) also experienced a strong change for full coverage, with short-range connections dominating the distribution at the expense of highly depleted mid- and long-range ones. These results confirm the hypothesis that crosses either trap axons in

a neighborhood or deflect them away, reducing the capacity of neuron to interconnect. The results also reveal that a small occupation of the substrate by obstacles only causes a minor effect in the distribution of connections. This was confirmed by investigating bigger substrates of 4 mm in diameter where the physical dimensions of the crosses were maintained, which thus occupied a very small area fraction (see **Table 1**). As can be seen in **Figures 4G–I**, the distributions for the empty, 1, 2 and 4 arrays are very similar to each other and fall within the fluctuations among replicates. Here, the effect of an increase in the number of obstacles is most noticeable on the distance distributions (**Figure 4I**). As the effect is much smaller than in the previous configurations, we computed the difference between each distribution and that of the empty configuration shown in the inset. It confirms the trend of excess short-range (< 0.7 mm) and depleted long-range connections with increased number of obstacles, as seen in the 2 mm configurations. However, these effects are much smaller in the 4 mm case due to the small area fraction occupied by the scaffolds.

An interesting trait of the distribution of distances is the presence of a plateau for the empty case (**Figure 4** and **Figure S1**). This plateau is associated with the broad range of possible axonal lengths, and whose average length ($\ell_a = 1.1$ mm) is an order of magnitude larger than the average radius of the dendritic tree (150 μm), effectively shaping a neighborhood around the neurons in which connection probability is independent of the distance. The presence of obstacles alters this plateau, particularly when they fully cover the substrate, since axons cannot extend freely for long distances.

We next explored the effect of substrate size on structural connectivity. We observed that alterations in k_{in} and k_{out} degree distributions—relative to the empty configuration—were more prominent when the substrate radius was similar to the characteristic axonal length, approximately 1 mm in our case. This is illustrated in **Figure 5**, where we compare the degree distributions among 3 networks grown on substrates whose diameters were scaled up from 2 to 12 mm (see **Table 2**). We considered the empty configuration and the crosses configurations with either 1 array or full coverage. The dimensions of the crosses were also scaled up according to the substrate diameter to preserve the area fraction occupied by obstacles at 2.4% for 1 array and 13.6% for the covered configuration. As shown in panels A and B of the figure, distributions corresponding to the 1-array configuration at different scaling factors were very similar among themselves except for the smallest size of 2 mm diameter, which markedly shifted to low k_{in} and k_{out} values. Only for this diameter differences between the empty and obstacles configurations could be appraised. However, for the configuration covered with crosses the effects were stronger as shown in **Figures 5C,D**. At 2 mm the mean value of k_{in} decreases by 26.5% and k_{out} by 24.1% in obstacles compared to empty configurations. At 4 and 6 mm the distributions of both in- and out-degree were still clearly shifted and narrower for the covered configuration. The effect was less clear but still perceptible at 8 and 12 mm scales. We thus conclude that size effects are very important and that they clearly attenuate effects of scaffolds at the area fractions explored here. When the system size is significantly larger than the characteristic axonal length, then metric correlations at short length scales mask out the alterations induced by the obstacles. Nonetheless, obstacles still have an impact on their neighborhood, but from a global perspective the network may appear unaffected.

To complete the analysis of connectivity, we studied the spatial variability in the degree distributions and in clustering coefficients (CCs) in the 2 mm substrate. For sake of simplicity, we considered only k_{in} in this analysis since it is the distribution that exhibits the strongest differences among configurations. We represented average values of k_{in} and CC in square regions of side 0.031 mm, containing each about 0.2 neurons. As shown in **Figures 6A,C**, the empty configuration portrayed strong inhomogeneities in k_{in} which originated from metric correlations. The addition of obstacles in the form of 1 array or 2 arrays of crosses reduced the in-degree values within the scaffolds and lead to higher values in localized areas outside the scaffolds, hence accentuating inhomogeneities in the network. We note that the k_{in} distributions shown in **Figures 4A,B** could not capture these inhomogeneities. Thus, this spatial analysis helps to highlight important fluctuations in the network that cannot be appreciated from solely inspecting the shape of the degree distributions averaged over replicates. The CC values, however, did not show a clear trend in spatial distribution upon the inclusion of scaffolds, although the maximum CCs increased by 20% and tended to concentrated around the scaffolds area, possibly as a consequence of the deflected axons and that facilitated the formation of a higher number of triangles.

The corresponding spatial analysis for obstacles fully covering the network is shown in **Figures 6D–F**. For crosses, the maximum k_{in} dropped by 15% relative to the empty configuration and the maximum values appeared concentrated in mostly a few adjacent cells, while patches of low in-degree were more evenly spread across the network. This behavior contrasts with the circles configuration, in which fluctuations among neighboring regions are much weaker, although very high k_{in} values occur near the border. For triangles, a strong gradient of k_{in} values emerged that extended across the entire network, with k_{in} decreasing sixfold in the direction of the tips of the triangles. This patterned distribution of k_{in} values highlights the strong guidance of the axons, which also favored an increase of the maximum k_{in} values by 15% compared to the empty configuration. The highest values were localized at the lower edge of the triangles pattern. The CC values for these configurations showed an overall increase of the maximum values by 30% for crosses and triangles, but only increased by 10% for circles. Spatial fluctuations in CCs were marked for crosses and milder for circles and triangles.

2.2. Dynamic and Effective Connectivity Alterations Induced by Obstacles

We simulated dynamics of excitatory cortical neurons in the generated structural networks through an integrate and fire model with adaptation, whose parameters were adjusted as in Orlandi et al. (2013) to provide rich spontaneous activity for the empty configuration. Activity was simulated for 30 min

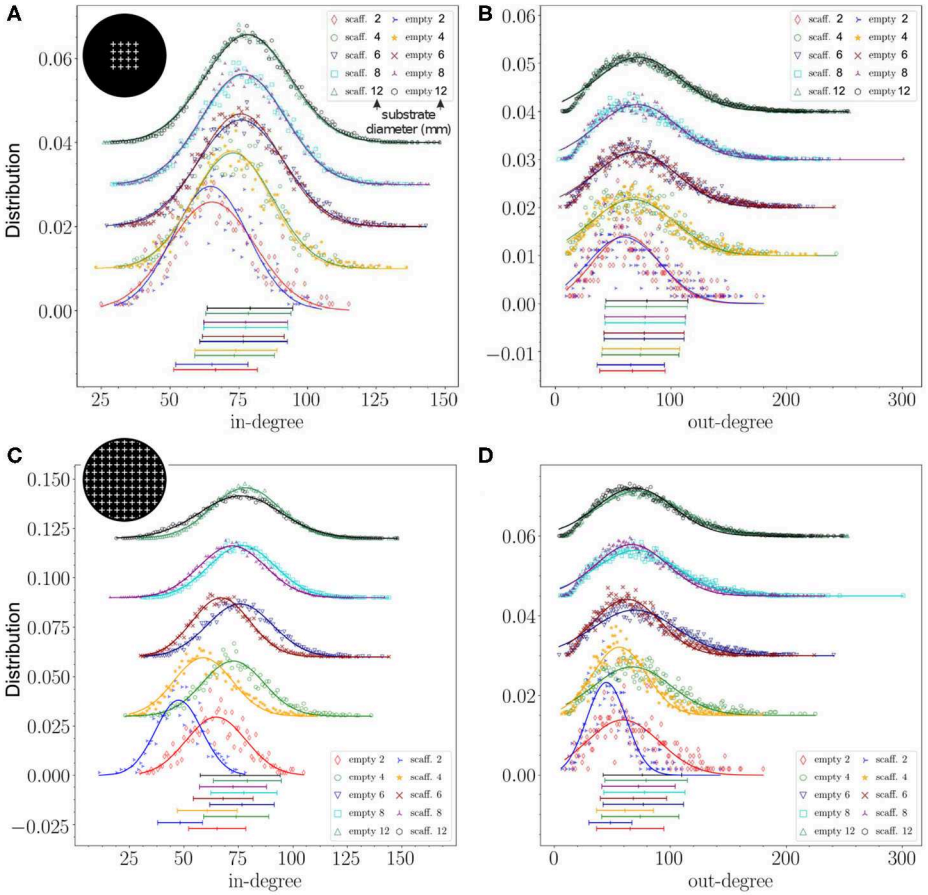

FIGURE 5 | Impact of network size scaling on degree distributions. **(A)** Distributions of in-degree for the empty configuration and for the configuration with 1 array of crosses scaled up to different diameters, preserving the relative size of the scaffolds with respect to substrate diameter. **(B)** Corresponding out-degree distributions. **(C,D)** Distributions of in- and out-degree of the configuration covered by crosses scaled up to different diameters, compared with the empty configuration. For all distributions, effects were more marked for smaller sizes. The data correspond to one representative replicate for each configuration. The lines show a Gaussian fit to the data, their parameters are given in **Table 2**.

for four replicates of each configuration. Then, we explored the changes in collective activity and effective connectivity due to the presence of obstacles in 2 mm diameter cultures which were the ones displaying the strongest effects in the above analyses of structural connectivity. We must note that spontaneous activity comprises both sporadic neuronal activations and network-wide coordinated episodes in the form of network bursts. An abundance of sporadic activations may mask the statistics of network activity and induce artifacts in the analysis of effective connectivity. Thus, in the analysis that follows we filtered out sporadic activity data to emphasize network bursting events, and retained only coordinated activations that encompassed at least 25% of the network.

We first considered the situation in which cross-shaped obstacles progressively populated a larger fraction of the substrate's area. As shown in **Figure 7A**, network bursting was high for the empty and 1 array configurations, and progressively diminished as the density of obstacles grew. Collective activity almost halted in the configuration in which the obstacles fully populated the area, suggesting that the substantially reduced structural in- and out-degree values strongly affected the capacity of the network to trigger activity and initiate bursts.

The corresponding analysis of the effective connectivity is shown in **Figures 7B,C**, which provide the adjacency matrices obtained through transfer entropy together with the network maps of community organization and effective out-degree distributions. In the maps, the size of a node is proportional to its out-degree. We chose to plot the out-degree since it reveals the initiation of activity, i.e., which neurons in the network tended to activate other neurons. The adjacency matrix for the empty configuration shows modular traits ($Q \simeq 0.28$) and reveals that some groups of neurons tended to coactivate more frequently with each other than with the rest of the network. The effective modules, however, did not shape compact areas in the network maps but were highly intermixed. This reveals that, despite modularity, network intercommunication was strong as indicated by the high global efficiency ($G_{\text{eff}} \simeq 0.41$). Activity

TABLE 2 | Scaled configurations without obstacles (empty), with 1 array of crosses and covered with crosses (full).

	diameter (mm)	$N_{neur.}$	k_{in} μ	k_{in} σ	k_{out} μ	k_{out} σ
Empty	2	625	64.77	13.67	59.13	31.50
Empty	4	2,500	72.57	14.02	67.32	33.44
Empty	6	5,625	75.68	14.92	69.05	36.11
Empty	8	10,000	76.69	15.18	70.78	35.49
Empty	12	22,500	78.24	15.57	72.23	36.01
1 array	2	625	65.26	15.60	61.23	29.07
1 array	4	2,500	72.92	14.49	66.98	35.81
1 array	6	5,625	75.87	15.69	69.44	34.78
1 array	8	10,000	77.06	15.25	70.73	35.65
1 array	12	22,500	77.70	15.48	71.23	35.82
Full	2	625	47.60	10.55	44.89	17.13
Full	4	2,500	58.78	13.35	56.08	23.17
Full	6	5,625	67.01	13.25	66.63	31.13
Full	8	10000	72.19	15.14	62.51	28.71
Full	12	22,500	75.88	18.68	69.26	33.17

We report the diameter of the configuration, the number of neurons ($N_{neur.}$), as well as the average value (μ) and standard deviation (σ) obtained for the Gaussian fits to the distributions of in- and out-degree (k_{in}, k_{out}).

also initiated in a similar manner throughout the culture, with the highest values of k_{out} spread out homogeneously.

A similar overall trend was observed for the configuration with 1 array of crosses, which yielded very similar values of Q and G_{eff}. However, the effective modules were more compact and no high out-degree values were observed in the center of the map, where the array is placed, indicating that activity did not initiate within the array. For the 2-array configuration, modularity increased by 25% relative to the empty case, which was accompanied by an increase in the number of modules. This is a sign of higher fragmentation of the dynamics. One of the modules was also compact when represented in the network map (pink-colored neurons), indicating that the obstacles weakened the capacity for whole-network interaction of activity. Most of the activity initiated in this module at the bottom of the map or in small regions at the top, and weak activity was detected within the arrays. These results indicate that the obstacles were capable of shaping effective microcircuits, i.e., a neighborhood of highly activate neurons that poorly interacted with the rest of the network. The isolation of these effective microcircuits strengthen for the configuration in which the crosses fully covered the area (**Figures 7B,C**, right panels). Here we observed a substantial increase in modularity by about 80% relative to the empty case, with some modules at the verge of full dynamic isolation, as recognized in the effective connectivity matrices by the few links outside the diagonal. G_{eff} practically fell to zero, indicating the severely reduced capacity of the network to exchange information. This appears in the map as a large number of disconnected neurons. Activity tended to start at the right edge of the culture (high density of out-degree values), possibly facilitated by the border of the substrate.

To complete the analysis of activity, we also looked at the spatiotemporal structure of network bursts. As shown in **Figure 7D**, bursting events propagated as circular or quasiflat fronts for the empty and 1 array configurations, reflecting a reduced sensitivity to connectivity inhomogeneities in the network. This neat propagation pattern was altered in the 2-array and full configurations, with propagation showing a richer structure that evinced the strong spatial fluctuations in connectivity.

We note that the dynamics in the 2-array and full coverage configurations were very sensitive to the details of the network replicate. We observed that in some instances the simulated networks were incapable of generating network bursts. We characterized this effect on network bursting by computing the spatial distribution of burst initiation events (**Figure 8**). For the 2 mm diameter network, burst initiation was distributed over most of the area in the empty configuration, but it became increasingly localized as more obstacles were incorporated. Initiation took place outside the scaffolds except for the full coverage configuration, for which the bursting fronts were so fragmented that the identification of initiation could not accurately be determined and most likely occurred near the edges of the network. For comparison, we also provide the results for 4 mm diameter networks. In those simulations the initiation was much richer and the impact of the obstacles was smaller. However, initiation never occurred within the arrays and became more localized as more arrays were added.

The equivalent effective connectivity analysis for the different types of obstacles fully covering the substrate is shown in **Figure 9**. The raster plots compare the characteristic dynamics across configurations. Although all of them displayed decreased activity due to the obstacles, bursting was least affected in circles, mildly in triangles and strongly in crosses, as discussed above. In all cases, however, the effective connectivity matrices (**Figure 9B**) showed a trend toward high Q values relative to

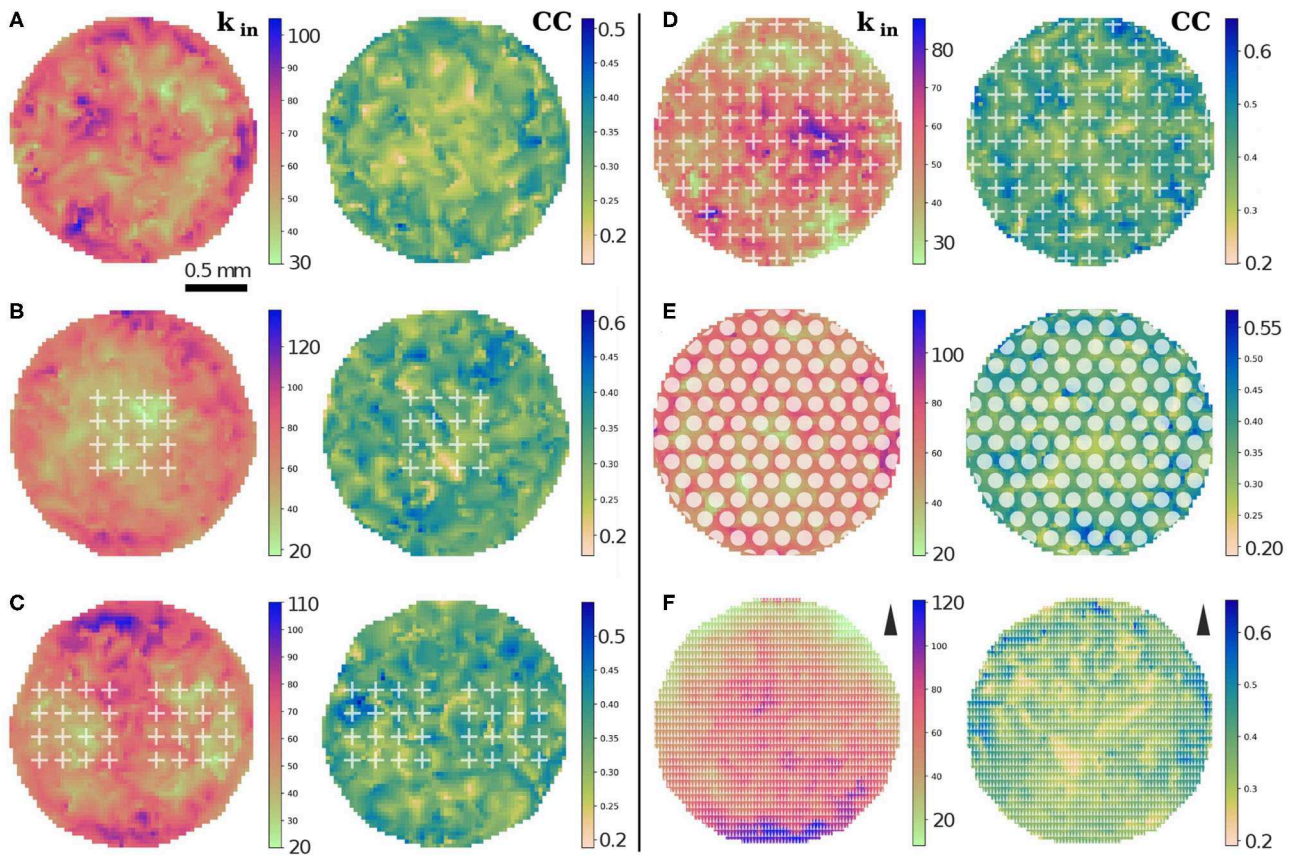

FIGURE 6 | Spatial distributions of in-degree k_{in} and clustering coefficient CC for the different configurations in the 2 mm substrate. **(A–C)** Distributions for the empty configuration, 1 array and 2 arrays of crosses. The value of k_{in} decreases by 30% inside the arrays of crosses, while CC is not appreciably affected. **(D–F)** Corresponding distributions for crosses, circles and triangles entirely covering the substrate. Strong effects are observed in the distribution of k_{in} for crosses and triangles. Compared to the empty configuration, crosses show a decrease of the maximum k_{in} by 15%; triangles show a gradient of k_{in} values, which increase in the direction opposite to the tips of the triangles which is indicated next to the colorbar **(F)**. The data correspond to one representative replicate for each configuration.

the empty configuration which was reflected in an abundance of small sized modules. Circles and triangles, as compared to crosses, exhibited well interlinked modules, with few silent neurons, and therefore their G_{eff} values were not as small as in the crosses configuration. The network maps (**Figure 9C**) illustrate the strong cohesion of the effective networks for circles and triangles, with modules extending all across the area. Effective out-degree values were well spread for circles, indicating that activity initiation equally occurred everywhere. For triangles there was a clear localization of out-degree values toward the bottom of the map, the region that contains also the highest structural k_{in} values. This correlation between structural and dynamical traits highlights that adequate configurations of obstacles help dictating activity initiation. The structure of spatiotemporal fronts (**Figure 9D**) shows that all configurations developed structured activity propagation patterns. We point out that the velocity of propagation varied among configurations. Propagating fronts crossed the network in about 30 ms for the empty and circles configurations, while this time increased to 60 ms for crosses and to 300 ms for triangles. The slow propagation observed in triangles is due to the strong connectivity differences between the direction parallel to the triangles' orientation (with high connectivity) and the direction orthogonal to it (weak connectivity), causing the front to advance faster in one direction but slower in the other.

To conclude our study, we compare the major dynamic and network characteristics—structural and effective—among configurations. **Figure 9A** provides a comparison of the distributions of inter-burst intervals (IBIs), showing the contrasting differences between crosses and the rest of configurations. **Figures 9B,C** provide the comparison of Q and G_{eff}, respectively. The main plots summarizes the data for the 2 mm diameter networks, while the insets provide the data for the 4 mm ones. All data is organized so that the magnitudes of a given property increase toward the right. For the 2 mm data, the structural network properties varied gently and with very small fluctuations. This contrasts with the effective network properties that exhibited strong changes among configurations and with substantial variability among replicates. For the 4 mm data, all network measures

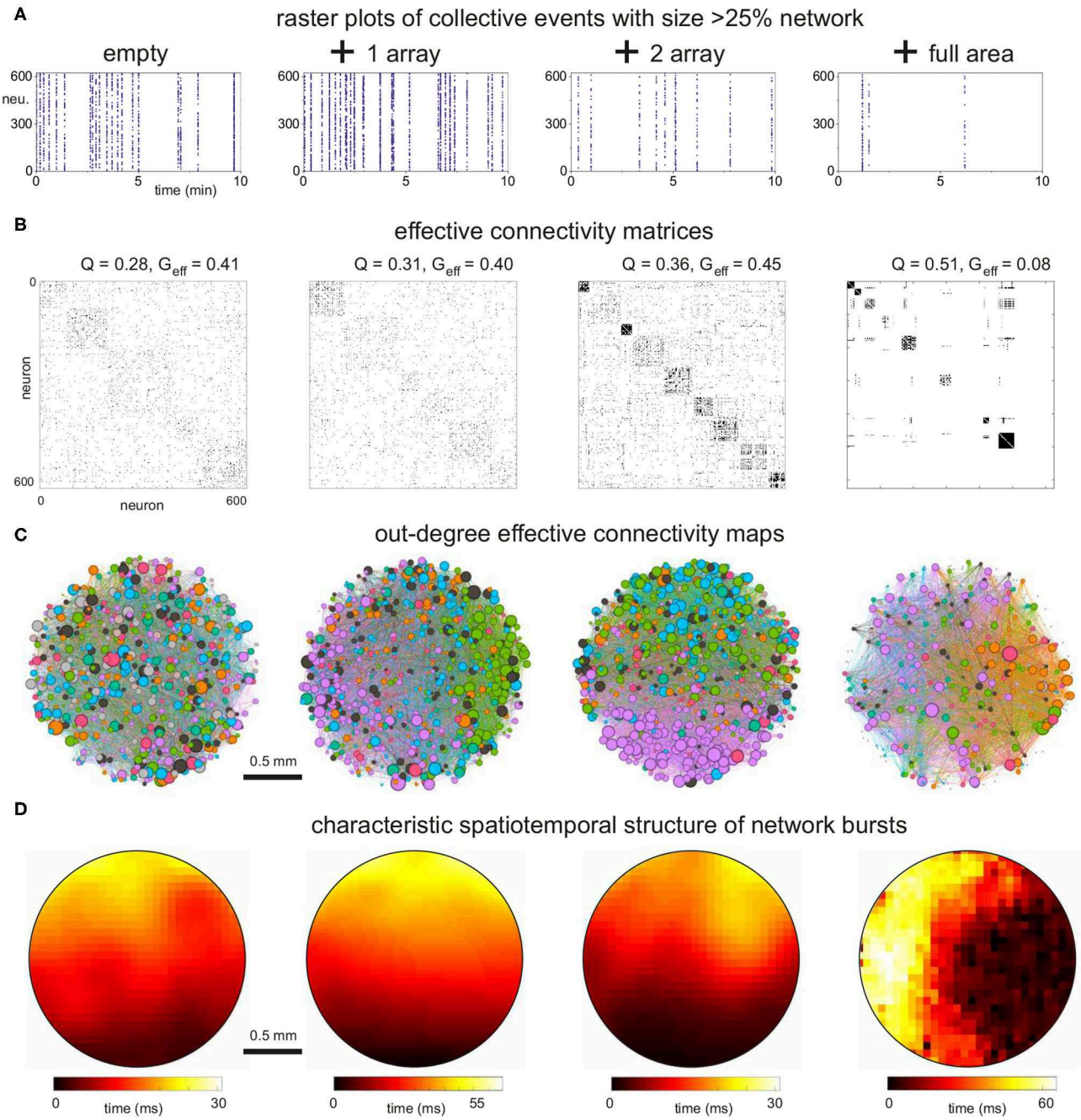

FIGURE 7 | Dynamics and effective connectivity in 2 mm diameter networks with increasing density of cross-shaped obstacles. **(A)** Representative raster plots of bursting events for the empty, 1 and 2 arrays of crosses and full coverage. Bursting activity decreases as the density of obstacles increases. **(B)** Corresponding effective adjacency matrices, with modules along the diagonal. Modularity increases as the density of crosses grows, with a boost by 80% for the full coverage relative to the empty configuration. **(C)** Spatial maps of the effective networks. Neurons are color-coded according to their modularity class, and their diameter is proportional to the out-degree of effective connectivity which reflects activity initiation. Edges are colored according to the outgoing module. The network fully populated with crosses exhibits strong spatial anisotropies. **(D)** Spatiotemporal patterns of representative network bursts. The activity fronts advance as a quasicircular front for the empty and 1 array configurations, to become more structured and erratic for higher densities of crosses. The data correspond to one representative replicate for each configuration.

varied gently, either structural or effective, which again highlights the importance of fully covering the substrate with obstacles to induce substantial changes in both structure and dynamics.

3. DISCUSSION

Our results show that obstacles imprint features on the structural connectivity that may lead to strong alterations in the collective

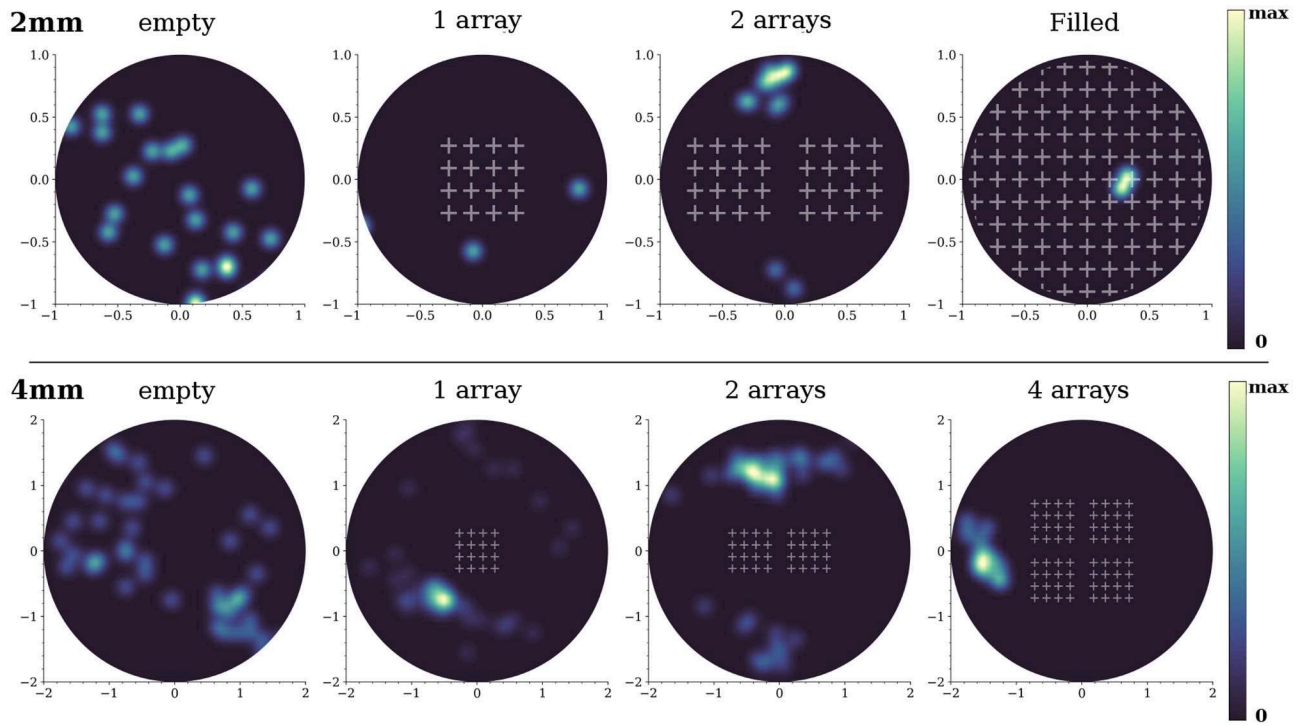

FIGURE 8 | Initiation points for the 2 and 4 mm diameter networks. The blue-yellow patterns are the locations where network bursts commenced. The brighter the color, the higher the occurrence of a burst initiation event in that area. The data shown correspond to one representative replicate of each configuration.

dynamics and effective connectivity of neuronal networks. With those designs of obstacles that we explored, the molding of structural connectivity can occur in two different ways. The first one is by funneling axons in a given direction, as observed with triangles, and the second one is by modifying the spatial density of incoming or outgoing connections in a given region, as observed with crosses. In either case, the capacity of the network to recruit, amplify and propagate activity is affected, thus causing alterations in the timing and spatiotemporal structure of network bursts whose details are sculpted by the underlying structure. When the obstacles fully populated the substrate, their shape was much more important than the total area they occupied. Circles and triangles configurations, both occupying an area fraction of about 45%, caused a twofold increase of the inter-burst interval (**Figure 10A**), while for crosses the increase was sixfold even though they occupied just 14% of the available area. The capacity of crosses to either trap or deflect axons emerged as a key property as compared to the funneling of axons by triangles or the gentle alteration of axonal paths by circles. In addition to shape, the ratio of typical axonal length to substrate diameter was also a key parameter. When obstacles occupied only a small region of the substrate, as the 1 or 2 arrays of crosses for instance, they induced local alterations whose global effects were masked by the connectivity traits of the rest of the network (**Figure 5**).

Our simulations reflect the importance of metric correlations in shaping connectivity and dynamics in neuronal circuits. Metric correlations appear naturally in spatially embedded networks (Orlandi et al., 2013; Tibau et al., 2020). As in our simulations, other studies pointed out the spatial distribution of neurons and the characteristic axonal length relative to system size as central ingredients in shaping local and global structural traits (Schmeltzer et al., 2014; Hernández-Navarro et al., 2017; Okujeni et al., 2017). The importance of metric correlations is that they facilitate spatial heterogeneities in the connectivity of the network which greatly influence the dynamic behavior of the entire system, in particular its capacity to initiate and propagate coherent activity in the form of network bursts (Orlandi et al., 2013; Okujeni et al., 2017; Faci-Lázaro et al., 2019). Our work goes a step further and shows that obstacles affect connectivity by changing the shape and average values of in- and out-degree distributions and by altering the range of connection distances, which promoted variations that could be locally very strong. The crosses and triangles configurations were the ones that more significantly altered the spatial structure of connectivity. The in-degree values dropped substantially within areas populated with crosses, while triangles induced a strong gradient of in-degrees along their orientation.

The mechanisms that caused a reduction of the spontaneous activity when obstacles were incorporated are complex. The detailed studies of burst initiation mechanisms by Orlandi et al. (2013) showed that a balance of different network observables was required to maximize bursting, which included in- and out-degrees, clustering coefficients, feed-forward loops and feed-backward loops, among others. Additionally, the study of Orlandi

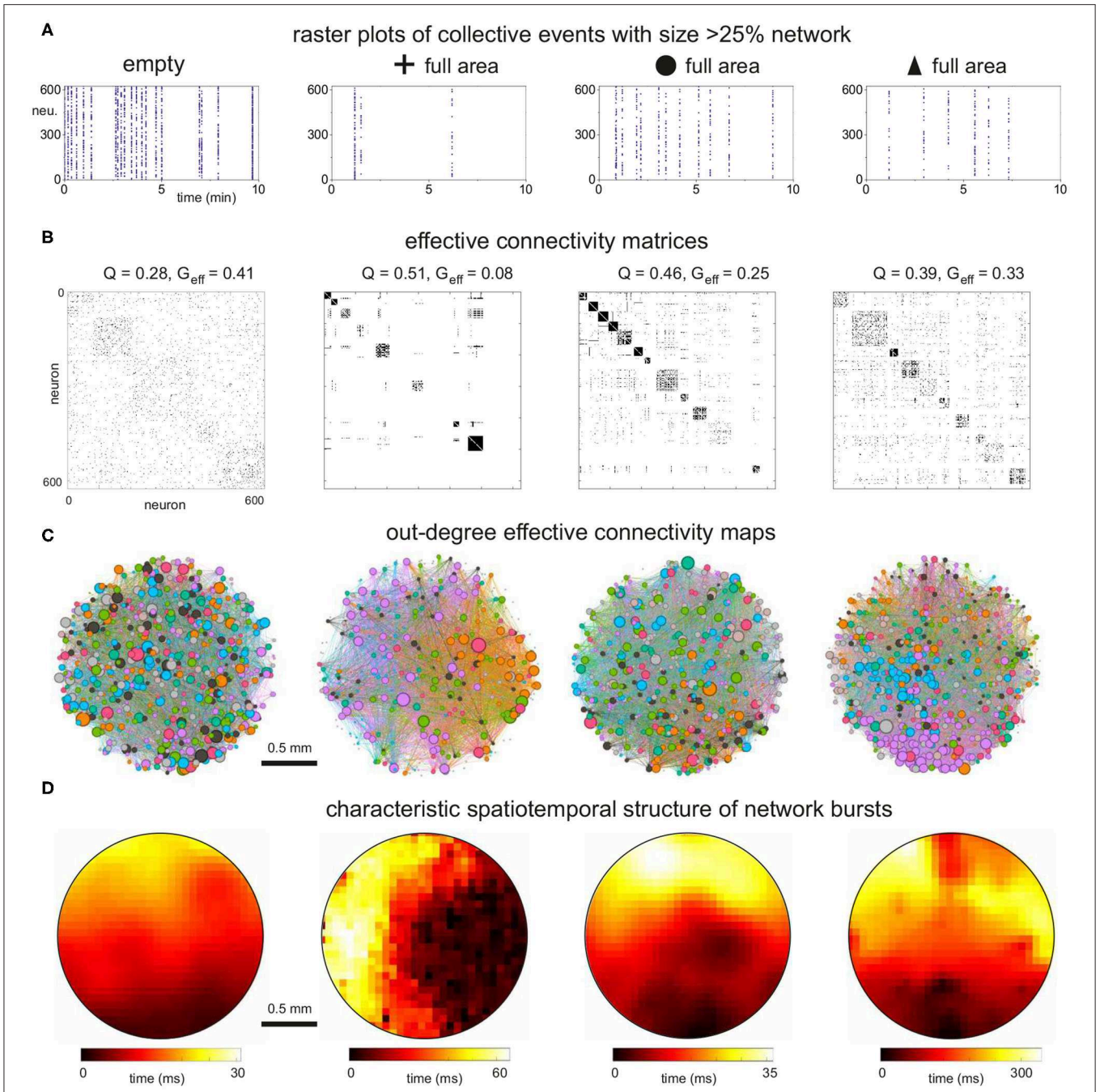

FIGURE 9 | Dynamics and effective connectivity in 2 mm diameter networks filled with obstacles. **(A)** Representative raster plots of bursting events for the empty configuration and full coverage by crosses, circles and triangles. **(B)** Corresponding effective adjacency matrices, with modules highlighted along the diagonal. The modularity Q depends on the specific obstacle design, but the number of communities is higher in all cases as compared to the empty configuration. **(C)** Spatial maps of the effective networks. Neurons are color-coded according to their modularity class, and their diameter is proportional to the out-degree of the effective connectivity which reflects activity initiation. Edges are colored according to the outgoing module. While the configuration of circles exhibits traits similar to the empty one, the triangles show a tendency for activity to initiate at the bottom of the network, where k_{in} is higher. **(D)** Spatiotemporal patterns of representative network bursts. Activity flow is structured for the networks with obstacles. The speed of propagation is a factor 2 and 10 lower for crosses and triangles, respectively, relative to the empty configuration. The data shown correspond to one representative replicate of each configuration.

and coworkers pointed out that an excess or deficit of some of these observables could substantially reduce bursting frequency. Our observation that the in- and out-degree distributions are substantially shifted to lower values suggests that they could be major actors in the alteration of activity. This is supported by a recent study of Faci-Lázaro et al. (2019), in which they observed

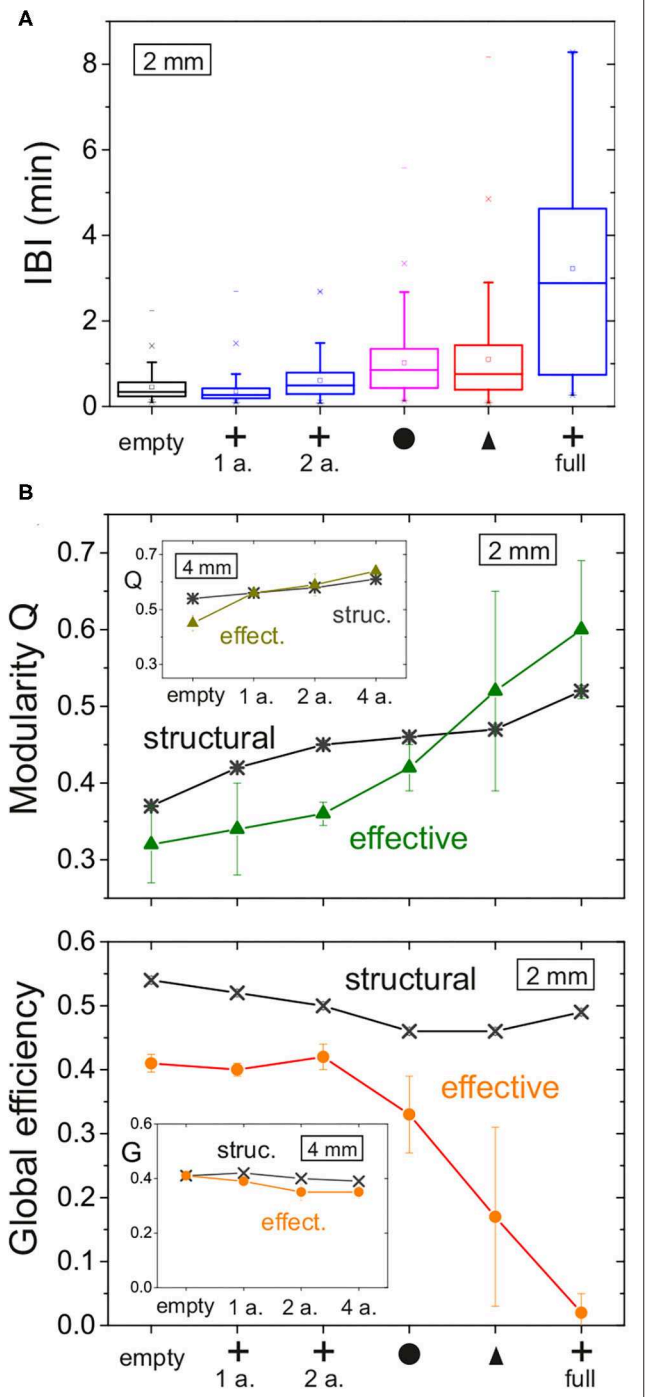

FIGURE 10 | Bursting and effective connectivity statistics for 2 mm networks with obstacles. **(A)** Box plots of the distribution of inter-burst intervals (IBIs) for all the explored configurations. For each box: the inner square is the mean; the central horizontal line the median; the top and bottom box edges are, respectively, the 25th and 75th percentiles; the bottom and top crosses are, respectively, the 1st and 99th percentiles; and the bottom and top dashes are the data range. The IBI in general increases as the density of obstacles grows. Among obstacles fully covering the network, crosses show the strongest alteration on the timing of activity. **(B)** Modularity Q and global efficiency G_{eff} for *(Continued)*

FIGURE 10 | the different configurations of obstacles, comparing structural network traits with effective ones. Main plots correspond to networks of 2 mm diameter, and insets to those of 4 mm diameter. Each data point is an average over four replicates, and error bars denote standard deviation.

in simulations of neuronal networks similar to ours that the loss of nodes with the highest out-degree precipitated a substantial drop in the number of bursting episodes. The important shift of the out-degree distribution toward lower values for cross-shaped obstacles in **Figure 4**, much stronger than for other types of obstacles, suggests that out-degree decrease could be one of the most important factors in activity reduction.

We observed that the structural network traits of the studied networks were very similar across network replicates. Even for the configurations in which the obstacles fully populated the substrate, the distributions of k_{in} and k_{out} and the values of Q and G_{eff} varied less than 5% among replicates of the same obstacle design (**Figure 10**, structural data). However, the effective traits substantially changed as evinced by the large dispersion of both Q and G_{eff} (**Figure 10**, effective data). Since the effective connectivity reflects dynamics, the strong contrast between these two network descriptions clearly shows the complex relationship between structure and dynamics, and that the former cannot be directly inferred from the latter with current methods. For instance, the network maps of the configurations with obstacles in **Figure 9** are qualitatively similar to one another, but by analyzing only them or the corresponding effective matrices we cannot deduce precisely which structural connectivity or obstacle configuration they emerged from. Thus, our work invites to proceed with caution when trying to infer structural connectivity features from effective ones.

The simulations showed that obstacles increased the modularity in the network, with an impact on both structure and dynamics (effective connectivity). We observed that the impact on structure was similar for all types of obstacles at full coverage of the substrate, with an increase of Q by about 30% with respect to the empty case. However, the impact on effective connectivity was much higher, with Q increasing by 50% for circles and 80% for crosses. This suggests that the sharp edges of the crosses configuration greatly facilitate the isolation of groups of neurons, a characteristic that is especially relevant for experimental, *in vitro* preparations aimed at enriching the dynamic and functional organization of neuronal networks. For instance, crosses could be placed in groups of four and closer to one another, shaping a structure similar to a hollow square with tiny entrances. Such a structure would create communities of strongly connected neurons with weak connectivity among communities, mimicking for instance the designs of Yamamoto et al. (2018).

Configurations of tailored obstacles could also help shaping networks-of-networks such as the experimentally observed aggregated neuronal networks (Sorkin et al., 2006; Teller et al., 2014) or fractal designs (Díaz Lantada et al., 2013). The latter can be employed to capture the non-Euclidean geometry of the human brain and its relation with developmental traits and

multi-scale dynamics (Werner, 2010; Hofman, 2014). Fractality and multi-scale organization are indeed inherent properties of cortical circuits and are closely related to the concept of criticality (Poil et al., 2008; Friedman et al., 2012; Haimovici et al., 2013; Massobrio et al., 2015; Marshall et al., 2016; Johnson et al., 2019), in which a neuronal circuit operates at the boundary between an ordered, strongly coupled state and a disordered, weakly coupled one. Neuronal systems at criticality exhibit long-range spatial and temporal correlations with power-law distributed statistics, facilitating a broad dynamic repertoire and swift communication among distant areas. Massobrio et al. (2015) showed through experiments and theoretical analysis that a critical state can be favored by combining short- and long-range connections, i.e., by imprinting small-world features into a neuronal circuit. Here, we observed that the presence of obstacles, particularly crosses, increased the "small-worldness" (Watts and Strogatz, 1998; Humphries and Gurney, 2008) from 2.46 to 4.03 (SM, **Table S1**). This confirms that the obstacles trap axons and increase connectivity locally while preserving some long-range connectivity.

Although our simulations aimed at providing a numerical playground to investigate the impact of physical constraints on structural connectivity and dynamics, they were limited by a number of simplifications that could be relaxed in future studies. A first simplification concerns the rule for the growth of axons. We disregarded for simplicity the interaction of axons with neurons or with other axons, and used a simple "reflection" rule to model the interaction between axons and obstacles. *In vitro* experiments in engineered neuronal cultures (Feinerman et al., 2008; Li et al., 2014; Casanova et al., 2018) and microfluidic chambers (Renault et al., 2016; Yamada et al., 2016; Holloway et al., 2019) have shown that axons interact in complex ways with obstacles and that axons often attach to and follow walls. Thus, for a more realistic representation of *in vitro* behavior those interactions should be incorporated in future simulations. A second simplification was the use of excitatory neurons only, which facilitated the inference and analysis of connectivity and its relation with overall network dynamics. The inclusion of inhibition, which typically comprises of about 20% of connections in cortical circuits (Soriano et al., 2008; Schröter et al., 2017), would reduce whole-network bursting and promote a richer spatio-temporal dynamics, as observed experimentally in two-dimensional homogeneous and engineered neuronal cultures (Cohen et al., 2008; Orlandi et al., 2013; Okujeni et al., 2017; Yamamoto et al., 2018). A third simplification was the use of soma and synapse dynamical models that shape cortical-only neuronal networks without plasticity. The inclusion of different cell types and activity-regulatory mechanisms could help investigating questions such as the capacity of the networks to reach activity set points or their response to neuronal loss, as recently explored experimentally *in vitro* (Slomowitz et al., 2015; Teller et al., 2019). And a fourth simplification was the use of solely two-dimensional networks, which only partially reflect the structural complexity and functional richness of naturally-formed brain circuits. Severino et al. (2016) recently showed experimentally and numerically that three-dimensional neuronal networks with fractal organization maintain modular characteristics while promoting long-range connections. As discussed above this facilitates the emergence of a small-world architecture and enhances whole-network bursting. Thus, fractal or three-dimensional patterns could be employed to design more realistic simulations aiming to mimic the dynamic behavior of *in vivo* circuits.

4. CONCLUSION

We have shown that it is possible to dictate the structure of neuronal circuits by incorporating obstacles, whose impact on dynamics and effective connectivity depends on their shape and density. Our work invites the exploration of various configurations in an effort to control the dynamics of the resulting networks. However, achieving precise control remains difficult due to the complex interplay between connectivity, intrinsic neuronal dynamics and noise. Nonetheless, our study provides a method and tools that will allow computational neuroscientists not only to explore a variety of configurations systematically, but eventually contribute to the understanding of the way in which geometry influences the emergence of patterns in growing networks of living neuronal circuits. Thereby, our study can assist in the design of substrates to guide the growth of networks *in vitro*, inviting a quicker and more efficient investigation of prototype geometries than in wet-lab experiments. This will help in finding and selecting suitable candidate geometries for scaffolds or complex architectures in brain-on-a-chip investigations.

5. METHODS

5.1. Geometric Patterns

Three types of obstacles were studied: crosses, circles and triangles. Arrays of obstacles were placed in circular areas of either 2 or 4 mm diameter. Patterns were set as white objects on a black substrate (**Figure 2**), and simulated neuronal soma and axons were only allowed to grow on the black areas. The neuronal density was set to 200 neurons/mm^2, leading to networks with 625 and 2500 neurons for the 2 and 4 mm diameter sizes, respectively. An empty configuration with the same number of neurons was also considered as reference (**Figure 2A**). The different obstacles' geometries are described in detail below.

Crosses: The cross-shaped obstacles were 130 μm high and wide, with a beam thickness of 20 μm. The spacing between crosses was 50 μm. They were arranged either in arrays of 4×4 crosses, each array covering a square area of side 670 μm, or filling the available substrate entirely (**Figures 2B–D**). For the latter, a ring 50 μm wide at the edge of the substrate, and free of obstacles, was incorporated to ensure that border effects were the same everywhere in the network. Arrays were placed at the center of the circular substrate. For the 2 mm diameter networks, simulation schemes considered 1 array, 2 arrays, and full occupation; for the 4 mm, simulations considered 1, 2, and 4 arrays. The spacing between arrays was 230 μm. The empty and 1 array configurations were also simulated in a version scaled up by factors 2, 3, 4 and 6. In these scaled versions, the dimensions of

the crosses changed according to the scaling factor. The number of neurons placed within the area was scaled to conserve the neuronal density of the smallest configuration (see **Table 2**).

Circles: This design consisted in circles of 120 μm in diameter that were placed in a hexagonal grid covering the entire substrate as shown in **Figure 2E**. The separation between circles was 50 μm. A 50 μm spacing at the edge of the substrate was incorporated as for the crosses.

Triangles: The triangle-shaped obstacles were designed to mimic the geometry of experimental scaffold structures (Crowe et al., 2020). Triangles were of isosceles shape with 50 μm height and 20 μm width. They were placed pointing upwards (**Figure 2F**). Triangles were arranged in an array that entirely filled the substrate excepted at the edge, that incorporated a ring 5 μm wide free of triangles. The horizontal and vertical separation between triangles at their base was 5 μm.

5.2. Network Generation

Neurons were randomly positioned without overlap in the black areas of the designed patterns. Neuronal soma were virtual objects that did not occupy physical space. Thus the axons interacted only with the obstacles and not with the neurons. Neuronal dendritic trees and axons were incorporated following (Orlandi et al., 2013). Briefly, dendritic trees were modeled as circular areas with radius drawn from a normal distribution (mean $\mu = 150$ μm and standard deviation $\sigma = 20$ μm), while axons grew at random angles from the neurons' center and followed a biased random walk of concatenated segments of length ℓ (**Figure 1**), with a total length drawn from a Rayleigh distribution with width $\sigma = 0.9$ mm and average axonal length $\ell_a = 1.1$ mm. Upon encountering an obstacle's edge an axon was reflected on the opposite side of the normal to the reflecting surface with a symmetric angle. Once the axons were positioned on the substrate, a connection was established whenever the axon of a given neuron intersected the dendritic tree of any other neuron. The whole network connectivity that resulted from this geometric construction was stored in the *structural* adjacency matrix $\mathbf{S} = \{\mathbf{s_{ij}}\}$, where $s_{ji} = 1$ corresponds to a connection $i \to j$ and $s_{ji} = 0$ otherwise.

5.3. Neuron and Synapse Dynamics

A quadratic integrate and fire model with adaptation, based on Izhikevich (Izhikevich, 2003, 2007; Alvarez-Lacalle and Moses, 2009), was used to model the soma dynamics. The equations governing a single neuron are

$$\tau_c \frac{d}{dt} v = k(v - v_r)(v - v_t) - u + I + \eta, \tag{1}$$

$$\tau_a \frac{d}{dt} u = b(v - v_r) - u, \tag{2}$$

$$\text{if} \quad v \geq v_p \quad \text{then} \quad v \leftarrow v_c,\ u \leftarrow u + d_0. \tag{3}$$

$$\frac{d}{dt} D = \frac{1}{\tau_D}(1 - D) - (1 - \beta) D \delta(t - t_m), \tag{4}$$

where the fast soma membrane potential is v, the slow inhibitory current is u, with τ_c and τ_a their respective time constants. The synaptic inputs are denoted by I, and the spontaneous emission of spikes is reflected by the noise term η. The resting membrane potential is v_r. Above the threshold potential v_t, v rises to its peak value v_p generating a spike, whereafter it is reset to v_c. The membrane potential u is reset with the parameter d_0 which describes high threshold conductances. Synaptic depression in Equation (4) is modeled as in Alvarez-Lacalle and Moses (2009), with the characteristic recovery time of synaptic vesicles τ_D (Cohen and Segal, 2011). Initially, D is 1 and after a current injection, i.e., an action potential, at time t_m it decreases as $D \to \beta D$ with $0 < \beta < 1$.

We used the same implementation as in (Orlandi et al., 2013; Tibau et al., 2020). Parameter values were similar to those used in (Orlandi et al., 2013) and were chosen so that the model reproduces typical behavior of cortical neurons. They are also given in the SM (**Table S2**). Here, all neurons were set to be excitatory for the sake of simplicity. Specifically, we set g_{AMPA} and g_{minis} equal to 9.5 for all simulations. These values facilitated the generation of network bursts, i.e., activity fronts that encompassed a large fraction of the network, although the timing and spatiotemporal structure of the fronts varied with the obstacles' designs. The time step in all simulations was set to 0.1 ms, with a total duration of 30 min.

5.4. Data Processing

5.4.1. Neuronal Activity, Data Filtering and Network Bursts

Simulated networks exhibited rich spontaneous activity that combined sporadic neuronal events with coherent activations of different sizes. Typically, neurons fired either individually or in a coordinated manner at a rate in the range $0.1 - 0.5$ Hz. Since effective connectivity inference was not reliable when sporadic activations were abundant, raster plots of neuronal activity were filtered to retain only coordinated activity episodes. The filtering consisted in computing first the size of coherent network activations in a window of 0.5 s, and next to inspect the distribution of sizes. About 95% of the collective events encompassed at least 25% of the network. Therefore this threshold was chosen to eliminate sporadic activations from the raster plots while only minimally affecting collective bursting episodes. The inter-burst interval (IBI) was then defined as the average time elapsing between two network bursts in which at least 25% of the network participated.

5.4.2. Initiation Points and Representative Spatiotemporal Activity Patterns

Network burst ignition events originated in specific areas of the network, which were termed "initiation points" as introduced in Orlandi et al. (2013). The spatial distribution of these events was obtained by first identifying the starting time of each burst in the raster plots. The neurons in each burst were then reindexed using the time of their first firing during the burst and that provided its spatiotemporal structure in the form of a wave front. This front was fitted to a space-time cone whose apex provided the spatial location of the origin of this burst. Wave

fronts that procured coordinates outside the area containing neurons were considered unreliable and excluded. The cone fitting assumed that activity propagated like a circular wave across the network, an assumption that was found valid only for the obstacles' configuration made of crosses. Therefore, the analysis of initiation points was carried out only for this configuration. Given the errors in the cone fitting, the final distributions of initiation points were smoothed versions of the spatially binned histograms of initiation points.

The information about the timing of burst and neuronal reindexing was also used to draw representative spatiotemporal activity patterns. The x and y coordinates of the neurons participating in the burst were mapped into a grid of 25×25 elements. The mapped data was then represented as an smoothed image plot with a color scheme proportional to the propagation time of the burst throughout the network.

5.4.3. Structural and Effective Connectivity

Structural connectivity: It corresponded to the ground truth topology that resulted from the geometric construction of the networks. Data was stored in the adjacency matrix $\mathbf{S} = \{\mathbf{s_{ij}}\}$ which is by construction directed and non-weighted. Their major topological traits were examined using the specified network measures.

Distributions of connection distances and angles for the structural connectivity: They were presented as histograms in the figures, and were obtained by combining the information about the spatial location of the neurons and their ground truth topology. The distance d_{ij} was the Euclidean distance between the centers of the somas of two physically connected neurons i and j. The corresponding angle θ_{ij} was measured as the angle between the vertical axis and the straight line corresponding to the distance d_{ij}.

Effective connectivity: It was inferred using a modified version of Transfer Entropy (TE) (Schreiber, 2000). For neurons X and Y with signals x_n and y_n indexed by $0 \leq n \leq n_{\max}$, where $n_{\max}$ is the total number of time steps in the data, TE was computed as

$$\mathrm{TE}_{Y\to X} = -\sum_{\substack{0\leq n\leq n_{\max}\\ 0\leq k\leq k_M}} p\left(x_{n+1}, x_n^{(k)}, y_n^{(k)}\right) \times \log_2 \frac{p\left(x_{n+1}\left|x_n^{(k)}, y_n^{(k)}\right.\right)}{p\left(x_{n+1}\left|x_n^{(k)}\right.\right)}, \tag{5}$$

where k is the index of the past time step considered, i.e., the length of the vectors $\{x_n^{(k)}\}$, and $k_M = 2$ is the Markov order of the model. Here, instantaneous feedback was assumed, meaning that X and Y could interact within a time bin, as in Generalized Transfer Entropy (Stetter et al., 2012; Orlandi et al., 2014). Thus, the Markov order superscript indices on $\{x_n^{(k)}\}$ and $\{y_n^{(k)}\}$ are identical. This assumption was justified because the synaptic time constants ($\simeq$ 1 ms) were much smaller than the time bins (50 ms) used. This binning also ensured that data analysis was feasible and reasonably fast. Effective connectivity was inferred for 30 min long raster plots ($n_{\max} = 36,000$) containing network bursting events only. For any connection X to Y, significance z was established by comparing the $\mathrm{TE}_{Y\to X}$ estimate with the joint distribution of TE for all input scores X' to Y and output scores X to Y' (for any X' and Y'), as

$$z = \frac{\mathrm{TE}_{Y\to X} - \langle \mathrm{TE}_{\mathrm{joint}} \rangle}{\sigma_{\mathrm{joint}}}, \tag{6}$$

where $\langle \mathrm{TE}_{\mathrm{joint}} \rangle$ is the average value of the joint distribution and σ_{joint} is its standard deviation. Significant connections were then set as those with $z \geq 2$. This threshold was considered optimal since it captured the flow of neuronal communication during activity at both local and global scales. A lower threshold of $z = 1$ yielded networks that excessively emphasized whole-network coordinated activity, effectively shaping random graphs in all studied cases. Thresholds $z \gtrsim 3$ emphasized the strongest neuron-to-neuron interactions and often yielded empty matrices. Significant connections were finally thresholded to 0 (absence of connection) and 1 (presence of connection). The final effective connectivity matrices $\mathbf{E}$ were then directed and non-weighted.

5.4.4. Network Analysis and Measures

The following network statistics and centrality measures were computed for both structural ($\mathbf{S}$, ground truth) and effective ($\mathbf{E}$) topologies.

In- and out-degree distributions and clustering coefficient: Degree statistics were computed in Python using the Brain Connectivity Toolbox (BCT) (Rubinov and Sporns, 2010). For the structural connectivity, these distributions reflected the capacity of the obstacles to shape or dictate a distinct circuitry. For the effective connectivity, they reflected the flow of activity. Clustering coefficients (CC) (Fagiolo, 2007) were computed using the Python module NetworkX (Hagberg et al., 2008). The spatial distributions represented as heatmaps in **Figure 6** show the average values of in-degree (k_{in}) and CC in square regions of side 0.031 mm, containing ≈ 0.2 neurons on average for the networks of 2 mm in diameter. Therefore, linear interpolation was used to improve readability of the heatmap. For larger network sizes, the size of the squares was scaled up proportionally to the diameter of the network.

Modularity Q: It quantified the likelihood that neurons were organized in communities, i.e., that neurons within a community were more connected with themselves than with neurons in other communities. Following Rubinov and Sporns (2010), Q was computed as

$$Q = \frac{1}{2m} \sum_{0\leq i,j\leq N} \left(A_{ij} - \frac{k_i k_j}{2m}\right) \delta(c_i, c_j), \tag{7}$$

where N is the number of neurons, A_{ij} represents the weight of the connection between i and j, $k_i = \sum_{j=1}^{N} A_{ij}$ is the sum of the weights of the connections attached to neuron i, c_i is the community to which neuron i belongs, $m = \frac{1}{2}\sum_{i,j=1}^{N} A_{ij}$, and the $\delta(u, v)$ function is 1 for $u = v$ and 0 otherwise. Optimal community structure was computed using the Louvain algorithm (Blondel et al., 2008). Q ranged from 0 to 1, with $Q \approx 0$ for a random, non-modular network and $Q \to 1$ for a strong modular organization.

Global efficiency G_{eff}: It quantified the integration capacity of the network, i.e., the performance of information exchange among neurons across the network. It was calculated using the BCT. Following (Latora and Marchiori, 2001; Rubinov and

Sporns, 2010), the efficiency E of a network of N nodes was computed as

$$E = \frac{1}{N(N-1)} \sum_{0 \leq i,j \leq N} \frac{1}{\lambda(i,j)}, \quad (8)$$

where N is the number of neurons and $\lambda(i,j)$ is the length of the shortest path connecting neurons i and j. The global efficiency G_{eff} is the relative value $G_{\text{eff}} = E/E_{\text{id}}$, where E_{id} refers to the efficiency of an ideal graph that has all $N(N-1)$ possible connections.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available upon request by the authors, without undue reservation. Source code pertaining to certain aspects of the analysis is available at https://github.com/adluinf/SpatialNetworkAnalysis.

AUTHOR CONTRIBUTIONS

A-AL designed and performed the simulations. All authors analyzed the data, contributed to the interpretation of results, and writing of the article.

FUNDING

This research was part of MESOBRAIN. MESOBRAIN has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement no. 713140. The project leading to these results has also received funding from La Caixa Foundation (ID 100010434) under the agreement LCF/PR/HR19/52160007. Research has been also supported by the Spanish Ministerio de Economia y Competitividad Projects FIS2016-78507-C2-2-P and FIS2017-90782-REDT (IBERSINC), and by the Generalitat de Catalunya Grant 2017-SGR-1061.

ACKNOWLEDGMENTS

We thank J. G. Orlandi, J. Crowe, and S. Faci-Lázaro for helpful discussions and insight.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncom.2020.00077/full#supplementary-material

REFERENCES

Achard, S., and Bullmore, E. (2007). Efficiency and cost of economical brain functional networks. *PLoS Comput. Biol.* 3:e17. doi: 10.1371/journal.pcbi.0030017

Aebersold, M. J., Dermutz, H., Forró, C., Weydert, S., Thompson-Steckel, G., Vörös, J., et al. (2016). "Brains on a chip": towards engineered neural networks. *Trends Anal. Chem.* 78, 60–69. doi: 10.1016/j.trac.2016.01.025

Alvarez-Lacalle, E., and Moses, E. (2009). Slow and fast pulses in 1-D cultures of excitatory neurons. *J. Comput. Neurosci.* 26, 475–493. doi: 10.1007/s10827-008-0123-5

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). Fast unfolding of communities in large networks. *J. Stat. Mech.* 2008:P10008. doi: 10.1088/1742-5468/2008/10/P10008

Bosi, S., Rauti, R., Laishram, J., Turco, A., Lonardoni, D., Nieus, T., et al. (2015). From 2D to 3D: novel nanostructured scaffolds to investigate signalling in reconstructed neuronal networks. *Sci. Rep.* 5, 1–11. doi: 10.1038/srep09562

Bullmore, E., and Sporns, O. (2012). The economy of brain network organization. *Nat. Rev. Neurosci.* 13, 336–49. doi: 10.1038/nrn3214

Cabral, J., Kringelbach, M. L., and Deco, G. (2017). Functional connectivity dynamically evolves on multiple time-scales over a static structural connectome: models and mechanisms. *Neuroimage* 160, 84–96. doi: 10.1016/j.neuroimage.2017.03.045

Casanova, A., Blatche, M. C., Ferre, C. A., Martin, H., Gonzalez-Dunia, D., Nicu, L., et al. (2018). Self-aligned functionalization approach to order neuronal networks at the single-cell level. *Langmuir* 34, 6612–6620. doi: 10.1021/acs.langmuir.8b00529

Cohen, D., and Segal, M. (2011). Network bursts in hippocampal microcultures are terminated by exhaustion of vesicle pools. *J. Neurophysiol.* 106, 2314–2321. doi: 10.1152/jn.00969.2010

Cohen, E., Ivenshitz, M., Amor-Baroukh, V., Greenberger, V., and Segal, M. (2008). Determinants of spontaneous activity in networks of cultured hippocampus. *Brain Res.* 1235, 21–30. doi: 10.1016/j.brainres.2008.06.022

Crowe, J., El-Tamer, A., Nagel, D., Koroleva, A., Madrid-Wolff, J., Olarte, O. E., et al. (2020). Development of two-photon polymerised scaffolds for optical interrogation and neurite guidance of human iPSC-derived cortical neuronal networks. *Lab Chip*. 20, 1792–1806. doi: 10.1039/C9LC01209E

de Santos-Sierra, D., Sendiña-Nadal, I., Leyva, I., Almendral, J. A., Anava, S., Ayali, A., et al. (2014). Emergence of small-world anatomical networks in self-organizing clustered neuronal cultures. *PLoS ONE* 9:e85828. doi: 10.1371/journal.pone.0085828

Díaz Lantada, A., Pareja Sánchez, B., Gómez Murillo, C., and Urbieta Sotillo, J. (2013). Fractals in tissue engineering: toward biomimetic cell-culture matrices, microsystems and microstructured implants. *Expert Rev. Med. Devices* 10, 629–648. doi: 10.1586/17434440.2013.827506

Downes, J. H., Hammond, M. W., Xydas, D., Spencer, M. C., Becerra, V. M., Warwick, K., et a;. (2012). Emergence of a small-world functional network in cultured neurons. *PLoS Comput. Biol.* 8:e1002522. doi: 10.1371/journal.pcbi.1002522

Faci-Lázaro, S., Soriano, J., and Gómez-Gardeñes, J. (2019). Impact of targeted attack on the spontaneous activity in spatial and biologically-inspired neuronal networks. *Chaos* 29:083126. doi: 10.1063/1.5099038

Fagiolo, G. (2007). Clustering in complex directed networks. *Phys. Rev. E* 76:026107. doi: 10.1103/PhysRevE.76.026107

Feinerman, O., Rotem, A., and Moses, E. (2008). Reliable neuronal logic devices from patterned hippocampal cultures. *Nat. Phys.* 4, 967–973. doi: 10.1038/nphys1099

Friedman, N., Ito, S., Brinkman, B. A., Shimono, M., Deville, R. E., Dahmen, K. A., et al. (2012). Universal critical dynamics in high resolution neuronal avalanche data. *Phys. Rev. Lett.* 108, 1–5. doi: 10.1103/PhysRevLett.108.208102

Gladkov, A., Pigareva, Y., Kutyina, D., Kolpakov, V., Bukatin, A., Mukhina, I., et al. (2017). Design of cultured neuron networks *in vitro* with predefined connectivity using asymmetric microfluidic channels. *Sci. Rep.* 7, 1–14. doi: 10.1038/s41598-017-15506-2

Hagberg, A., Swart, P., and S Chult, D. (2008). *Exploring Network Structure, Dynamics, and Function Using NetworkX*. Technical report, Los Alamos National Lab (LANL), Los Alamos, NM.

Haimovici, A., Tagliazucchi, E., Balenzuela, P., and Chialvo, D. R. (2013). Brain organization into resting state networks emerges at criticality on a model of the human connectome. *Phys. Rev. Lett*. 110:178101. doi: 10.1103/PhysRevLett.110.178101

Hernández-Navarro, L., Orlandi, J. G., Cerruti, B., Vives, E., and Soriano, J. (2017). Dominance of metric correlations in two-dimensional neuronal cultures described through a random field ising model. *Phys. Rev. Lett*. 118, 1–5. doi: 10.1103/PhysRevLett.118.208101

Hofman, M. A. (2014). Evolution of the human brain: when bigger is better. *Front. Neuroanat*. 8:15. doi: 10.3389/fnana.2014.00015

Holloway, P. M., Hallinan, G. I., Hegde, M., Lane, S. I., Deinhardt, K., and West, J. (2019). Asymmetric confinement for defining outgrowth directionality. *Lab Chip* 19, 1484–1489. doi: 10.1039/C9LC00078J

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, 10240–5. doi: 10.1073/pnas.0701519104

Honey, C. J., Thivierge, J.-P., and Sporns, O. (2010). Can structure predict function in the human brain? *Neuroimage* 52, 766–776. doi: 10.1016/j.neuroimage.2010.01.071

Humphries, M. D., and Gurney, K. (2008). Network 'small-world-ness': a quantitative method for determining canonical network equivalence. *PLoS ONE* 3:e0002051. doi: 10.1371/journal.pone.0002051

Izhikevich, E. (2003). Simple model of spiking neurons. *IEEE Trans. Neural Netw*. 14, 1569–1572. doi: 10.1109/TNN.2003.820440

Izhikevich, E. M. (2007). *Dynamical Systems in Neuroscience*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/2526.001.0001

Johnson, J. K., Wright, N. C., Xiá, J., and Wessel, R. (2019). Single-cell membrane potential fluctuations evince network scale-freeness and quasicriticality. *J. Neurosci*. 39, 4738–4759. doi: 10.1523/JNEUROSCI.3163-18.2019

Kunze, A., Giugliano, M., Valero, A., and Renaud, P. (2011). Micropatterning neural cell cultures in 3D with a multi-layered scaffold. *Biomaterials* 32, 2088–2098. doi: 10.1016/j.biomaterials.2010.11.047

Larramendy, F., Yoshida, S., Maier, D., Fekete, Z., Takeuchi, S., and Paul, O. (2019). 3D arrays of microcages by two-photon lithography for spatial organization of living cells. *Lab Chip* 19, 875–884. doi: 10.1039/C8LC01240G

Latora, V., and Marchiori, M. (2001). Efficient behavior of small-world networks. *Phys. Rev. Lett*. 87:198701. doi: 10.1103/PhysRevLett.87.198701

Li, W., Xu, Z., Huang, J., Lin, X., Luo, R., Chen, C.-H., et al. (2014). Neuroarray: a universal interface for patterning and interrogating neural circuitry with single cell resolution. *Sci. Rep*. 4:4784. doi: 10.1038/srep04784

Marshall, N., Timme, N. M., Bennett, N., Ripp, M., Lautzenhiser, E., and Beggs, J. M. (2016). Analysis of power laws, shape collapses, and neural complexity: new techniques and MATLAB support via the NCC toolbox. *Front. Physiol*. 7:250. doi: 10.3389/fphys.2016.00250

Massobrio, P., Pasquale, V., and Martinoia, S. (2015). Self-organized criticality in cortical assemblies occurs in concurrent scale-free and small-world networks. *Sci. Rep*. 5, 1–16. doi: 10.1038/srep10578

Messé, A., Rudrauf, D., Benali, H., and Marrelec, G. (2014). Relating structure and function in the human brain: relative contributions of anatomy, stationary dynamics, and non-stationarities. *PLoS Comput. Biol*. 10:e1003530. doi: 10.1371/journal.pcbi.1003530

Okujeni, S., Kandler, S., and Egert, U. (2017). Mesoscale architecture shapes initiation and richness of spontaneous network activity. *J. Neurosci*. 37, 3972–3987. doi: 10.1523/JNEUROSCI.2552-16.2017

Orlandi, J. G., Soriano, J., Alvarez-Lacalle, E., Teller, S., and Casademunt, J. (2013). Noise focusing and the emergence of coherent activity in neuronal cultures. *Nat. Phys*. 9, 582–590. doi: 10.1038/nphys2686

Orlandi, J. G., Stetter, O., Soriano, J., Geisel, T., and Battaglia, D. (2014). Transfer entropy reconstruction and labeling of neuronal connections from simulated calcium imaging. *PLoS ONE* 9:e98842. doi: 10.1371/journal.pone.0098842

Pernice, V., Deger, M., Cardanobile, S., and Rotter, S. (2013). The relevance of network micro-structure for neural dynamics. *Front. Comput. Neurosci*. 7:72. doi: 10.3389/fncom.2013.00072

Poil, S. S., Van Ooyen, A., and Linkenkaer-Hansen, K. (2008). Avalanche dynamics of human brain oscillations: Relation to critical branching processes and temporal correlations. *Hum. Brain Mapp*. 29, 770–777. doi: 10.1002/hbm.20590

Renault, R., Durand, J.-B., Viovy, J.-L., and Villard, C. (2016). Asymmetric axonal edge guidance: a new paradigm for building oriented neuronal networks. *Lab Chip* 16, 2188–2191. doi: 10.1039/C6LC00479B

Rubinov, M., and Sporns, O. (2010). Complex network measures of brain connectivity: uses and interpretations. *Neuroimage* 52, 1059–1069. doi: 10.1016/j.neuroimage.2009.10.003

Schmeltzer, C., Soriano, J., Sokolov, I. M., and Rüdiger, S. (2014). Percolation of spatially constrained Erdös-Rényi networks with degree correlations. *Phys. Rev. E* 89:012116. doi: 10.1103/PhysRevE.89.012116

Schreiber, T. (2000). Measuring information transfer. *Phys. Rev. Lett*. 85, 461–464. doi: 10.1103/PhysRevLett.85.461

Schroeter, M. S., Charlesworth, P., Kitzbichler, M. G., Paulsen, O., and Bullmore, E. T. (2015). Emergence of rich-club topology and coordinated dynamics in development of hippocampal functional networks *in vitro*. *J. Neurosci*. 35, 5459–5470. doi: 10.1523/JNEUROSCI.4259-14.2015

Schröter, M., Paulsen, O., and Bullmore, E. T. (2017). Micro-connectomics: probing the organization of neuronal networks at the cellular scale. *Nat. Rev. Neurosci*. 18, 131–146. doi: 10.1038/nrn.2016.182

Severino, F. P. U., Ban, J., Song, Q., Tang, M., Bianconi, G., Cheng, G., et al. (2016). The role of dimensionality in neuronal network dynamics. *Sci. Rep*. 6, 1–14. doi: 10.1038/srep29640

Shein-Idelson, M., Ben-Jacob, E., and Hanein, Y. (2011). Engineered neuronal circuits: a new platform for studying the role of modular topology. *Front. Neuroeng*. 4:10. doi: 10.3389/fneng.2011.00010

Simitzi, C., Ranella, A., and Stratakis, E. (2017). Controlling the morphology and outgrowth of nerve and neuroglial cells: the effect of surface topography. *Acta Biomater*. 51, 21–52. doi: 10.1016/j.actbio.2017.01.023

Slomowitz, E., Styr, B., Vertkin, I., Milshtein-Parush, H., Nelken, I., Slutsky, M., et al. (2015). Interplay between population firing stability and single neuron dynamics in hippocampal networks. *eLife* 4:e04378. doi: 10.7554/eLife.04378

Soriano, J., Martínez, M. R., Tlusty, T., and Moses, E. (2008). Development of input connections in neural cultures. *Proc. Natl. Acad. Sci. U.S.A*. 105, 13758–13763. doi: 10.1073/pnas.0707492105

Sorkin, R., Gabay, T., Blinder, P., Baranes, D., Ben-Jacob, E., and Hanein, Y. (2006). Compact self-wiring in cultured neural networks. *J. Neural Eng*. 3, 95–101. doi: 10.1088/1741-2560/3/2/003

Sporns, O. (2011). The non-random brain: efficiency, economy, and complex dynamics. *Front. Comput. Neurosci*. 5:5. doi: 10.3389/fncom.2011.00005

Stetter, O., Battaglia, D., Soriano, J., and Geisel, T. (2012). Model-free reconstruction of excitatory neuronal connectivity from calcium imaging signals. *PLoS Comput. Biol*. 8:e1002653. doi: 10.1371/journal.pcbi.1002653

Stiso, J., and Bassett, D. S. (2018). Spatial embedding imposes constraints on neuronal network architectures. *Trends Cogn. Sci*. 22, 1127–1142. doi: 10.1016/j.tics.2018.09.007

Tang-Schomer, M. D., White, J. D., Tien, L. W., Schmitt, L. I., Valentin, T. M., Graziano, D. J., et al. (2014). Bioengineered functional brain-like cortical tissue. *Proc. Natl. Acad. Sci. U.S.A*. 111, 13811–13816. doi: 10.1073/pnas.1324214111

Teller, S., Estévez-Priego, E., Granell, C., Tornero, D., Andilla, J., Olarte, O. E., et al. (2019). Spontaneous functional recovery after focal damage in neuronal cultures. *eNeuro*. 7, 1–13. doi: 10.1523/ENEURO.0254-19.2019

Teller, S., Granell, C., De Domenico, M., Soriano, J., Gómez, S., and Arenas, A. (2014). Emergence of assortative mixing between clusters of cultured neurons. *PLoS Comput. Biol*. 10:e1003796. doi: 10.1371/journal.pcbi.1003796

Tibau, E., Ludl, A. A., Rudiger, S., Orlandi, J. G., and Soriano, J. (2020). Neuronal spatial arrangement shapes effective connectivity traits of *in vitro* cortical networks. *IEEE Trans. Netw. Sci. Eng*. 7, 435–448. doi: 10.1109/TNSE.2018.2862919

Voges, N., and Perrinet, L. (2012). Complex dynamics in recurrent cortical networks based on spatially realistic connectivities. *Front. Comput. Neurosci*. 6:41. doi: 10.3389/fncom.2012.00041

Watts, D. J., and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature* 393, 440–442. doi: 10.1038/30918

Werner, G. (2010). Fractals in the nervous system: conceptual implications for theoretical neuroscience. *Front. Physiol*. 1:15. doi: 10.3389/fphys.2010.00015

Yamada, A., Vignes, M., Bureau, C., Mamane, A., Venzac, B., Descroix, S., et al. (2016). In-mold patterning and actionable axo-somatic compartmentalization for on-chip neuron culture. *Lab Chip* 16, 2059–2068. doi: 10.1039/C6LC00414H

Yamamoto, H., Moriya, S., Ide, K., Hayakawa, T., Akima, H., Sato, S., et al. (2018). Impact of modular organization on dynamical richness in cortical networks. *Sci. Adv*. 4:eaau4914. doi: 10.1126/sciadv.aau4914

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.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 04 September 2020
doi: 10.3389/fncom.2020.00075

A Computational Model of the Cholinergic Modulation of CA1 Pyramidal Cell Activity

Adam Mergenthal, Jean-Marie C. Bouteiller, Gene J. Yu and Theodore W. Berger*

Biomedical Engineering Department, Center for Neural Engineering, University of Southern California, Los Angeles, CA, United States

Dysfunction in cholinergic modulation has been linked to a variety of cognitive disorders including Alzheimer's disease. The important role of this neurotransmitter has been explored in a variety of experiments, yet many questions remain unanswered about the contribution of cholinergic modulation to healthy hippocampal function. To address this question, we have developed a model of CA1 pyramidal neuron that takes into consideration muscarinic receptor activation in response to changes in extracellular concentration of acetylcholine and its effects on cellular excitability and downstream intracellular calcium dynamics. This model incorporates a variety of molecular agents to accurately simulate several processes heretofore ignored in computational modeling of CA1 pyramidal neurons. These processes include the inhibition of ionic channels by phospholipid depletion along with the release of calcium from intracellular stores (i.e., the endoplasmic reticulum). This paper describes the model and the methods used to calibrate its behavior to match experimental results. The result of this work is a compartmental model with calibrated mechanisms for simulating the intracellular calcium dynamics of CA1 pyramidal cells with a focus on those related to release from calcium stores in the endoplasmic reticulum. From this model we also make various predictions for how the inhibitory and excitatory responses to cholinergic modulation vary with agonist concentration. This model expands the capabilities of CA1 pyramidal cell models through the explicit modeling of molecular interactions involved in healthy cognitive function and disease. Through this expanded model we come closer to simulating these diseases and gaining the knowledge required to develop novel treatments.

Keywords: hippocampus, acetylcholine, CA1, muscarinic, compartmental model, pyramidal, computational

OPEN ACCESS

Edited by:
Tatyana Sharpee,
Salk Institute for Biological Studies, United States

Reviewed by:
Carmen Castro Canavier,
Louisiana State University, United States
Ervin Wolf,
University of Debrecen, Hungary

****Correspondence:***
Adam Mergenthal
mergenth@usc.edu

Received: *09 April 2020*
Accepted: *17 July 2020*
Published: *04 September 2020*

Citation:
Mergenthal A, Bouteiller J-MC, Yu GJ and Berger TW (2020) A Computational Model of the Cholinergic Modulation of CA1 Pyramidal Cell Activity. Front. Comput. Neurosci. 14:75. doi: 10.3389/fncom.2020.00075

1. INTRODUCTION

Acetylcholine (ACh) directly modulates the activity of neurons within every subregion of the hippocampus, including both principal neurons and interneurons (Aznavour et al., 2002; Takács et al., 2018). The dense distribution of the cholinergic terminals within the hippocampus suggests that this neurotransmitter plays an important role in healthy hippocampal functioning. This important role is further evidenced by the correlation of dysfunctions reported in cholinergic terminals with cognitive impairment. The progression of Alzheimer's disease (AD) has long been associated with the decline of cholinergic markers in the hippocampus (Schliebs and Arendt, 2011). Other cognitive disorders such as depression and schizophrenia are also associated with

alterations in cholinergic disregulation (Higley and Picciotto, 2014). On an even broader scale, changes in cholinergic expression are associated with the cognitive decline due to advanced age (Schliebs and Arendt, 2011). The variety of cognitive dysfunctions related to ACh suggests that it plays not only an important, but a complex role. In a 2-year double-blind study, 35% of patients taking an acetylcholinesterase inhibitor to slow cognitive decline due to AD had a recurrence of major depressive episodes vs. 19% of those on a placebo (Reynolds et al., 2011). In other words, a drug meant to counteract one form of cholinergic dysfunction exacerbated a separate form of cholinergic dysfunction. Developing better treatments for these disorders requires a better understanding of the dynamics of healthy cholinergic modulation. Currently, ACh is understood to play a role in a variety of cognitive processes. We will summarize some of these effects briefly but for fuller reviews (see Dannenberg et al., 2017; Solari and Hangya, 2018). Acetylcholine has long been understood to be involved in the generation of theta oscillations (4–12 Hz) in the hippocampus. Theta oscillations are theorized to organize memory encoding and retrieval into distinct phases (Hasselmo et al., 2002). Acetylcholine seems to be involved with the generation of the lower frequency portion of theta oscillations, as these frequencies can be blocked by the cholinergic receptor antagonist atropine (Kramis et al., 1975). On a behavioral level, the blockade of cholinergic receptors in animal models leads to a variety of memory deficits involving both spatial navigation and the acquisition of conditioned fear responses (Jiang et al., 2016; Solari and Hangya, 2018). These effects result from the activation of a variety of cholinergic receptors in the hippocampus. These receptors can be sorted into two types. The first type, nicotinic receptors, act as ionotropic receptors and allow the passage of ions through the plasma membrane. In the CA1, nicotinic receptors primarily modulate interneuron activity (McQuiston, 2014), but they also appear in low densities on pyramidal cells (Kalappa et al., 2010). The second type, muscarinic receptors, have a much larger effect in modulating CA1 pyramidal cell activity (Dasari and Gulledge, 2011). These receptors are G protein coupled receptors with their activation setting off a cascade of intracellular reactions. Among the five subtypes of muscarinic acetylcholine receptors (mAChRs), the subtypes that primarily modulate CA1 pyramidal activity are the M1 and M4 mAChRs. M4 mAChRs suppress glutamatergic release from excitatory synapses originating from the CA3 subregion (Dasari and Gulledge, 2011). M1 mAChRs are present throughout the cell's morphology and alter its overall excitability along with altering the intracellular calcium dynamics (Dasari and Gulledge, 2011). Thus, the M1 mAChRs are responsible for the majority of the cholinergic response in this cell type. The M1 mAChR, as a G-protein coupled receptor, activates a cascade of intracellular reactions (Falkenburger et al., 2010a,b). It is through these reactions that the M1 receptor is able to modulate the behavior of a variety of ion channels. Teithehe M-current was given that name due to muscarinic receptors suppressing its activity (Brown and Adams, 1980). Inhibition of this current in CA1 pyramidal cells through bath application of the M-current antagonist XE991 lead to a depolarized resting membrane potential and increased spiking activity (Shah et al., 2008). This current was also shown to be inhibited after bath application of the muscarinic agonist Oxotremorine-M (Oxo-M) (Carver and Shapiro, 2019). The channels responsible for the M-current, Kv7 Potassium channels, require phosphatidylinositol 4,5-bisphosphate (PIP_2) in the cell membrane to maintain its open state. M1 activation leads to the activation of phospholipase C (PLC) which hydrolyzes PIP_2 into inositol(1,4,5)triphosphate (IP_3) and diacylglycerol (Falkenburger et al., 2010a,b). It is through this depletion of PIP_2 that mAChRs inhibit the M-current. Also, by producing IP_3, M1 receptors trigger the release of calcium from the endoplasmic reticulum (ER) via IP_3 receptors. This leads to an increase in intracellular calcium which activates calcium dependent potassium (SK) channels. In CA1 pyramidal cells M1 activation is followed by a hyperpolarization which is able to inhibit action potentials. These hyperpolarizations can be blocked through the application of apamin, an SK channel antagonist (Dasari and Gulledge, 2011). **Figure 1** provides both a flowchart and a cartoon illustrate these processes. One long term goal of our lab has been to create a large scale model of the hippocampus and through this model, gain a better understanding of the underlying dynamics of this system (Hendrickson et al., 2015), thereby facilitating the development of better treatments (electrical or pharmaceutical) to alleviate hippocampal dysfunctions. Experimental evidence has demonstrated that cholinergic modulation plays an important role in controlling the dynamics of this system. This has driven the development of this single cell model, which will act as a foundation for integrating cholinergic modulation into our efforts for a large-scale hippocampal model. We have chosen to build the single cell model on a biophysically realistic basis wherever possible. This is for two reasons. First, the collection of experimental data for calibrating a model gives a perspective on the depth of understanding and raises questions to guide further *in vitro* or *in vivo* experimental efforts. Second, the inclusion of biochemical mechanisms allows for broad parametric manipulations which (i) facilitate the simulation of pathological processes and disease states and (ii) provide useful insights for the identification and development of novel treatment options. By creating a biophysically realistic model, we have developed a tool that allows more cohesive collaboration with other experimental efforts. What follows is a description of a model for the cholinergic modulation of the somatic activity of pyramidal cells within the CA1 region of the hippocampus. Within the hippocampus, this cell type is the most studied in terms of cholinergic modulation and will constitute a solid foundation for the construction of larger cell network models.

2. MATERIALS AND METHODS

The primary task of this research was to evaluate and bring together a variety of mechanisms and models to accurately capture the dynamic response of CA1 pyramidal cells to acetylcholine. As a starting point, we used a compartmental model of the CA1 pyramidal cell (mpg141209_A_idA as downloaded from ModelDB) (Migliore et al., 2018) previously developed for the NEURON simulation environment (Carnevale

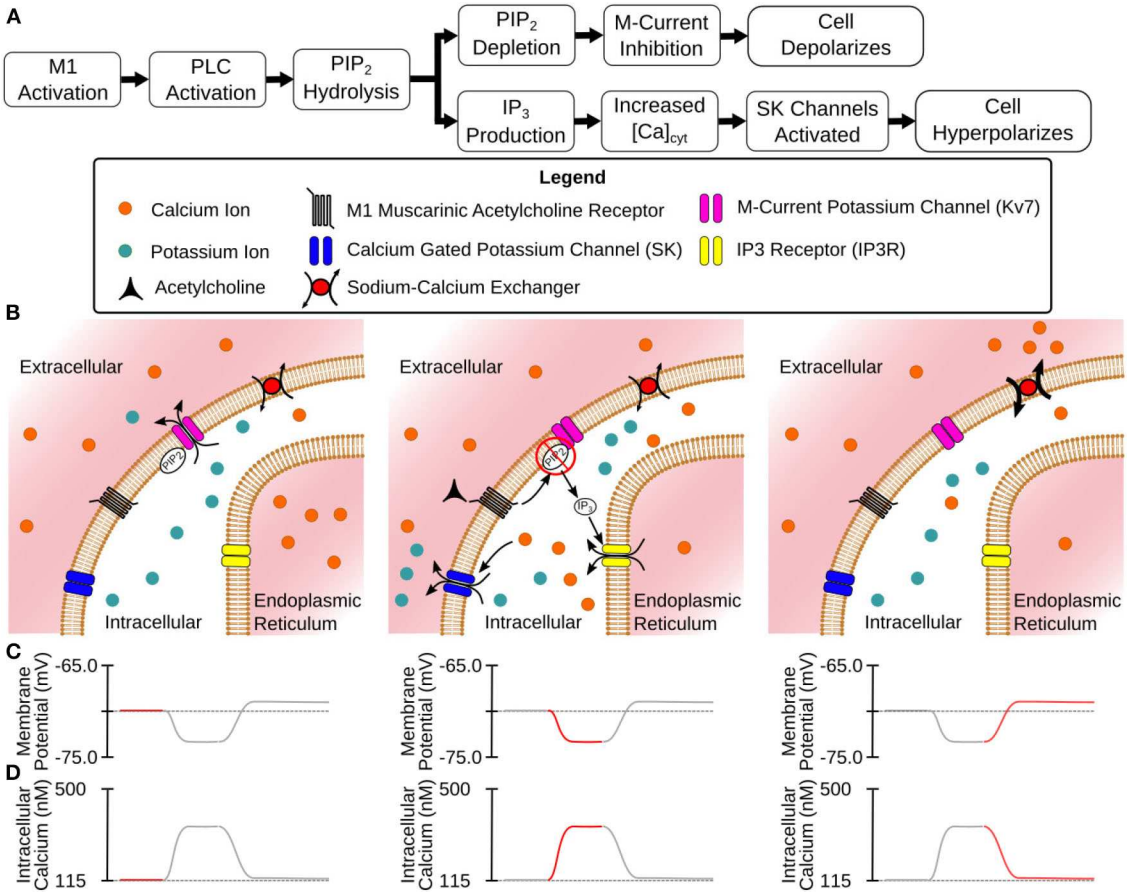

FIGURE 1 | Mechanisms of CA1 pyramidal cell modulation by M1 MAChRs. **(A)** Flowchart description of the steps between M1 activation and modulated membrane potential. **(B)** Illustrated stages of activation: (Left) Before activation the cell is at rest with Kv7 channels open. (Center) M1 activation leads to hydrolysis of PIP_2 from cell membrane (inhibiting Kv7 channels) and the release of intracellular Ca^{2+} (activating SK channels) through the generation of IP_3. (Right) As Ca^{2+} is extruded from the intracellular space SK channels close while Kv7 channels remain closed. **(C)** Membrane potential at different stages of activation. **(D)** Intracellular calcium levels at different stages of activation.

and Hines, 2006). We chose to use this simulation environment as its RXD module (McDougal et al., 2013) allowed us to efficiently expand the model's intracellular calcium mechanisms. The code for these simulations was developed in the Python programming language. The base model included mechanisms for the M-current, SK channels, voltage-gated calcium channels (VGCC). Entry through VGCCs was the only mechanism through which intracellular calcium increased, while calcium efflux was simulated as an exponential decay of the intracellular calcium to its resting value. As one of the focuses of this work was to simulate intracellular calcium release we needed to insert and calibrate all of the mechanisms for simulating the storage and release of calcium from the endoplasmic reticulum, buffering the intracellular calcium concentration, and extrusion of excess calcium into the extracellular space. Without these mechanisms, none of the inhibitory effects seen in **Figure 1** could be replicated. These calcium mechanisms were only expanded in the sections that comprise the soma and the first 200 μm of the apical dendritic trunk. **Figure S1** illustrates which sections within the full morphology were given expanded calcium mechanisms. One reason for the decision to only expand the calcium model into these sections was that the calcium dynamics in these regions are the most studied due their diameters being large enough for calcium imaging using fluorescent dyes. Second, cholinergic modulation in synaptic spines seems to play a role in plasticity (Dennis et al., 2016). However, plasticity in these synapses is also dependent upon postsynaptic spiking activity. To properly simulate how plasticity is altered by cholinergic modulation requires we first make a working model of how cholinergic modulation alters cell excitability and spike generation. Finally, the mechanisms of action differ between synaptic and somatic modulation. For instance, the hyperpolarization seen at the soma is due to the activation of SK channels as evidenced by its blockade by apamin (Dasari and Gulledge, 2011), while synaptic cholinergic modulation has been tied to the inhibition of SK channels (Buchanan et al., 2010). Calibrating these differing mechanisms requires a separate series of simulations and would be best

explained in a separate work. Our goal in selecting additional mechanisms was to create a relatively simple model capable of replicating intracellular calcium dynamics. Disease and age have been reported to alter several of the mechanisms included (e.g., calcium buffering; Gant et al., 2006; Oh et al., 2013). By incorporating mechanistic models for these altered states, we can explore how the cell behavior changes, and how these changes impact network-level outcomes. Of importance, fidelity to the biochemistry of the intracellular space must be balanced against the realities of computational modeling. A model that includes all of the known molecular interactions would have too many parameters to constrain with the available experimental evidence. Additionally, simulations using this model would be computationally expensive even for a single cell model. In addition, our goal of including this model into large scale network simulations only exacerbates this limitation. We have thus strived to include the minimum collection of mechanisms that is necessary for capturing cholinergic modulation in the soma and apical trunk. Expanding the model to other regions and to include other mechanisms will be performed in subsequent work. A visualization of the mechanisms in the expanded calcium model can be found in **Figure 2**, while the concentrations and kinetic parameters for these mechanisms can be found in **Tables S1**, **S2**, respectively. The addition of a mechanism often required constraining parameter values to properly replicate experimental results. In order to simplify the calibration process, the mechanisms were divided into groups based upon region of action (e.g., endoplasmic reticulum vs. intracellular). These groups were then calibrated in a specific order, starting from protocols that required the smallest number of mechanisms and comprised a minimum number of interdependent parameters. For example, the rate at which the endoplasmic reticulum (ER) regains depleted calcium at rest is based upon the balance between the rate of calcium uptake from sarco/endoplasmic reticulum calcium pumps (SERCA) vs. the rate of calcium leakage from the ER. Since the conductance of VGCCs does not factor into this result it can be ignored. Conversely, replicating intracellular calcium transients after an action potential requires constraining parameters for VGCC conductance and calcium extrusion, in addition to SERCA and ER leak flux, as ER calcium sequestration alters the dynamics in the intracellular space. Since we could relatively isolate the ER mechanisms, those parameters were calibrated first. This simplified the calibration of later parameters based on results that depend on more mechanisms. The following sections describe the mechanisms that were implemented and the experimental data from which constrained these parameter values.

2.1. Calibrating the Endoplasmic Reticulum

The first step in creating the model was to calibrate the parameters pertaining to the ER. We chose to model the ER as an idealized 10% of the intracellular volume to avoid explicitly modeling the intricate and dynamic geometry of the ER. Reconstructions of the ER in CA1 pyramidal cells have focused on the organelle's volume in either the dendritic branches or synaptic spines while ignoring the volume of the ER in the soma and apical trunk. Using smaller values for the percentage of ER volume, such as those found in dendritic reconstructions (2–8% of dendrite volume) (Spacek and Harris, 1997), decreased the capacity of calcium storage such that the model could not replicate the amplitude of calcium release events. The 10% value is therefore a compromise that allows larger intracellular calcium release events while remaining near the experimentally measured range. The resting concentration inside the ER was initialized at 175 μM (Solovyova et al., 2002). For the initial calibration, there were three mechanisms that defined the ER calcium dynamics: calreticulin (CALR) concentration, SERCA pumps, and calcium leak. The inositol(1,4,5)triphosphate receptor mechanism (IP_3R) was calibrated at a later stage as the IP_3R model produced negligible currents at resting IP_3 concentrations. CALR acts as the major calcium buffer in the lumen of the ER and its concentration defines the amount of buffered calcium reserves for a given lumenal calcium concentration. We used the CALR kinetics and concentration found in an earlier ER model (Doi et al., 2005).

Expressions (1) and (2) were used for the SERCA pump mechanism while Expression (3) was the formula used to calculate the leak of calcium from the ER into the cytosol. Expression (4) shows the chemical formula used for CALR binding to calcium.

$$Ca^{2+}_{cyt} \xrightarrow{k^S_f} Ca^{2+}_{ER} \tag{1}$$

$$k^S_f = \frac{g_S \cdot [Ca^{2+}_{cyt}]^2}{[Ca^{2+}_{cyt}]^2 + 0.0013^2} \tag{2}$$

$$Ca^{2+}_{ER} \xrightarrow{k_f^{leak\ ER}} Ca^{2+}_{cyt} \tag{3}$$

$$\mathrm{CALR} + \mathrm{Ca}^2 + \underset{k_b^{calr}}{\overset{k_f^{calr}}{\rightleftharpoons}} \mathrm{CALRCa} \tag{4}$$

From these mechanisms we calibrated two parameters, g_S, the SERCA conductance, and $k_f^{leak\ ER}$, the rate of leakage from the ER. Due to the model ER not having a set geometry it also lacks a set surface area. Therefore these mechanisms were implemented as direct fluxes between the two volumes without consideration of surface density. To constrain the SERCA and leak mechanisms, we used an experimental result (Garaschuk et al., 1997) for which the return of Ca^{2+}_{ER} to resting concentrations was fit with an exponential function with a time constant of 59 s. This time constant along with the experimental Ca^{2+}_{ER} resting concentration gave us a target with which we manually calibrated g_S and $k_f^{leak\ ER}$. The results of this calibration are illustrated in **Figure 2B**.

2.2. Calibrating Intracellular Calcium and Indicator Model

With the ER related mechanisms calibrated, we moved to calibrating the mechanisms related to the intracellular space. The major mechanisms of interest in this portion of the model pertain to the extrusion of excess calcium into the extracellular space [i.e., plasma membrane calcium pumps (PMCA)] and the conductance of VGCCs. However, experimental evidence to constrain these parameters required the addition of mechanisms

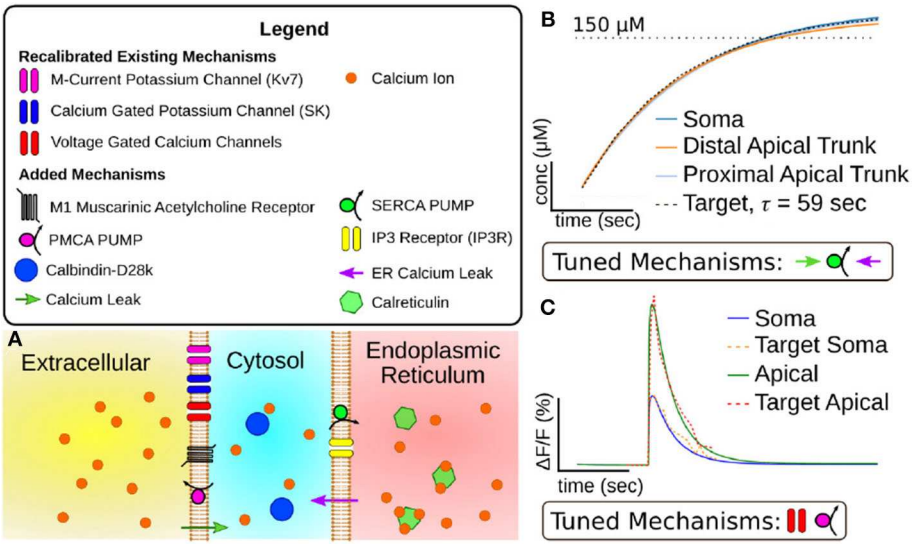

FIGURE 2 | (A) Visualization of mechanisms included in expanded CA1 pyramidal cell calcium model. **(B)** Results of tuning model mechanisms that replenish calcium stores in the endoplasmic reticulum without action potentials. Plotted value is simulated calcium stored in endoplasmic reticulum over time (x-axis = 50 s, y-axis = 50 μM). Target data is based on experiments in Garaschuk et al. (1997). **(C)** Results of tuning model mechanisms that determine calcium dynamics following an action potential. Plotted value is the percent change in fluorescence of a simulated calcium indicator (OGB-1) over time (Scale: x-axis = 1 s, y-axis= 10% change in fluorescence). Target data is based on results from Power and Sah (2002).

to replicate experiments using fluorescent calcium indicators. Fluorescence measurements constitute the primary method to visualize calcium dynamics. However these indicators act as a high affinity calcium buffer and alter the very dynamics they are supposed to report. We therefore included mechanisms to simulate the binding of calcium to Oregon Green BAPTA-1 (OGB-1), as this was the indicator used in the experimental results we sought to replicate. The kinetic parameters we used for the OGB-1 mechanism were based on measurements in an intracellular environment as interactions with intracellular ions can change the affinity from its reported *in vitro* value (Thomas et al., 2000). Expression (5) was the chemical formula used for the binding of calcium to OGB-1. According to the product information sheet OGB-1 bound to calcium fluoresces 14 times the rate of the unbound state (Molecular Probes, 2005). We used this fact to create Expression (6), which provides a method to calculate the simulated fluorescence. With this mechanism we could use fluorescence experiments that used this calcium indicator to constrain the other intracellular mechanisms. This OGB-1 mechanism for creating a simulated fluorescence was only used in this portion of the calibration process and was not included in later parameter calibrations.

$$\mathrm{OGB} + \mathrm{Ca}^2 + \underset{k_b^{ogb1}}{\overset{k_f^{ogb1}}{\rightleftharpoons}} \mathrm{OGBCa} \tag{5}$$

$$F = f_{mult} \cdot OGB1Ca + OGB1 \tag{6}$$

Deciding what concentration of Calbindin-D28k (CB) to use in our simulations was another obstacle. The presence of CB is among the ways that CA1 pyramidal cells display heterogeneity, with only around 50% of cells expressing this protein (Müller et al., 2005). Expression of CB is not correlated with the bursting/regular firing characteristic that serves as the major dichotomy within CA1 pyramidal cells (Baimbridge et al., 1991), so the spiking response of base compartmental model could not be used as a constraint. Additionally CB is mostly mobile, so a large portion of CB likely diffused out of the cell and into the electrodes used to inject the fluorescent indicators as demonstrated in Müller et al. (2005). These factors make it difficult to have full confidence in the intracellular concentration of CB during the fluorescent measurements we used to calibrate the parameters. In simulations replicating fluorescent data, we assumed these cells did express this protein but that the concentration was diminished. We set the diminished concentration to 20% of its regular value as that proportion was estimated to be immobile in neurons (Schmidt et al., 2005). Expression (7) is the chemical formula for the binding of CB to calcium.

$$\mathrm{CB} + \mathrm{Ca}^2 + \underset{k_b^{cb}}{\overset{k_f^{cb}}{\rightleftharpoons}} \mathrm{CBCa} \tag{7}$$

Expression (8) describes the binding of pmca to calcium while Expression (9) describes the release of calcium into the extracellular volume. This series of reactions describes how PMCA acts as the mechanism for the extrusion of calcium from the cytosol to the extracellular space.

$$\mathrm{PMCA} + \mathrm{Ca}^2 + \underset{k_b^{pmca\ ca}}{\overset{k_f^{pmca\ ca}}{\rightleftharpoons}} \mathrm{PMCACa} \tag{8}$$

$$\text{PMCACa} \xrightarrow{k^{pmca\ rel}} \text{PMCA} \tag{9}$$

Among these mechanisms there were two parameters that required calibration. First, we needed to tune the overall rate of calcium extrusion due to PMCA. While we had the parameters for its binding kinetics, we needed to tune the overall flux by altering the mechanism's surface density. Second, we had to alter the conductances of the VGCCs. In altering the model we had expanded the volume the model tracked while calculating calcium concentration. The original conductance values were tuned assuming a thin shell on the inner surface of the cell membrane. This expanded volume required increasing the channel conductances such that the calcium influx was enough to drive the fluctuations seen in the target experimental data.

As our target for constraining the intracellular calcium dynamics we chose the calcium fluorescence following an action potential (AP) (Power and Sah, 2002) as this protocol minimized the amount of calcium released from the ER. These experiments measured fluorescence transients in both the soma and the apical dendritic trunk, allowing us to calibrate separate parameter values for different section types. To calibrate these mechanisms we induced a simulated AP. We then modified the parameters values by hand until the simulated calcium fluorescence matched the target data. By altering the VGCC conductances we could alter the overall amplitude of the calcium transient. Due to differences in target amplitude, separate VGCC conductance values were calibrated for the soma and apical dendritic trunk. Increasing the density of PMCA decreased the maximum amplitude of the calcium along with increasing the rate the transient decayed to resting concentrations. The results of this calibration can be seen in **Figure 2C**.

2.3. Calibrating Calcium Release and Spike Acceleration

With the components for the calcium dynamics in place, the next step was to calibrate the production of IP_3 following M1 mAChR activation. For the M1 mAChR model we turned to the kinetic models developed by the Hille lab (Falkenburger et al., 2010a,b, 2013; Kruse et al., 2016). This model included mechanisms that describe the process from the receptor activation by its agonist to PIP_2 hydrolysis into IP_3 and DAG. A schematic representation of this model including all of the associated reactions can be found in **Figure S2**. However, the model required notable modifications to fit our purpose. First, the Hille model simulated the agonist Oxo-M, not acetylcholine (ACh). While Oxo-M is an important muscarinic agonist, the goal of simulating endogenous cholinergic modulation required the mechanism to include ACh. Our work added the action of ACh on the M1 mAChR model through the calibration of additional parameters. Second, the rate of IP_3 production was extremely slow compared to the behavior seen in CA1 pyramidal cells. Recordings of spiking CA1 pyramidal cells exposed to brief (40 ms) pulses of ACh were provided by the authors of Gulledge and Kawaguchi (2007). From these recordings we selected a subset of traces demonstrating regular spiking activity where the pre-ACh spiking frequency was $<$15 Hz. This provided 16 cell voltage traces. From these selected recordings it was determined that the regenerative release occurred within 200 ms of receptor activation as spikes were inhibited by this time. Both of these issues required the alteration of parameter values in order to achieve the desired responses.

In Falkenburger et al. (2010b), the authors used fluorescence resonance energy transfer (FRET) to measure the binding of M1 mAChRs to Oxo-M. A separate study performed an analysis of ACh binding to M1 using similar FRET techniques (Ziegler et al., 2011). From this study we took the half maximal effective concentration value of ACh and used that value to calibrate the parameters for agonist binding the receptor (see reaction 1 in **Table S2**). The change in the receptor's response to agonist concentration can be seen in **Figure S3**.

The discrepancy between the rapid release of intracellular calcium after M1 activation seen in CA1 pyramidal cells and the slow generation of IP_3 in the Kruse et al. (2016) model was solved by increasing a subset of kinetic parameters in two portions of the M1 model. This discrepancy is most likely due to the original model being constrained to fit the response within sympathetic neurons. Activation of M1 channels in this neuron type leads to PIP_2 depletion but not large releases of intracellular calcium. The rate of PLC activity will differ depending on the specific isozymes present within the cell type. Hippocampal cells contain PLC isozymes which are activated by increased intracellular calcium (Nakahara et al., 2005), creating a positive feedback for the hydrolysis of PIP_2. It stands to reason that PIP_2 hydrolysis would be triggered more rapidly than in the original model. The first portion of the model that needed faster dynamics was the activation and inactivation of PLC through its binding and unbinding to the G protein. Expressions (10) and (11) describe these reactions.

$$\text{PLC} + G\alpha - \text{GTP} \xrightarrow{k^{PLCassoc}} G\alpha - \text{GTP-PLC} \tag{10}$$

$$G\alpha - \text{GDP-PLC} \xrightarrow{k^{PLCdiss}} G\alpha - \text{GDP} + \text{PLC} \tag{11}$$

From these reactions we recalibrated the two forward rates ($k^{PLCassoc}$ and $k^{PLCdiss}$). If we examine the original dynamics as seen in **Figures 3A,B**, one can see that the PLC activation peaks around 2 s after the ACh pulse and that IP_3 levels peak around the same time. However, looking at the cell recordings (see **Figure 4A** for an example), by this time the calcium transients have already largely ended by 2 s as the cells have largely resumed spiking by then. Using the original parameter values led to a longer weak release of calcium from the ER as opposed to the approximately 1 s duration strong release we required to reach higher ($>1\ \mu$M) intracellular calcium concentrations. The two parameters were therefore both increased by a factor of 10. The difference in dynamics can be seen in **Figure 3A**.

The second portion of the M1 model that required changed kinetics was the hydrolysis of PIP_2 into DAG and IP_3. Expressions (12) and (13) describe this reaction.

$$\text{PIP2} \xrightarrow{k^{PLC}} \text{IP}_3 + \text{DAG} \tag{12}$$

$$k^{PLC} = r_{PLC} * \text{G}\alpha\text{-GTP-PLC} \tag{13}$$

Here the parameter, r_{PLC}, was increased by a factor of 100. If we look at **Figure 3B**, we can see how this altered the dynamics of

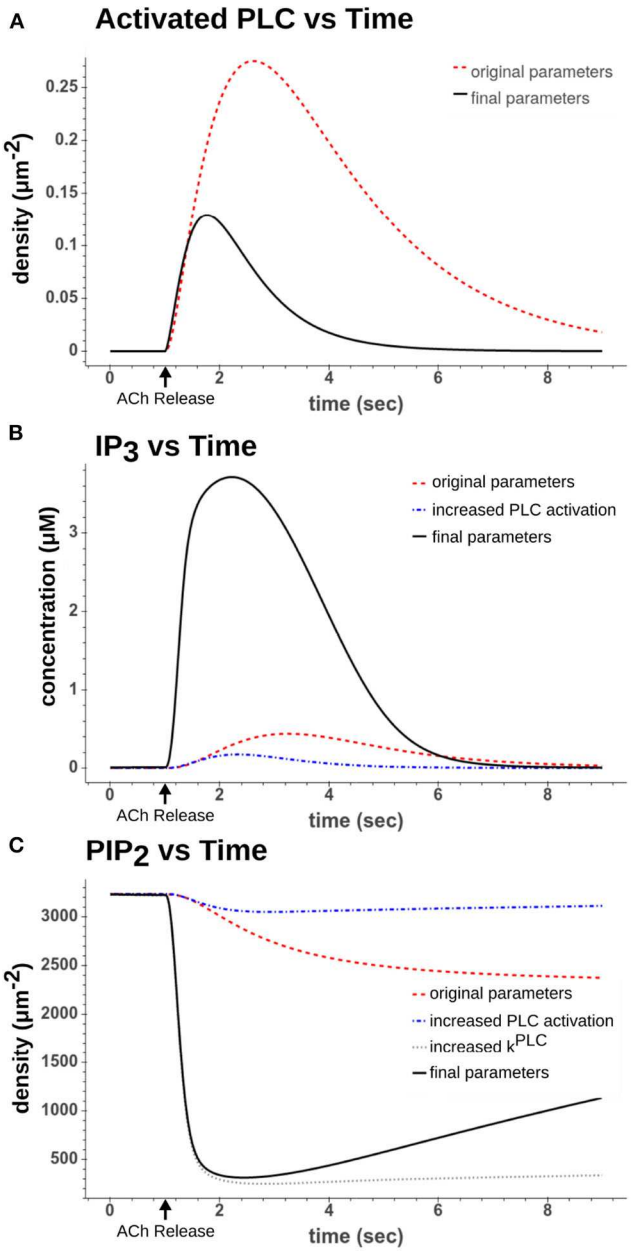

FIGURE 3 | Comparison of dynamics after a simulated 50 ms 100 μM ACh pulse. **(A)** Activated PLC dynamics using original parameter values from Kruse et al. (2016) (red dash) to recalibrated parameter values used in final model (black solid). **(B)** IP_3 dynamics using either the original parameters (red dash), increased PLC activation parameters (blue dash-dot), or increased PLC activation parameters along with increased hydrolysis rate (k^{PLC}) (black solid) **(C)** PIP_2 dynamics using either using either the original parameters (red dash), increased PLC activation parameters (blue dash-dot), increased PLC activation parameters along with increased hydrolysis rate (k^{PLC}) (gray dot), or all increased parameters including those that drive synthesis of PIP_2 (black solid).

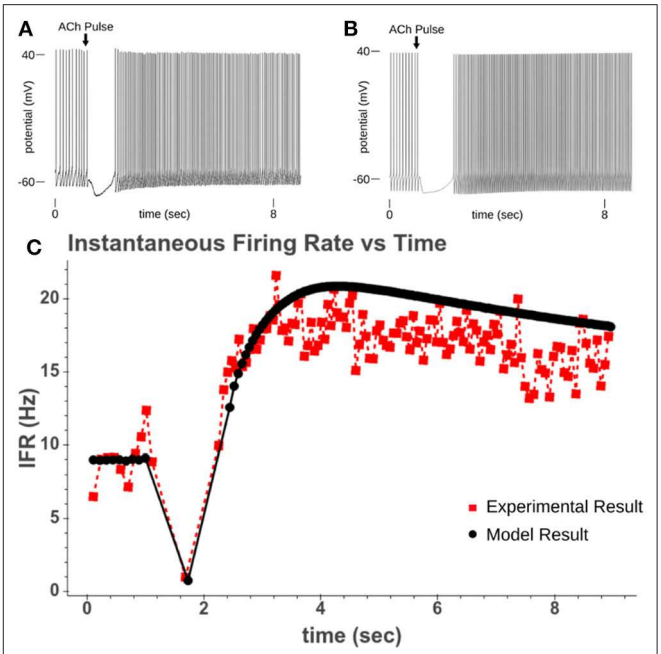

FIGURE 4 | (A) Experimental recording of CA1 pyramidal cell responding to a 40 ms pulse of 100 μM while driven to spiking. **(B)** Model response to 50 ms pulse of 100 μM ACh while cell is driven to regular spiking. **(C)** Comparison of instantaneous firing rate of experimental response and model response.

the reaction. By increasing the rate of the hydrolysis along with increasing the rates in Expressions (10) and (11), the production of IP_3 occurred far more rapidly and was largely complete within 2 s of the ACh pulse. This also rapidly depleted the PIP_2 as seen in **Figure 3C**.

The next goals were to calibrate the calcium release required for spike inhibition and the rate of PIP_2 synthesis. This latter process controls the rate of reactivation of the M-current, and thereby controls the duration of spike acceleration. With the rate of IP_3 increased in previous calibration steps, overall IP_3 levels were controlled by altering the total concentrations of IP_3 Kinase (IP3K) (Expressions 16, 17, and 18) and IP 5-phosphatase (IP5P) (Expressions 14, 15). Similar reasoning to the increased rate of IP_3 production drove tuning the rate of IP_3 removal. The concentration of IP_3 needed to return to near resting levels quickly enough that calcium release ended within seconds of the ACh pulse. This allowed mechanisms to restore calcium to the ER and also allowed the cell's activity to sharply transition from hyperpolarization to accelerated spiking.

$$\mathrm{IP_5P} + \mathrm{IP3} \underset{k_b^{ip5p}}{\overset{k_f^{ip5p}}{\rightleftharpoons}} \mathrm{IP_5P} - \mathrm{IP_3} \tag{14}$$

$$\mathrm{IP_5P} - \mathrm{IP_3} \xrightarrow{k^{ip2}} \mathrm{IP_5P} + \mathrm{IP_2} \tag{15}$$

$$\mathrm{IP_3K} + 2\mathrm{Ca}^{2+} \underset{k_b^{ip3k\ ca}}{\overset{k_f^{ip3k\ ca}}{\rightleftharpoons}} \mathrm{IP_3K} - 2\mathrm{Ca_2} \tag{16}$$

$$\mathrm{IP_3K} - \mathrm{Ca_2} + \mathrm{IP_3} \underset{k_b^{ip3k\ ip3}}{\overset{k_f^{ip3k\ ip3}}{\rightleftharpoons}} \mathrm{IP_3K} - 2\mathrm{Ca_2} - \mathrm{IP_3} \tag{17}$$

$$IP_3K - Ca_2 - IP_3 \xrightarrow{k^{ip4}} IP_3K - Ca_2 + IP_4 \quad (18)$$

The mechanism for calcium efflux through IP_3 receptors is described by Expressions (19) and (20). Here R_{Open} refers to the open state of the full kinetic model. For the full kinetic scheme of the IP_3 receptor model (see **Figure S2B**).

$$Ca_{ER}^{2+} \xrightarrow{k_f^{IP3R}} Ca_{cyt}^{2+} \quad (19)$$

$$k_f^{IP3R} = g_{IP3R} \cdot R_{Open} \quad (20)$$

By manipulating the maximum flux, g_{IP3R}, and the rate of IP_3 breakdown we were able to produce calcium transients with peaks reached >1 μM that resolved within the desired duration range (1–3 s).

The final set of parameters we altered to replicate the CA1 pyramidal cells' behavior were involved in the synthesis of PIP_2. The resynthesis of PIP_2 is due to activity in the ER that varies with the proteins a cell type expresses (Blunsom and Cockcroft, 2020). As these species are not characterized within the CA1 pyramidal cell, we chose to use the mechanisms present in the model and to calibrate the kinetic parameters to match the behavior seen in *in vitro* experiments. The following expressions describe these reactions.

$$PI \xrightarrow{k^{4K}} PI(4)P \quad (21)$$

$$PI(4)P \xrightarrow{k^{5K}} PIP_2 \quad (22)$$

$$PI(4)P \xrightarrow{k^{4P}} PI \quad (23)$$

$$PIP_2 \xrightarrow{k^{5P}} PI(4)P \quad (24)$$

From these expressions four parameters needed to be recalibrated (k^{4K}, k^{5K}, k^{4P}, and k^{5P}). These parameters were recalibrated based upon the rate the instantaneous firing rate (IFR) returned to its pre-ACh value. In our model, this increased spiking is due to the inhibition of the M-current following PIP_2 depletion. By simulating the original experiments used in Gulledge and Kawaguchi (2007), we could replicate the altered spiking behavior and calibrate the kinetic parameters for PIP_2 synthesis so that the resolution of spike acceleration matched the experimental results. **Figure 3C** shows how the recalibrated parameters changed the synthesis of PIP_2. **Figure 4** shows a simulated experiment along with an example of a cell recording and demonstrates the model's ability to replicate the changes to IFR over time.

2.4. Depolarization Activated Calcium Store Replenishment

The final mechanism we resolved to include in this model was the role of store operated calcium entry (SOCE). Briefly, this is a process by which the depletion of lumenal calcium causes the activation of calcium channels on the plasma membrane. These channels are positioned in membrane junctions or regions where the distance between the plasma and ER membrane is <100 nm. This allows the calcium that enters through SOCE to almost directly enter the ER without altering the overall intracellular calcium concentration. For a more in depth review of this process (see Majewski and Kuznicki, 2014). Interestingly this process is also dependent on depolarization of the cell (Dasari et al., 2017). Without depolarization, repeated phasic exposure fails to demonstrate repeated hyperpolarizing responses to intracellular calcium release.

While this process is well documented and the responsible actors have been partly identified, the kinetics of this process have not been quantified. Not including a mechanism to replicate SOCE would make it impossible to simulate network activity with synaptic release of ACh, as the cell model would only be able to respond to one release event. To overcome this limitation we included a mechanism that replicated the behavior of SOCE without explicitly modeling the underlying molecular events. This mechanism is based on a series of assumptions.

- Depolarization is required for activation (Dasari et al., 2017).
- Hyperpolarization does not cause a leak from intracellular stores.
- Ca_{ER} depletion is required for activation (Majewski and Kuznicki, 2014).
- The action of this mechanism bypasses the intracellular calcium concentration as calcium directly moves from the extracellular space to the ER lumen.

Expression (25) was the mechanism we used which fit the above criteria.

$$dCa_{ER}^{2+} = g_{SOCE} \cdot \ln(1 + e^{v_m - v_{init}}) \cdot e^{\frac{-(Ca_{ER} - Ca_d)}{k_{SOCE}}} \quad (25)$$

This mechanism avoids altering the intracellular calcium concentration by directly changing the value of Ca_{ER}. Through the use of a softplus function this mechanism will have minimum activation except when the cell's membrane potential is depolarized from it's resting value (v_{init}). Also as Ca_{ER} approaches its resting value, this mechanism deactivates, ensuring it is maximally activated after calcium store depletion. As seen in **Figure S4** this mechanism allows repeated hyperpolarizations following intracellular calcium release if the cell depolarizes, but intracellular calcium release cannot repeatedly occur if the cell maintains a near resting membrane potential. This replicates behavior seen in cortical pyramidal cells (Dasari et al., 2017).

3. RESULTS

3.1. Acetylcholine and Cell Excitability

With the compartmental model able to replicate experimental responses of CA1 cells, we sought to explore how variations in the concentration of ACh would alter the model's behavior. As our model only captures the modulation in the soma, axon, and apical trunk of the cell, we focused on simulating how ACh alters the cell's excitability. Experiments tend to use agonist concentrations that will drive a significant and unambiguous response. These experimental concentrations may not be biologically relevant, however. While some measurements of *in vivo* ACh concentrations have been made, these measurements were made under the assumption of volumetric transmission. Recent work, however, has demonstrated that cholinergic terminals

form synapses, undermining this volumetric assumption (Takács et al., 2018). It therefore remains unclear what concentrations muscarinic receptors see during cognition. This particular aspect will be further discussed later. To avoid testing an exact concentration profile we sought to explore how our model responds to a wide range of concentrations. We also explored how the model behaved differently under a synaptic release of ACh vs. a steady state exposure. Short pulses (50 ms) replicated a simultaneous synaptic release we termed "phasic," while long term (>5 s) exposures simulated a steady state exposure referred to as "tonic" exposure.

We first replicated phasic exposure while the cell is at a resting membrane potential. **Figure 5** demonstrates how the model replicates these changes to membrane potential. At 100 μM the model produces a −9.05 mV hyperpolarization followed by a depolarization of 1.98 mV. From **Figures 5C,D**, it is clear that the amplitudes of these reactions are highly concentration dependent. From **Figure 5D**, its clear that for concentrations <0.1 μM, the release of calcium from intracellular stores is negligible and consequently no hyperpolarization occurs. In **Figure 5C**, we see that the hyperpolarizing effect is induced at lower concentrations ($EC_{50} = 0.499\ \mu$M) than the depolarization ($EC_{50} = 1.95\ \mu$M). This suggests that for cells at rest, within a certain range of concentrations, short pulses of acetylcholine would only have an inhibitory effect on the cell without producing much excitatory modulation.

We then sought to explore how phasic ACh exposure would alter the spiking activity of the cell model. We ran a series of simulations in which current injections drove the cell to spike at a constant rate of 10 Hz. At the 1 s mark we then modeled the injection of a 50 ms pulse of ACh with each simulation having a different concentration. The results of these simulations can be seen in **Figure 6**. Looking at **Figure 6C**, we can see that once spiking resumes, the IFR increases rapidly reaching a peak around 2–3 s after the ACh pulse. The increased firing rate then slowly returned to its baseline rate over several seconds. We calculated the peak percent increase in IFR, or spike acceleration, for each simulated concentration. As shown in **Figure 6D**, this value formed a smooth sigmoidal curve when plotted against concentration. The duration of the spike inhibition, seen in **Figure 6E** varied in a way similar to the variation seen in the resting conditions, abruptly beginning at concentrations above 0.1 μM. That spike inhibition begins so abruptly means that for ACh concentrations of 0.1 μM or less the firing rate will have noticeably increased without a period of spike inhibition. Additionally, while the peak acceleration followed a sigmoidal curve, the longest spike inhibition occurred at 1 μM, with the duration of inhibition decreasing thereafter. If we examine **Figure 7**, we can see the cause of this nonlinearity. In **Figure 7A**, we can see that as the concentration of the ACh pulse increases, the peak concentration of the intracellular calcium transient increases. **Figure 7B** demonstrates that these peak values form a sigmoidal curve. However, while the peak values is increasing, the duration of the calcium transient is also decreasing. This is due to larger ACh pulses driving increased IP_3 production and thereby causing a more rapid depletion of ER calcium stores. As the stores are depleted,

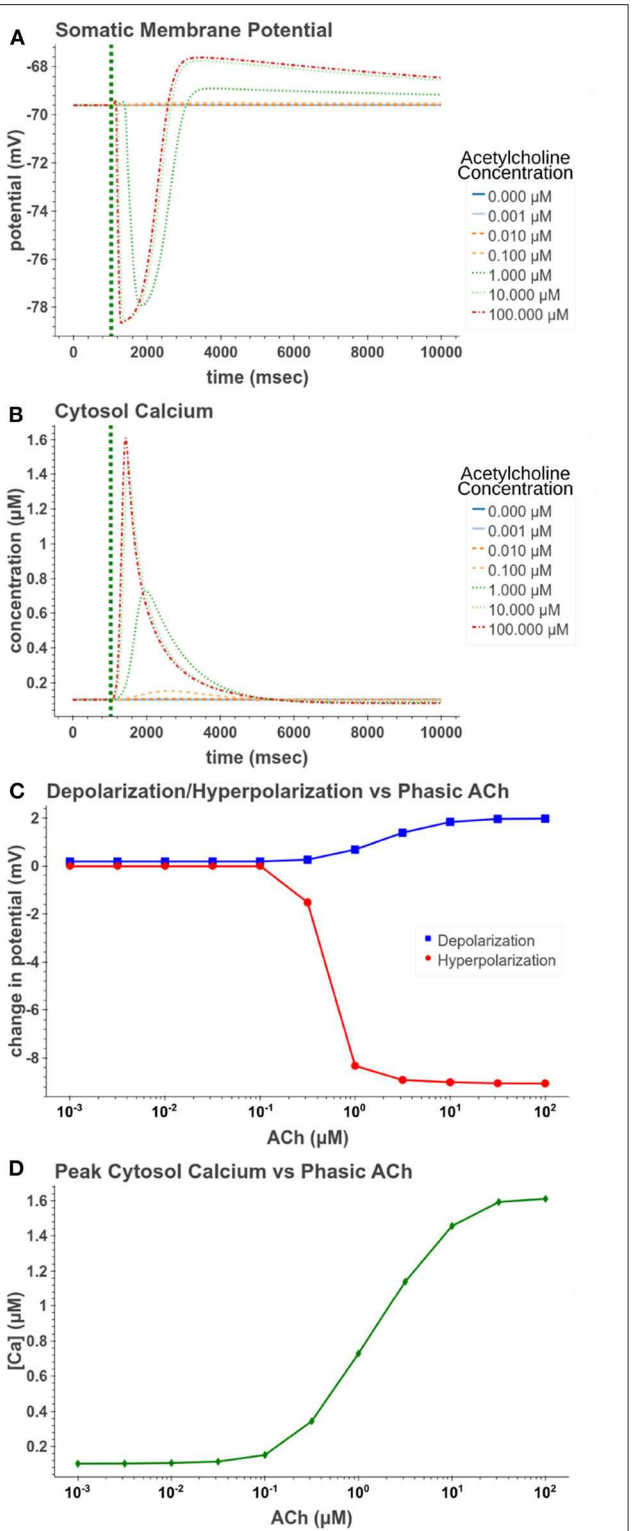

FIGURE 5 | Simulated response to phasic (50 ms) exposure at varying concentrations of acetylcholine. **(A)** Cell membrane potential with no other stimulation besides acetylcholine pulse. **(B)** Simulated intracellular calcium release. **(C)** Peak hyperpolarization and depolarization values at different concentrations of acetylcholine. **(D)** The peak intracellular calcium concentration following acetylcholine pulse.

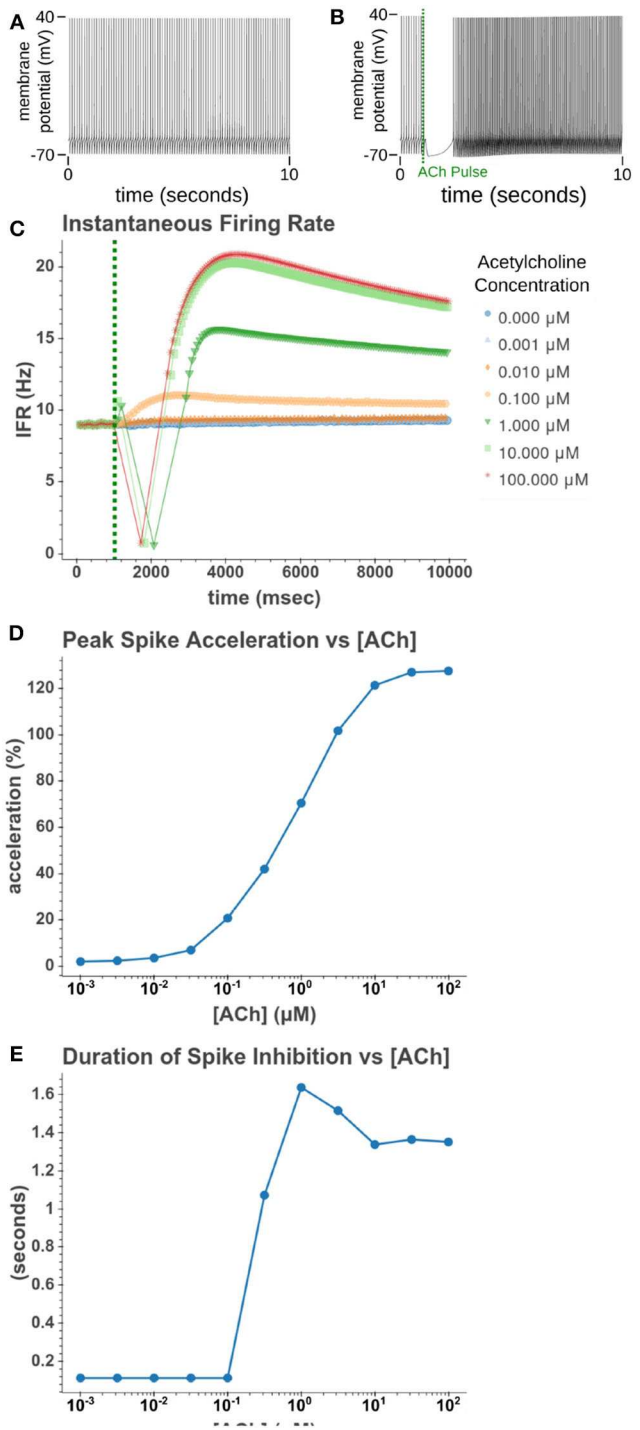

FIGURE 6 | Simulated response to phasic (50 ms) exposure of varying concentrations of acetylcholine. **(A)** Cell injected with a current amplitude such that it spikes at a steady rate of 10 Hz. Current injection is present throughout simulation. **(B)** Repeat of experiment with the addition of a 100 μM acetylcholine pulse that starts at $t = 1$ s and lasts for the duration of the simulation. **(C)** Instantaneous firing rate over time for different concentrations of ACh. This rate is the inverse of the inter spike interval. **(D)** The peak spike acceleration increased with higher concentrations of ACh. **(E)** The duration of the pause in spiking vs. the concentration of ACh.

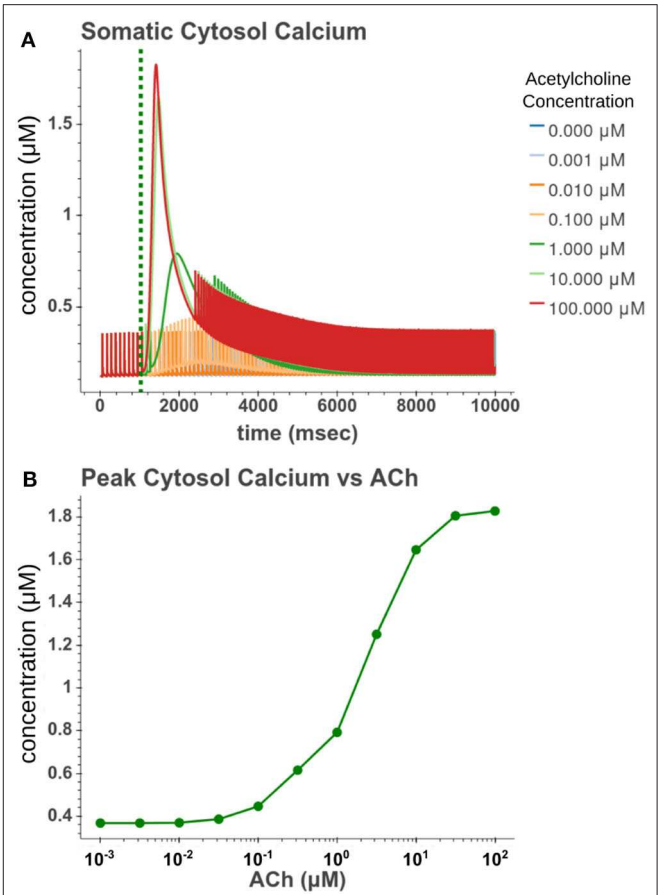

FIGURE 7 | (A) Time series of intracellular calcium concentration after phasic (50 ms) exposure to ACh. **(B)** Peak cytosol calcium vs. the concentration of the phasic ACh pulse.

the calcium transient begins to decay. It is this accelerated depletion of ER calcium which leads to the shorter duration of inhibition for higher concentrations of ACh. This nonlinearity in spike inhibition could have interesting implications for network activity, and will be a subject of discussion later in this paper.

Under tonic exposure to ACh, we noted multiple ways that the cell model displayed increased excitability. As can be seen in **Figure 8** the rheobase (defined here as the minimum amplitude of a 200 ms current pulse required to elicit an action potential) decreased with increasing concentrations of ACh. Starting at a value of 263 pA, the rheobase decreased 40.5% to a value of 156 pA with 14.3 nM of ACh producing half of the maximum decrease. The cell model also demonstrated increased excitability, illustrated by an increase of 39.1% in the input resistance measured at the soma. The increase in simulated input resistance can be see in **Figure 9**. This increased excitability plateaus in the high hundred nanomolar range with the increased excitability starting within nanomolar concentrations. This suggests that even relatively

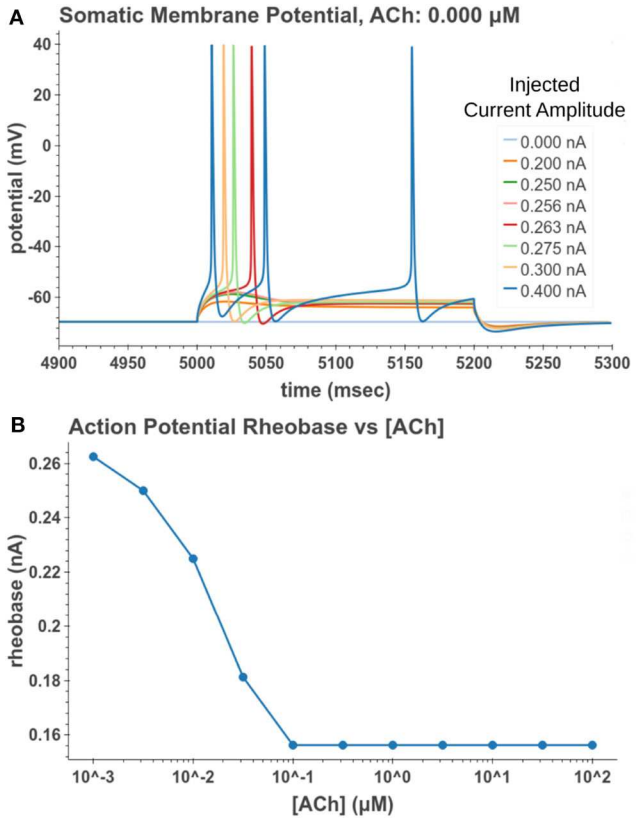

FIGURE 8 | Measuring reduction in rheobase due to tonic acetylcholine exposure. **(A)** A 200 ms current pulse of varying amplitude is applied at a time sufficiently after the start of a simulated acetylcholine exposure such that the system is at steady state. A binary search was performed to find the minimum current injection amplitude which would generate an action potential. **(B)** Cell rheobase decreases with increased concentration of acetylcholine.

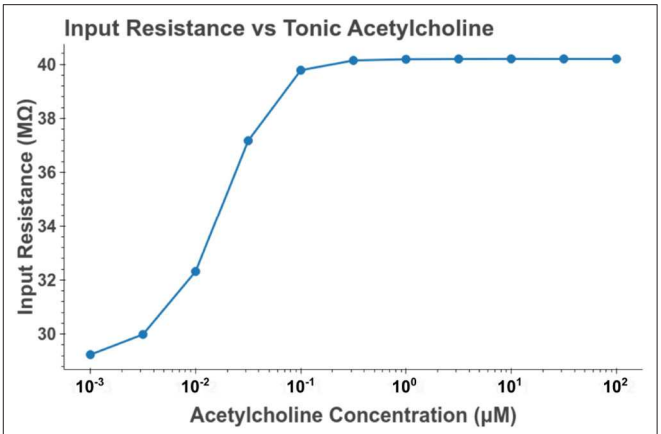

FIGURE 9 | Increasing the concentration of tonic acetylcholine increases the input resistance of the cell model as measured at the soma. Input resistance was measured by performing a series of somatic current injections and then performing linear regression on the relation between membrane depolarization to current amplitude. The values plotted are the slopes of the estimated linear functions. The current amplitudes used were 0, −100, and 100 pA.

low background concentrations should be able to alter cell spiking behavior. Long term exposure of ACh also produced a depolarization that persisted through the duration of exposure. The amplitude of this depolarization varied as a function of ACh concentration as demonstrated in **Figure 10**. Finally, tonic exposure caused accelerated spiking for a given amplitude of current injected at the soma. This accelerated spiking is demonstrated in **Figure 11**.

3.2. Intracellular Calcium Release

Focal application of muscarinic agonists and stimulation of cholinergic terminals were demonstrated to generate calcium waves that progressed from the apical dendritic trunk to the soma (Power and Sah, 2002). Our model, as demonstrated in **Figure 12**, replicates many aspects of these calcium waves. As seen experimentally, the sections in the apical trunk reached a higher peak calcium concentration more rapidly than the somatic section. This is likely due to dendritic regions having higher surface area to volume ratios.

4. DISCUSSION

4.1. Novel Additions to CA1 Compartmental Model

This model includes a number of mechanisms that have largely been absent from previous compartmental computational models of the CA1 pyramidal cell. In addition to including the M1 mAChR model, the intracellular calcium related mechanisms have been greatly expanded. Among these new calcium mechanisms were calbindin and PMCA. We have also been able to replicate the calcium wave phenomenon by including the endoplasmic reticulum. The parameters for these mechanisms were calibrated using experimental measurements obtained in CA1 pyramidal cells to ensure the resulting model accurately replicates this cell type's behavior. The addition of these novel mechanisms allows our model to replicate several molecular interactions that have been heretofore ignored in whole cell computational models of CA1 pyramidal cells.

4.2. Predictions From Model

The expanded CA1 pyramidal cell model have allowed us to generate some predictions which could be tested experimentally. First, the model predicts that intracellular calcium release can be triggered over a wide range of ACh concentrations. Later experimental evidence may show a more tightly regulated threshold that the transition between minimum and maximum responses occurs over a narrower range of concentrations. These hypothetical results would then suggest that there are mechanisms involved which introduce additional nonlinearities that increase the threshold for calcium release. For example the rate of PIP_2 hydrolysis into IP_3 being dependent on calcium would likely cause a sharper threshold for calcium waves.

A second prediction is that the duration of spike suppression as seen in **Figures 6E**, **11D** does not increase monotonically;

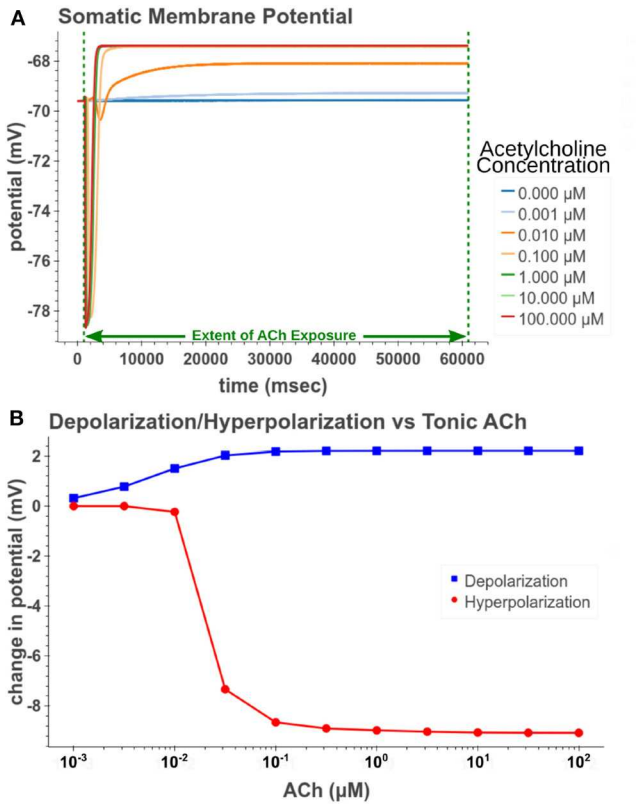

FIGURE 10 | Measuring depolarization and hyperpolarization after tonic (60 s) exposure to acetylcholine. **(A)** Simulated response of somatic membrane potential to different concentrations. **(B)** Amplitude of steady depolarization and temporary hyperpolarization vs. acetylcholine concentration.

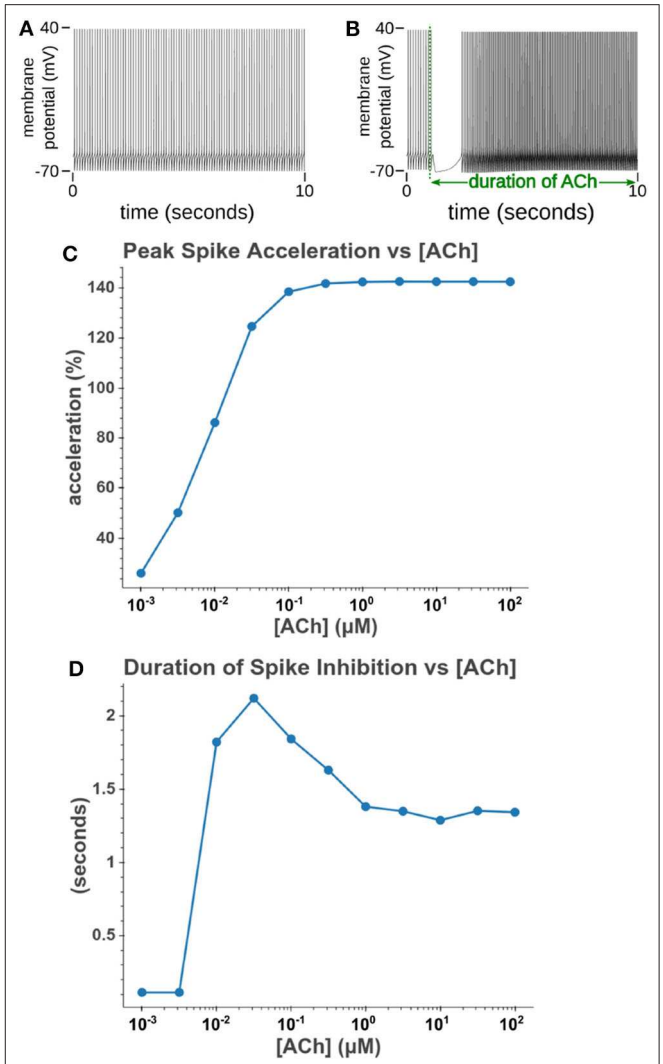

FIGURE 11 | Increasing the concentration of tonic acetylcholine increases spike rate for a given injected current amplitude. **(A)** Cell injected with a current amplitude such that it spikes at a steady rate of 10 Hz. Current injection is present throughout simulation. **(B)** Repeat of experiment with the addition of a 100 μM acetylcholine pulse that starts at $t = 1$ s and lasts for the duration of the simulation. **(C)** The maximum spike frequency acceleration vs. acetylcholine concentration. Spike frequency acceleration was measured as the percent increase from the rate before acetylcholine exposure. **(D)** Duration of spike inhibition vs. tonic acetylcholine concentration as measured as the longest inter spike interval after the initiation of the acetylcholine pulse.

instead the maximum duration occurs at intermediate concentrations (0.1 μM for tonic and 1 μM for phasic exposure). This result is likely due to the interplay between two competing processes: regenerative calcium release and the rate of calcium store depletion. The concentrations with the longest suppressions generate enough IP_3 to drive regenerative calcium release through IP_3Rs while minimizing the rate of Ca^{2+} release from intracellular stores. Higher ACh concentrations drive higher IP_3 production and so IP_3Rs open more fully and deplete intracellular stores more rapidly. This modulation of the length of inhibition is interesting when considering the possible functional roles calcium waves play in CA1 pyramidal cells.

If the function of calcium waves is to provide an inhibitory signal, then this inhibitory signal would have some interesting properties. First, as IP_3 is the trigger for this inhibition, multiple sources (whether mGluRs or mAChRs) could be required to work in concert to generate this signal. The non-monotonically increasing duration of spike cessation suggests that coactivation of additional IP_3 sources after the regenerative calcium release threshold has been passed may cause a shorter inhibition as the additional IP_3 will only lead to faster calcium stores depletion. Second, the rate at which intracellular calcium stores are depleted depends upon the amount of calcium stored. A cell with more calcium buffered in the ER will have a longer inhibitory reaction to cholinergic modulation. Since every action potential increases the amount of calcium in the ER, calcium waves would be longer for cells that have had more action potentials in the recent past. This mechanism would thereby act as an internal inhibition which encodes each cell's past activity. These two properties suggest scenarios where cholinergic modulation causes shorter inhibition for cells that are currently receiving a large glutamatergic signal but have not been spiking much in the past while cells that have been consistently spiking and only

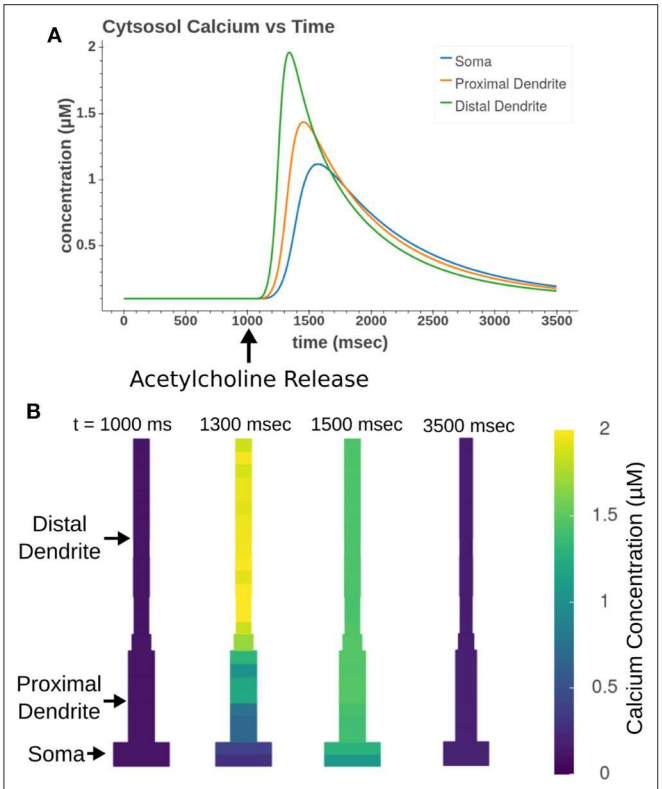

FIGURE 12 | Acetylcholine leads to the release of intracellular calcium. All sections in the apical trunk and soma were simultaneously exposed to a 50 ms pulse of 100 μM ACh. **(A)** Simulated time course of intracellular calcium after acetylcholine exposure. More distal sections achieve higher concentrations more rapidly than somatic sections. **(B)** Display of concentrations in all model compartments at different time steps throughout calcium wave.

receive cholinergic modulation are inhibited for a longer period. The implications of these properties may have a crucial impact on the downstream effects of calcium homeostasis, amongst which are excitotoxicity and learning and memory; these aspects will be studied in later work (see section 4.4 below).

4.3. Refining Model

In developing this model there have been gaps in experimental evidence which have made it difficult to model all of the experimental reactions to cholinergic modulation. First of all the dynamics of the phosphoinositides in the CA1 pyramidal cell plasma membranes are not well understood. This has forced us to make assumptions based on electrophysiological results, but further research into this area would aid in refining the model. As the ER plays a role in the production of these phospholipids, it is likely that the depletion of calcium stores leads to changed dynamics. Indeed experiments in CA1 pyramidal cells have suggested that prolonged activation of mAChRs can drive oscillations in PIP_2 levels (Hackelberg and Oliver, 2018). The signaling cascade that drives these oscillations, however, is not well-understood and therefore could not be included in this model iteration. As these dynamics become better understood, more explicit cascades can be incorporated into the model allowing for a better simulation of the depletion and synthesis of PIP_2.

Tonic cholinergic activation was also shown to inhibit the early portions of slow after hyperpolarization (sAHP) following trains of action potentials (Dasari and Gulledge, 2011). Experimental evidence suggests this sAHP is largely due to sodium-potassium exchange pumps (Gulledge et al., 2013; Tiwari et al., 2018). It is unclear how these exchange pumps interact with the mechanisms involved with mAChR activation. Without this clearer understanding, we have no way to properly calibrate the level of sAHP inhibition to variations in ACh concentration.

The model could also be expanded through adding mechanisms which model mitochondrial calcium dynamics. The mitochondria, along with being vital for the energy metabolism of the cell, play a large role in calcium dynamics through interactions with the ER (Krols et al., 2016). As mitochondrial dysfunction has a well established link with Alzheimer's disease (Cenini and Voos, 2019), this expansion would provide a method for exploring the functional consequences to network behavior and how best to intercede.

4.4. Future Uses

Though the present work is a significant advancement for modeling the interactions between cholinergic input, intracellular calcium, and neuronal dynamics, the model is far from encompassing all of the mechanisms that participate in cholinergic response. Yet this work represents a framework within which additional mechanisms can be added as the knowledge of the system evolves. We have sought to use best practices while generating the code base to facilitate its understanding and allow future users to expand upon its capabilities. The current cell model focused on cholinergic modulation in the apical dendritic trunk and the somatic region and consequently does not incorporate the modulation of synaptic transmission. Experimental evidence has shown that in synapses originating from the CA3 region, the activation of presynaptic M4 mAChRs suppresses the amplitude of excitatory postsynaptic potentials (EPSPs) (Dasari and Gulledge, 2011). This signal suppression has been suggested to shift control of CA1 pyramidal cell activity away from the CA3 toward synaptic inputs from the entorhinal cortex (EC). This is theorized to set the CA1 network into a state more conducive for encoding the sensory information encoded by the EC synapses (Hasselmo and McGaughy, 2004). However, the synaptic connections from the EC are located in the most distal portions of the CA1 pyramidal cell dendritic tree. In order for these inputs to become dominant, the CA1 pyramidal cell would need to become more sensitive to distal inputs. Our model has demonstrated that it is capable of replicating an increased excitability as measured by increased input resistance and lower rheobase at higher concentrations of ACh. This increased excitability, in conjunction with suppressed CA3 synaptic activity would replicate increased sensitivity to distal inputs. Our model thereby constitutes a solid foundation for future work exploring the consequences of this modulation for the integration of inputs from EC vs. CA3.

Additionally, synaptic connections from CA3 pyramidal cells to CA1 pyramidal cells demonstrate plasticity that is dependent

upon postsynaptic calcium concentrations (Cummings et al., 1996). M1 mAChRs are known to be present at these synapses (Yamasaki et al., 2010) and have been linked to both long-term potentiation and long-term depression (Dennis et al., 2016). Our lab has already developed a kinetic model for the postsynaptic calcium seen in the spine head (Hu et al., 2018). While this previous model did not include any mechanisms to link M1 activation to intracellular calcium release, these missing mechanisms could easily be added. This would allow us to expand upon previous modeling efforts that sought to tie calcium dynamics to plasticity (Shouval et al., 2002) to explore how cholinergic modulation alters the network dynamics through long term changes in connectivity.

Another direction of interest would consist of expanding the model to explicitly model cholinergic synapses. Currently our simulations treat acetylcholine concentration as a fixed value which we change in a step-wise manner. The addition of cholinergic synapses would allow us to explore how the model responds to varying synaptic parameters. For example, tonic ACh concentration is based upon both the amount of ACh released but also the rate of hydrolysis due to acetylcholinesterase (AChE). This enzyme is the target for a class of drugs, AChE inhibitors, used in the treatment of AD. Exploring how these drugs alter CA1 network dynamics could point to better treatment strategies. Additionally, cholinergic synapses have been shown to cotransmit ACh with GABA (Granger et al., 2016; Takács et al., 2018). This cotransmission could have dramatic effects on network coherence.

Finally, although pyramidal cells are the most numerous cell type in the CA1, they are not alone. There are a variety of interneuron cell types which are also the subject of cholinergic modulation. If ACh does play a role in shifting the focus of information processing from synapses from the CA3 to synapses from the EC, interneurons likely contribute to this process. This is due to certain interneurons' ability to disinhibit CA1 pyramidal cells (e.g., CCK+ Basket Cells; Karson et al., 2009). Furthermore, interneurons also participate in the generation of network oscillations which help organize network processes (e.g., OLM cells; Mikulovic et al., 2018). Understanding how these cells, through ACh modulation regulate the overall network activity would aid our understanding of the complex role ACh plays in the hippocampus.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://github.com/armerg/ca1_muscarinic_modulation/.

AUTHOR CONTRIBUTIONS

AM and J-MB: concept and design of study, analysis and/or interpretation of data, and drafting of the manuscript. AM: data acquisition. AM, J-MB, GY, and TB: critical revision and approval of the manuscript to be published. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by NIBIB grants U01 GM104604 and P41 EB001978-24.

ACKNOWLEDGMENTS

We would like to thank Dr. Allan Gulledge for providing cell recordings which aided in this project. Computation for the work described in this paper was supported by the University of Southern California's Center for Advanced Research Computing (https://carc.usc.edu/).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncom.2020.00075/full#supplementary-material

REFERENCES

Aznavour, N., Mechawar, N., and Descarries, L. (2002). Comparative analysis of cholinergic innervation in the dorsal hippocampus of adult mouse and rat: a quantitative immunocytochemical study. *Hippocampus* 12, 206–217. doi: 10.1002/hipo.1108

Baimbridge, K. G., Peet, M. J., McLennan, H., and Church, J. (1991). Bursting response to current-evoked depolarization in rat ca1 pyramidal neurons is correlated with Lucifer yellow dye coupling but not with the presence of calbindin-D28k. *Synapse* 7, 269–277. doi: 10.1002/syn.890070404

Blunsom, N. J., and Cockcroft, S. (2020). Phosphatidylinositol synthesis at the endoplasmic reticulum. *Biochim. Biophys. Acta* 1865:158471. doi: 10.1016/j.bbalip.2019.05.015

Brown, D. A., and Adams, P. R. (1980). Muscarinic suppression of a novel voltage-sensitive K^+ current in a vertebrate neurone. *Nature* 283, 673–676. doi: 10.1038/283673a0

Buchanan, K. A., Petrovic, M. M., Chamberlain, S. E., Marrion, N. V., and Mellor, J. R. (2010). Facilitation of long-term potentiation by muscarinic M1 receptors is mediated by inhibition of SK channels. *Neuron* 68, 948–963. doi: 10.1016/j.neuron.2010.11.018

Carnevale, N. T., and Hines, M. L. (2006). *The NEURON Book*. New York, NY: Cambridge University Press. doi: 10.1017/CBO9780511541612

Carver, C. M., and Shapiro, M. S. (2019). Gq-coupled muscarinic receptor enhancement of KCNQ2/3 channels and activation of TRPC channels in multimodal control of excitability in dentate gyrus granule cells. *J. Neurosci.* 39, 1566–1587. doi: 10.1523/JNEUROSCI.1781-18.2018

Cenini, G., and Voos, W. (2019). Mitochondria as potential targets in Alzheimer disease therapy: an update. *Front. Pharmacol.* 10:902. doi: 10.3389/fphar.2019.00902

Cummings, J. A., Mulkey, R. M., Nicoll, R. A., and Malenka, R. C. (1996). Ca^{2+} signaling requirements for long-term depression in the hippocampus. *Neuron* 16, 825–833. doi: 10.1016/S0896-6273(00)80102-6

Dannenberg, H., Young, K., and Hasselmo, M. (2017). Modulation of hippocampal circuits by muscarinic and nicotinic receptors. *Front. Neural Circuits* 11:102. doi: 10.3389/fncir.2017.00102

Dasari, S., and Gulledge, A. T. (2011). M1 and M4 receptors modulate hippocampal pyramidal neurons. *J. Neurophysiol.* 105, 779–792. doi: 10.1152/jn.00686.2010

Dasari, S., Hill, C., and Gulledge, A. T. (2017). A unifying hypothesis for M1 muscarinic receptor signalling in pyramidal neurons. *J. Physiol.* 595, 1711–1723. doi: 10.1113/JP273627

Dennis, S. H., Pasqui, F., Colvin, E. M., Sanger, H., Mogg, A. J., Felder, C. C., et al. (2016). Activation of muscarinic M1 acetylcholine receptors induces long-term potentiation in the hippocampus. *Cereb. Cortex* 26, 414–426. doi: 10.1093/cercor/bhv227

Doi, T., Kuroda, S., Michikawa, T., and Kawato, M. (2005). Inositol 1,4,5-trisphosphate-dependent Ca^{2+} threshold dynamics detect spike timing in cerebellar purkinje cells. *J. Neurosci.* 25, 950–961. doi: 10.1523/JNEUROSCI.2727-04.2005

Falkenburger, B. H., Dickson, E. J., and Hille, B. (2013). Quantitative properties and receptor reserve of the DAG and PKC branch of Gq-coupled receptor signaling. *J. Gen. Physiol.* 141, 537–555. doi: 10.1085/jgp.201210887

Falkenburger, B. H., Jensen, J. B., and Hille, B. (2010a). Kinetics of M1 muscarinic receptor and G protein signaling to phospholipase C in living cells. *J. Gen. Physiol.* 135, 81–97. doi: 10.1085/jgp.200910344

Falkenburger, B. H., Jensen, J. B., and Hille, B. (2010b). Kinetics of PIP2 metabolism and KCNQ2/3 channel regulation studied with a voltage-sensitive phosphatase in living cells. *J. Gen. Physiol.* 135, 99–114. doi: 10.1085/jgp.200910345

Gant, J. C., Sama, M. M., Landfield, P. W., and Thibault, O. (2006). Early and simultaneous emergence of multiple hippocampal biomarkers of aging is mediated by Ca^{2+}-induced Ca^{2+} release. *J. Neurosci.* 26, 3482–3490. doi: 10.1523/JNEUROSCI.4171-05.2006

Garaschuk, O., Yaari, Y., and Konnerth, A. (1997). Release and sequestration of calcium by ryanodine-sensitive stores in rat hippocampal neurones. *J. Physiol.* 502, 13–30. doi: 10.1111/j.1469-7793.1997.013bl.x

Granger, A. J., Mulder, N., Saunders, A., and Sabatini, B. L. (2016). Cotransmission of acetylcholine and GABA. *Neuropharmacology* 100, 40–46. doi: 10.1016/j.neuropharm.2015.07.031

Gulledge, A. T., Dasari, S., Onoue, K., Stephens, E. K., Hasse, J. M., and Avesar, D. (2013). A sodium-pump-mediated afterhyperpolarization in pyramidal neurons. *J. Neurosci.* 33, 13025–13041. doi: 10.1523/JNEUROSCI.0220-13.2013

Gulledge, A. T., and Kawaguchi, Y. (2007). Phasic cholinergic signaling in the hippocampus: functional homology with the neocortex? *Hippocampus* 17, 327–332. doi: 10.1002/hipo.20279

Hackelberg, S., and Oliver, D. (2018). Metabotropic acetylcholine and glutamate receptors mediate PI(4,5)P2 depletion and oscillations in hippocampal CA1 pyramidal neurons *in situ*. *Sci. Rep.* 8, 1–15. doi: 10.1038/s41598-018-31322-8

Hasselmo, M. E., Bodelón, C., and Wyble, B. P. (2002). A proposed function for hippocampal theta rhythm: separate phases of encoding and retrieval enhance reversal of prior learning. *Neural Comput.* 14, 793–817. doi: 10.1162/089976602317318965

Hasselmo, M. E., and McGaughy, J. (2004). High acetylcholine levels set circuit dynamics for attention and encoding and low acetylcholine levels set dynamics for consolidation. *Prog. Brain Res.* 145, 207–231. doi: 10.1016/S0079-6123(03)45015-2

Hendrickson, P. J., Yu, G. J., Song, D., and Berger, T. W. (2015). Interactions between inhibitory interneurons and excitatory associational circuitry in determining spatio-temporal dynamics of hippocampal dentate granule cells: a large-scale computational study. *Front. Syst. Neurosci.* 9:155. doi: 10.3389/fnsys.2015.00155

Higley, M. J., and Picciotto, M. R. (2014). Neuromodulation by acetylcholine: examples from schizophrenia and depression. *Curr. Opin. Neurobiol.* 29, 88–95. doi: 10.1016/j.conb.2014.06.004

Hu, E., Mergenthal, A., Bingham, C. S., Song, D., Bouteiller, J. M., and Berger, T. W. (2018). A glutamatergic spine model to enable multi-scale modeling of nonlinear calcium dynamics. *Front. Comput. Neurosci.* 12:58. doi: 10.3389/fncom.2018.00058

Jiang, L., Kundu, S., Lederman, J. D., López-Hernández, G. Y., Ballinger, E. C., Wang, S., et al. (2016). Cholinergic signaling controls conditioned fear behaviors and enhances plasticity of cortical-amygdala circuits. *Neuron* 90, 1057–1070. doi: 10.1016/j.neuron.2016.04.028

Kalappa, B. I., Gusev, A. G., and Uteshev, V. V. (2010). Activation of functional α7-containing nAChRs in hippocampal CA1 pyramidal neurons by physiological levels of choline in the presence of PNU-120596. *PLoS ONE* 5:e13964. doi: 10.1371/journal.pone.0013964

Karson, M. A., Tang, A.-H., Milner, T. A., and Alger, B. E. (2009). Synaptic cross talk between perisomatic-targeting interneuron classes expressing cholecystokinin and parvalbumin in hippocampus. *J. Neurosci.* 29, 4140–4154. doi: 10.1523/JNEUROSCI.5264-08.2009

Kramis, R., Vanderwolf, C. H., and Bland, B. H. (1975). Two types of hippocampal rhythmical slow activity in both the rabbit and the rat: relations to behavior and effects of atropine, diethyl ether, urethane, and pentobarbital. *Exp. Neurol.* 49, 58–85. doi: 10.1016/0014-4886(75)90195-8

Krols, M., Bultynck, G., and Janssens, S. (2016). ER-Mitochondria contact sites: a new regulator of cellular calcium flux comes into play. *J. Cell Biol.* 214, 367–370. doi: 10.1083/jcb.201607124

Kruse, M., Vivas, O., Traynor-Kaplan, A., and Hille, B. (2016). Dynamics of phosphoinositide-dependent signaling in sympathetic neurons. *J. Neurosci.* 36, 1386–1400. doi: 10.1523/JNEUROSCI.3535-15.2016

Majewski, L., and Kuznicki, J. (2014). SOCE in neurons: signaling or just refilling? *Biochim. Biophys. Acta* 1853, 1940–1952. doi: 10.1016/j.bbamcr.2015.01.019

McDougal, R. A., Hines, M. L., and Lytton, W. W. (2013). Reaction-diffusion in the NEURON simulator. *Front. Neuroinform.* 7:28. doi: 10.3389/fninf.2013.00028

McQuiston, A. R. (2014). Acetylcholine release and inhibitory interneuron activity in hippocampal CA1. *Front. Synapt. Neurosci.* 6:20. doi: 10.3389/fnsyn.2014.00020

Migliore, R., Lupascu, C. A., Bologna, L. L., Romani, A., Courcol, J.-D., Antonel, S., et al. (2018). The physiological variability of channel density in hippocampal CA1 pyramidal cells and interneurons explored using a unified data- driven modeling workflow. *PLoS Comput. Biol.* 14:e1006423. doi: 10.1371/journal.pcbi.1006423

Mikulovic, S., Restrepo, C. E., Siwani, S., Bauer, P., Pupe, S., Tort, A. B., et al. (2018). Ventral hippocampal OLM cells control type 2 theta oscillations and response to predator odor. *Nat. Commun.* 9:3638. doi: 10.1038/s41467-018-05907-w

Molecular Probes (2005). *Long-Wavelength Calcium Indicators*. Available online at: https://assets.thermofisher.com/TFS-Assets/LSG/manuals/mp03010.pdf

Müller, A., Kukley, M., Stausberg, P., Beck, H., Müller, W., and Dietrich, D. (2005). Endogenous Ca^{2+} buffer concentration and Ca^{2+} microdomains in hippocampal neurons. *J. Neurosci.* 25, 558–565. doi: 10.1523/JNEUROSCI.3799-04.2005

Nakahara, M., Shimozawa, M., Nakamura, Y., Irino, Y., Morita, M., Kudo, Y., et al. (2005). A novel phospholipase C, PLCη2, is a neuron-specific isozyme. *J. Biol. Chem.* 280, 29128–29134. doi: 10.1074/jbc.M503817200

Oh, M. M., Oliveira, F. A., Waters, J., and Disterhoft, J. F. (2013). Altered calcium metabolism in aging CA1 hippocampal pyramidal neurons. *J. Neurosci.* 33, 7905–7911. doi: 10.1523/JNEUROSCI.5457-12.2013

Power, J. M., and Sah, P. (2002). Nuclear calcium signaling evoked by cholinergic stimulation in hippocampal CA1 pyramidal neurons. *J. Neurosci.* 22, 3454–3462. doi: 10.1523/JNEUROSCI.22-09-03454.2002

Reynolds, C. F., Butters, M. A., Lopez, O., Pollock, B. G., Dew, M. A., Mulsant, B. H., et al. (2011). Maintenance treatment of depression in old age: a randomized, double-blind, placebo-controlled evaluation of the efficacy and safety of donepezil combined with antidepressant pharmacotherapy. *Arch. Gen. Psychiatry* 68, 51–60. doi: 10.1001/archgenpsychiatry.2010.184

Schliebs, R., and Arendt, T. (2011). The cholinergic system in aging and neuronal degeneration. *Behav. Brain Res.* 221, 555–563. doi: 10.1016/j.bbr.2010.11.058

Schmidt, H., Schwaller, B., and Eilers, J. (2005). Calbindin D28k targets myo-inositol monophosphatase in spines and dendrites of cerebellar Purkinje neurons. *Proc. Natl. Acad. Sci. U.S.A.* 102, 5850–5855. doi: 10.1073/pnas.0407855102

Shah, M. M., Migliore, M., Valencia, I., Cooper, E. C., and Brown, D. A. (2008). Functional significance of axonal Kv7 channels in hippocampal pyramidal neurons. *Proc. Natl. Acad. Sci. U.S.A.* 105, 7869–7874. doi: 10.1073/pnas.0802805105

Shouval, H. Z., Castellani, G. C., Blais, B. S., Yeung, L. C., and Cooper, L. N. (2002). Converging evidence for a simplified biophysical model of synaptic plasticity. *Biol. Cybern.* 87, 383–391. doi: 10.1007/s00422-002-0362-x

Solari, N., and Hangya, B. (2018). Cholinergic modulation of spatial learning, memory and navigation. *Eur. J. Neurosci.* 48, 2199–2230. doi: 10.1111/ejn.14089

Solovyova, N., Veselovsky, N., Toescu, E. C., and Verkhratsky, A. (2002). Ca^{2+} dynamics in the lumen of the endoplasmic reticulum in sensory neurons: direct visualization of Ca^{2+}-induced Ca^{2+} release triggered by physiological Ca^{2+} entry. *Eur. Mol. Biol. Organ. J.* 21, 622–630. doi: 10.1093/emboj/21.4.622

Spacek, J., and Harris, K. M. (1997). Three-dimensional organization of smooth endoplasmic reticulum in hippocampal CA1 dendrites and dendritic spines of the immature and mature rat. *Neuroscience* 17, 190–203. doi: 10.1523/JNEUROSCI.17-01-00190.1997

Takács, V. T., Cserép, C., Schlingloff, D., Pósfai, B., Sznyi, A., Sos, K. E., et al. (2018). Co-transmission of acetylcholine and GABA regulates hippocampal states. *Nat. Commun*. 9:2848. doi: 10.1038/s41467-018-05136-1

Thomas, D., Tovey, S. C., Collins, T. J., Bootman, M. D., Berridge, M. J., and Lipp, P. (2000). A comparison of fluorescent Ca^{2+} indicator properties and their use in measuring elementary and global Ca2+ signals. *Cell Calcium* 28, 213–223. doi: 10.1054/ceca.2000.0152

Tiwari, M. N., Mohan, S., Biala, Y., and Yaari, Y. (2018). Differential contributions of Ca^{2+}-activated K+ channels and Na^{+}/K^{+}-ATPases to the generation of the slow after hyperpolarization in CA1 pyramidal cells. *Hippocampus* 28, 338–357. doi: 10.1002/hipo.22836

Yamasaki, M., Matsui, M., and Watanabe, M. (2010). Preferential localization of muscarinic M1 receptor on dendritic shaft and spine of cortical pyramidal cells and its anatomical evidence for volume transmission. *J. Neurosci.* 30, 4408–4418. doi: 10.1523/JNEUROSCI.5719-09.2010

Ziegler, N., Bätz, J., Zabel, U., Lohse, M. J., and Hoffmann, C. (2011). FRET-based sensors for the human M1-, M3-, and M5-acetylcholine receptors. *Bioorgan. Med. Chem*. 19, 1048–1054. doi: 10.1016/j.bmc.2010.07.060

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.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 10 September 2020
doi: 10.3389/fncom.2020.00081

Assessing the Impact of I_h Conductance on Cross-Frequency Coupling in Model Pyramidal Neurons

Melvin A. Felton Jr.[1]*, *Alfred B. Yu*[2], *David L. Boothe*[2], *Kelvin S. Oie*[2] and *Piotr J. Franaszczuk*[2,3]

[1] Combat Capabilities Development Command (CCDC)—Army Research Laboratory, Adelphi, MD, United States, [2] Combat Capabilities Development Command (CCDC)—Army Research Laboratory, Aberdeen Proving Ground, MD, United States, [3] Department of Neurology, Johns Hopkins University School of Medicine, Baltimore, MD, United States

Large cortical and hippocampal pyramidal neurons are elements of neuronal circuitry that have been implicated in cross-frequency coupling (CFC) during cognitive tasks. We investigate potential mechanisms for CFC within these neurons by examining the role that the hyperpolarization-activated mixed cation current (I_h) plays in modulating CFC characteristics in multicompartment neuronal models. We quantify CFC along the soma-apical dendrite axis and tuft of three models configured to have different spatial distributions of I_h conductance density: (1) exponential gradient along the soma-apical dendrite axis, (2) uniform distribution, and (3) no I_h conductance. We simulated two current injection scenarios: distal apical 4 Hz modulation and perisomatic 4 Hz modulation, each with perisomatic, mid-apical, and distal apical 40 Hz injections. We used two metrics to quantify CFC strength—modulation index and height ratio—and we analyzed CFC phase properties. For all models, CFC was strongest in distal apical regions when the 40 Hz injection occurred near the soma and the 4 Hz modulation occurred in distal apical dendrite. The strongest CFC values were observed in the model with uniformly distributed I_h conductance density, but when the exponential gradient in I_h conductance density was added, CFC strength decreased by almost 50%. When I_h was in the model, regions with much larger membrane potential fluctuations at 4 Hz than at 40 Hz had stronger CFC. Excluding the I_h conductance from the model resulted in CFC either reduced or comparable in strength relative to the model with the exponential gradient in I_h conductance. The I_h conductance also imposed order on the phase characteristics of CFC such that minimum (maximum) amplitude 40 Hz membrane potential oscillations occurred during I_h conductance deactivation (activation). On the other hand, when there was no I_h conductance, phase relationships between minimum and maximum 40 Hz oscillation often inverted and occurred much closer together. This analysis can help experimentalists discriminate between CFC that

OPEN ACCESS

Edited by:
Christoph Metzner, Technical University of Berlin, Germany

Reviewed by:
Athanasia Papoutsi, Foundation of Research and Technology-Hellas (FORTH), Greece
Frances K. Skinner, University Health Network & University of Toronto, Canada

***Correspondence:**
Melvin A. Felton Jr.
melvin.a.felton.civ@mail.mil

Received: *27 April 2020*
Accepted: *31 July 2020*
Published: *10 September 2020*

Citation:
Felton MA Jr, Yu AB, Boothe DL, Oie KS and Franaszczuk PJ (2020) Assessing the Impact of I_h Conductance on Cross-Frequency Coupling in Model Pyramidal Neurons. Front. Comput. Neurosci. 14:81. doi: 10.3389/fncom.2020.00081

originates from different underlying physiological mechanisms and can help illuminate the reasons why there are differences between CFC strength observed in different regions of the brain and between different populations of neurons based on the configuration of the I_h conductance.

Keywords: cross-frequency coupling, phase-amplitude coupling, theta-gamma coupling, pyramidal neuron, I_h conductance

INTRODUCTION

Cross-frequency coupling (CFC) has been associated with mental processes like perceptual and memory-related tasks, and is often observed via electroencephalogram (EEG) and local field potential (LFP) measurements (Jensen and Colgin, 2007; Tort et al., 2009; Canolty and Knight, 2010; Lisman and Jenson, 2013). Different types of network properties can yield distinct CFC signatures [see Hyafil et al. (2015) for a review], and there are a variety of physiological mechanisms believed to contribute to CFC in pyramidal neurons, such as the timing of upstream inputs (Fernández-Ruiz et al., 2017), or fast synaptic inhibition (Wuff et al., 2009), and NMDA-mediated excitation of related interneuronal populations (Korotkova et al., 2010). However, the detailed nature of the role that intracellular mechanisms play in CFC of pyramidal neurons is still uncertain.

The way individual pyramidal neurons, and therefore the neuronal networks they are a part of, respond to synaptic input in particular frequency ranges is largely governed by ion channels (Lai and Jan, 2006; Nusser, 2012). In particular, the hyperpolarization-activated mixed cation current (I_h) plays a multitude of roles in the regulation of neuronal and network excitability impacting both membrane resting potential and rhythmic activity, as well as the magnitude of excitatory post-synaptic potentials (EPSPs) (Nusser, 2009; Brennan et al., 2016). The different effects of the I_h conductance suggest that it may play an important role in the occurrence of CFC within individual pyramidal neurons, however this remains uncertain. Vaidya and Johnston (2013) report gamma-theta correlation of synaptic currents observed in single CA1 pyramidal neurons that is distinct from the phenomenon of cross-frequency phase coupling observed at the network level. In addition, using a model of hippocampal CA3, Neymotin et al. (2013) showed that the density of pyramidal neuron I_h conductance modulated the amplitude of CFC observed in the simulated LFP (network-level CFC). Because there are differences in the subcellular distribution of I_h in distinct classes of pyramidal neurons (Bullis et al., 2006; Nusser, 2009), I_h may have a variable impact on CFC depending upon factors like regional specialization or underlying pathological conditions.

As an extension of our previous work examining coupling between perisomatic and distal apical functional zones in cortical layer five pyramidal neurons (Felton et al., 2018), we examined CFC along the soma-apical dendrite axis and tuft of realistic compartmental models of large pyramidal neurons. In particular, our baseline model neuron possessed an exponentially-increasing gradient of I_h conductance density along the apical dendrite like cortical layer five and hippocampal CA1 pyramidal neurons are known to possess (Hu et al., 2009; Nusser, 2009; Hay et al., 2011). To assess the role of the I_h conductance in the occurrence of CFC within neocortical and limbic pyramidal neurons, we configured a total of three models, each with the same morphology, but with different spatial distributions of I_h conductance density. We also examine the effect on CFC of fundamentally different input configurations to large pyramidal neurons. We simulated two modulation scenarios, one based on distal 4 Hz modulation and the other on perisomatic 4 Hz modulation. For each of these modulation types, 40 Hz current injections were simulated in key locations throughout the apical dendrite, namely in perisomatic, middle apical, and distal apical compartments. We used two metrics to quantify the strength of CFC—modulation index and height ratio (Tort et al., 2010)—and we analyzed the phase properties of CFC.

METHODS

Models

For this study, we adapted the model of a large cortical pyramidal neuron used in Felton et al. (2018). This model was obtained by modifying the regular-spiking, layer five pyramidal neuron model used by Traub et al. (2005), so that it possessed characteristics now known to be common among this class of neuron. The characteristics of this model included an exponentially increasing gradient of the I_h conductance density ascending along the soma-apical dendrite axis, and a distal apical Ca^{2+} hot zone where the conductance densities for the high-threshold and low-threshold Ca^{2+} channels are 10 and 100 times higher, respectively, than anywhere else in the apical dendrite and tuft (Hay et al., 2011). However, our focus in the current study is primarily on the impact of I_h on CFC. Therefore, the conductance densities for the two Ca^{2+} conductances were reverted back to Traub et al. (2003, 2005).

To evaluate the impact of the I_h conductance on CFC, we evaluated three models distinguished by the configuration of the I_h conductance throughout the neuronal membrane. The first model had a somato-apical dendritic exponential gradient in the I_h conductance common to many pyramidal neurons in layer five of the cortex or in the hippocampus (Hu et al., 2009; Nusser, 2009; Hay et al., 2011). The second model had almost an entirely uniform I_h conductance density distribution which can be found in various neuron types throughout the cortex, and in some cases, a uniformly-distributed I_h conductance can underlie pathological conditions like epilepsy (Nusser, 2009; Brennan et al., 2016). The third model did not have an I_h conductance, which may not be common at all in the cortex or hippocampus for large pyramidal neurons, but can be induced with the use

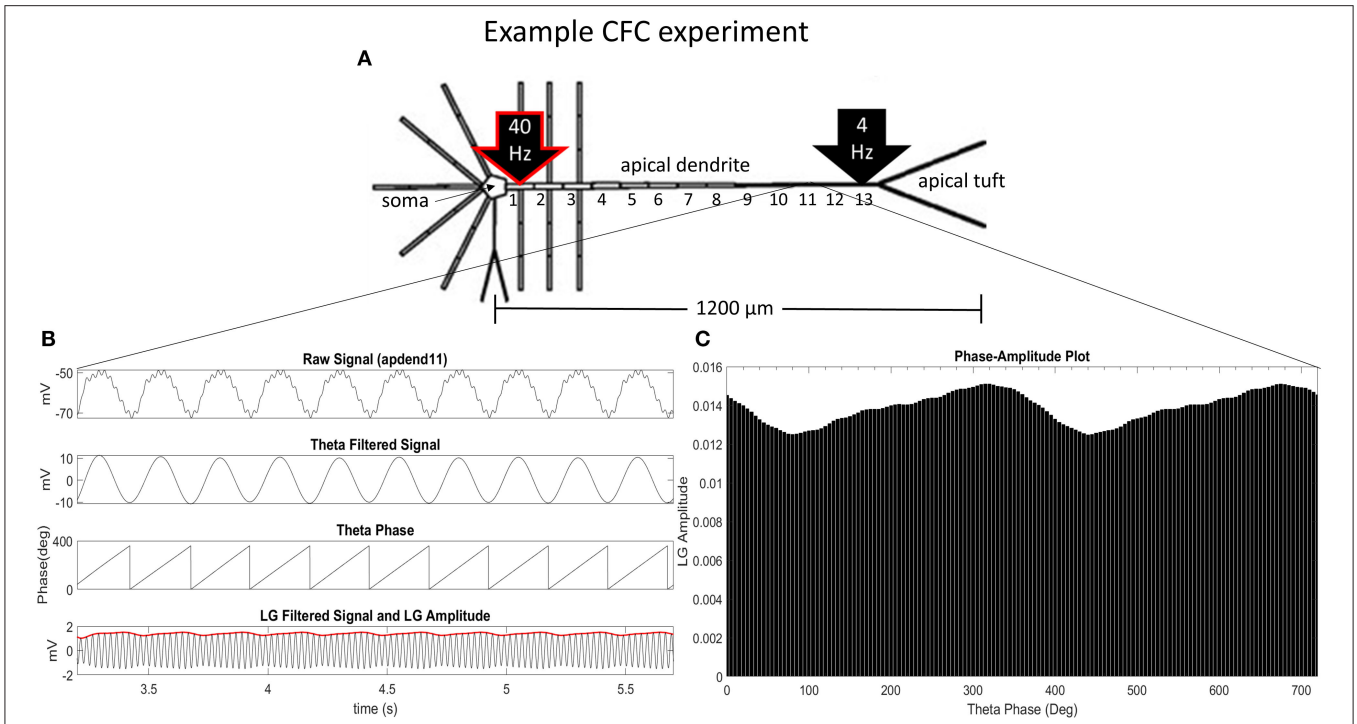

FIGURE 1 | CFC analysis performed on model pyramidal neuron. **(A)** A 1.5 nA 40 Hz sinusoidal current was injected into the base of the apical dendrite while a 1.5 nA 4 Hz sinusoidal current was injected into the distal apical dendrite. Additonal labels include: soma, apical dendrite, and apical tuft; apical dendrite compartments 1–13 (apdend1, apdend2, …, apdend13); and the distance from soma to end of apical tuft−1,200 µm. **(B)** From top to bottom: membrane potential oscillations in the 11th distal apical dendrite compartment (apdend11) resulting from the 40 Hz and 4 Hz current injections, filtered 4 Hz component of the membrane potential oscillation, time series of 4 Hz phase, and filtered 40 Hz component of membrane potential oscillation (black) and amplitude envelope of 40 Hz oscillation (red). **(C)** Phase-amplitude plot for two complete 4 Hz cycles with a phase bin size of 5° (LG–low gamma, ∼40 Hz).

of an I_h channel blocker, such as 4-(N-ethyl-N-phenylamino)-1, 2-dimethyl-6-(methylamino) pyrimidinium chloride (ZD7288) (Zhang et al., 2016). See **Table 1** for detailed ionic conductance density configuration for each of these three models.

CFC Analysis

Our CFC quantification analysis was based on the work of Tort et al. (2010). In our study, we focus on theta-gamma phase-amplitude coupling within large pyramidal neurons, which is commonly observed in cortex and hippocampus (Tort et al., 2010). **Figure 1A** illustrates an example simulation scenario where CFC was induced by injecting a 40 Hz sinusoidal current at the base of the apical dendrite while a 4 Hz sinusoidal current was injected in the distal apical dendrite. **Figures 1B,C** show several diagnostic plots that result from the CFC quantification analysis, in this case, for the 11th apical dendrite compartment from the soma (distal).

Following Tort et al. (2010), we obtain the diagnostic plots in **Figure 1B** that, from top to bottom, show the unfiltered membrane potential oscillations for the compartment that result from the two current injections, the filtered 4 Hz membrane potential oscillation for the compartment, the time series of the 4 Hz oscillation phase for the compartment, and the amplitude envelope (red) and filtered 40 Hz oscillation for the compartment (black). In this example, the simulation was run on the model with an exponential gradient in I_h conductance density.

To obtain the phase-amplitude plot from which the two CFC metrics used in this analysis are calculated, we first bin the phases of 4 Hz oscillations over the time period of analysis, and then calculate the mean of the 40 Hz oscillation amplitude envelope within each phase bin, j, denoted by $\langle A_{40} \rangle (j)$. After normalizing by the sum over all phase bins (N), we obtain the following expression for the normalized amplitude (P) distribution:

$$P(j) = \frac{\langle A_{40} \rangle (j)}{\sum_{k=1}^{N} \langle A_{40} \rangle (k)} \tag{1}$$

Figure 1C is a phase-amplitude plot that is obtained by plotting P as a function of phase bin. For our analysis, we chose $N = 72$ to obtain high-resolution CFC phase information (5° bins of the 4 Hz signal phase).

We use two metrics to quantify CFC. We calculate a modulation index, *MI*, introduced by Tort et al. (2010). With values between 0 and 1, this metric quantifies the distance between the normalized amplitude distribution, P, and the uniform distribution, or, the case when there is no CFC.

$$MI = \frac{\log N + \sum_{j=1}^{N} P(j) \log P(j)}{\log N} \tag{2}$$

TABLE 1 | Membrane conductance densities (mS/m^2) by level in the model structure for the three model pyramidal neurons.

Level	$g_{Na(F)}$	$g_{Na(P)}$	$g_{K(DR)}$	$g_{K(C)}$	$g_{K(A)}$	$g_{K(M)}$	g_{K2}	$g_{Ca(H)}$	$g_{Ca(L)}$	g_{h_exp}	g_{h_unif}	g_{h_0}
0	0	5	450	0	6	600	15	0	0	0	0	0
1	0	5	450	0	6	300	15	0	0	0	0	0
2	0	5	750	288	200	250	15	16	1	1	1	0
3	0	0.6	250	288	40	136	0	16	1	1.7635	1	0
4	0	0.12	0	9	40	136	0	16	1	2.4515	1	0
5	0	0.12	0	9	40	136	0	16	1	3.3194	1	0
6	0	1.2	500	288	40	136	15	16	1	1.7635	1	0
7	0	0.6	150	288	40	136	15	16	1	2.4515	1	0
8	0	0.12	2	9	40	136	15	16	1	3.3194	1	0
9	0	0.12	2	9	40	40	0	4	1	4.4141	1	0
10	0	0.12	2	9	40	40	0	4	1	5.7950	1	0
11	0	0.12	2	9	40	40	0	4	1	7.5368	1	0
12	0	0.12	2	9	40	40	0	4	1	9.7338	1	0
13	0	0.12	2	9	40	40	0	4	1	12.5050	1	0
14	0	0.12	2	9	40	40	0	4	1	16.0006	1	0
15	0	0.024	2	9	40	40	0	4	1	20.4098	1	0
16	0	0.024	2	9	40	10	10	4	1	25.9714	1	0
17	0	0.024	2	9	40	10	10	4	1	32.9866	1	0
18	0	0.024	2	9	40	10	10	4	1	41.8355	1	0
19	0	0.024	2	9	40	10	10	10.8	1	52.9971	2	0
20	0	0.024	2	9	40	10	10	2.4	1	52.9971	2	0

I_h columns are highlighted in maroon because their conductance density values were varied to obtain the three scenarios considered in this study: g_{h_exp}- exponential gradient (Hay et al., 2011), g_{h_unif}- uniform (Traub et al., 2005), g_{h_0}- no I_h. Explanation of levels: level 0 = initial axon; level 1 = rest of axon; level 2 = soma; level 3 = proximal basal and oblique dendrites; level 4 = middle basal and oblique dendrites; level 5 = distal basal and oblique dendrites; levels 6–18 = progressively more distal apical dendrite shaft (apdend1 to apdend13); level 19 = proximal apical tuft; level 20 = distal apical tuft.

The second metric we use to quantify CFC is height ratio, which is defined as:

$$height\ ratio = \frac{h_{max} - h_{min}}{h_{max}} \tag{3}$$

where h_{max} and h_{min} are the maximal and minimal normalized amplitudes, respectively, in the phase-amplitude plot (**Figure 1C**). The *MI* and height ratio values for the phase-amplitude plot in **Figure 1C** are 3.97×10^{-4} and 0.1735, respectively.

Sensitivity Analysis: Number of Bins and Injection Current Amplitude

The dependence of Equations (1–3) on N indicates that the number of bins chosen affects the calculated value for MI and height ratio. To better understand these dependencies, we ran the CFC analysis for several values of N (10, 12, 18, 20, 36, and 72). For this test, we injected a 1.5 nA 4 Hz sinusoidal current into the soma and a 1.5 nA 40 Hz sinusoidal current into the base of the apical dendrite in the model with an exponential gradient in I_h conductance density along the soma-apical dendrite axis. Over the range of values for N considered, MI decreased monotonically with some flattening as N increased (**Figure 2**). By contrast, height ratio remained mostly constant, particularly as the value of N increased (**Figure 2**). This result suggests that height ratio is the CFC metric that can more readily be compared across experiments that use different bin sizes. For this reason, we use height ratio as the CFC metric to present the results of this study. The corresponding MI values are presented in **Figures S1, S2**. For reference, Tort et al. (2010) determined that the significance threshold for concluding if CFC is present is MI $= 4.30 \times 10^{-5}$ (number of bins $N = 18$).

It has been reported that the amplitude of the membrane potential oscillations involved in CFC can affect the strength of CFC. In particular, this has been observed for slow frequencies, such as delta (0.5–4 Hz), theta (4–8 Hz), and alpha (8–12 Hz) (Hyafil et al., 2015). To examine this effect, we performed an amplitude sensitivity test for both the 4 and 40 Hz current injections (**Figure 3**) in the model with a soma-apical dendrite exponential gradient in I_h conductance density. In one scenario, we kept the amplitude of the 40 Hz current injection into the base of the apical dendrite constant at 1.5 nA while we performed the CFC analysis three times for 4 Hz somatic current injections with amplitudes of 1.5, 1.0, and 0.5 nA. Similarly, we ran the CFC analysis with constant 4 Hz amplitude of 1.5 and 40 Hz amplitudes of 1.5, 1.0, and 0.5 nA.

For both MI and height ratio, differences in amplitude of the 4 Hz current injection had a much bigger impact on CFC strength. Because we employ a variety of injection scenarios and

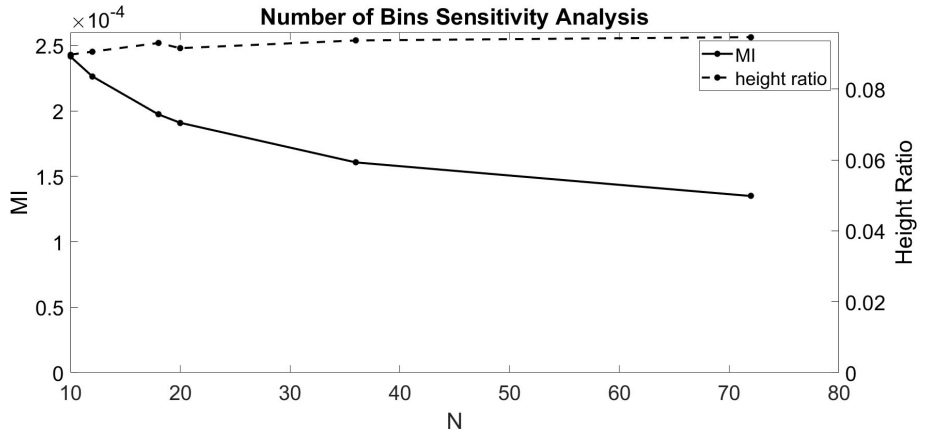

FIGURE 2 | Sensitivity analysis for the number of bins (*N*). MI index (solid, left ordinate) and height ratio (dashed, right ordinate) as a function of *N*. Values for *N* evaluated (10, 12, 18, 20, 36, and 72) indicated by dots.

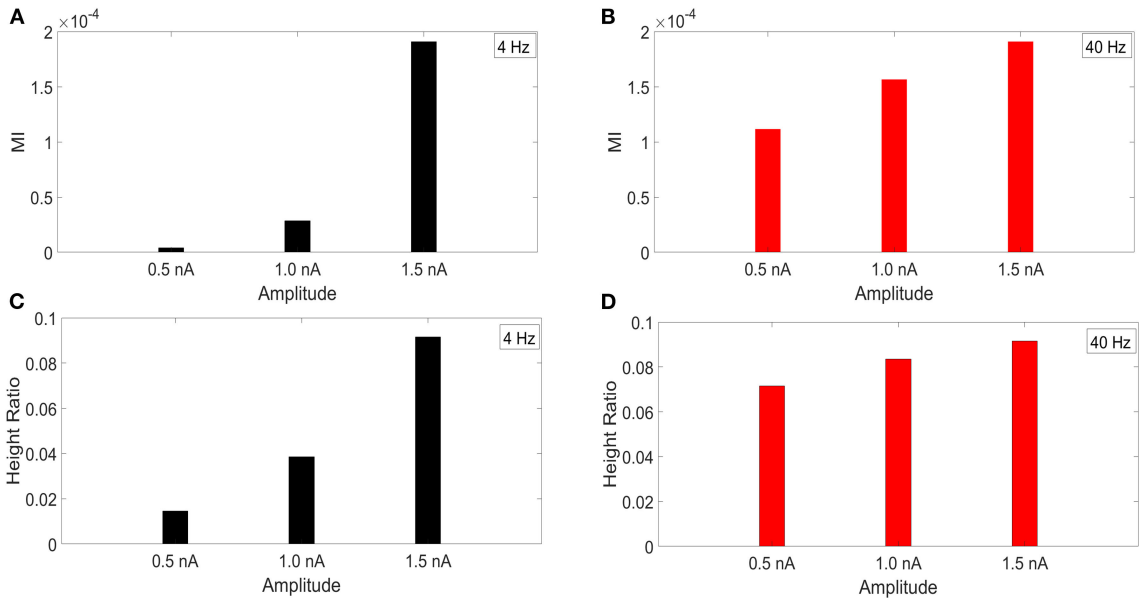

FIGURE 3 | Current injection amplitude sensitivity analysis. **(A,C)** MI and height ratio, respectively, for amplitudes of the 4 Hz current injection of 0.5, 1.0, and 1.5 nA while the amplitude of the 40 Hz injection was held constant at 1.5 nA. **(B,D)** MI and height ratio, respectively, for amplitudes of the 40 Hz current injection of 0.5, 1.0, and 1.5 nA while the amplitude of the 4 Hz injection was held constant at 1.5 nA.

there is a differential degree of filtering for the 4 and 40 Hz oscillations as they spread from their respective injection sites, the results of this amplitude sensitivity test will be useful for interpreting the CFC observed in our simulations.

RESULTS

Two Injection Scenarios

We used two injection scenarios to test for a range of possible input configurations of large cortical and limbic pyramidal neurons. In the first injection scenario, we simulated a 1.5 nA 4 Hz sinusoidal current injection into the most distal apical dendrite compartment while a 1.5 nA 40 Hz sinusoidal current was injected into a perisomatic (base of apical dendrite), middle apical, and distal apical dendrite compartment on successive runs (not simultaneous). This injection scenario is consistent with slow distal modulation of cortical pyramidal neurons via matrix thalamocortical or higher-order feedback (VanRullen and Koch, 2003; Spruston, 2008; Larkum, 2013).

A second injection scenario simulates perisomatic 4 Hz modulation. A 1.5 nA 4 Hz sinusoidal current was injected into the soma while a 1.5 nA 40 Hz sinusoidal current was injected into a perisomatic, middle apical, and distal apical dendrite compartment on successive runs. This type of slow

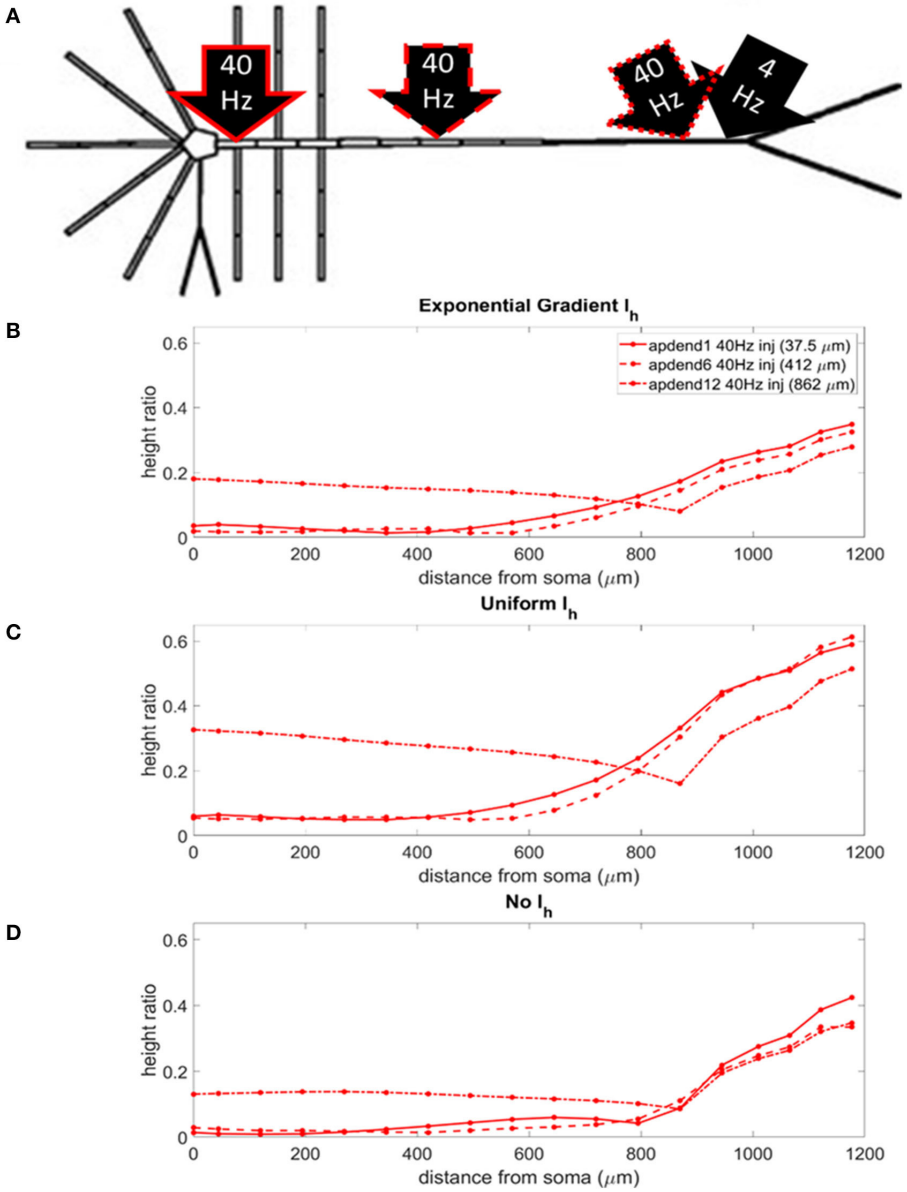

FIGURE 4 | (A) Distal apical dendrite 1.5 nA, 4 Hz modulation (black arrow) with 1.5 nA 40 Hz current injections in base of apical dendrite (black arrow, solid red outline), middle apical dendrite (black arrow, dashed red outline), and distal apical dendrite (black arrow, dot-dashed red outline). **(B)** Height ratio calculated for the soma, apical dendrite, and apical tuft in the model with exponential gradient in I_h conductance density along apical dendrite. **(C)** Height ratio calculated for the soma, apical dendrite, and apical tuft in model with uniform I_h conductance density. **(D)** Height ratio calculated for the soma, apical dendrite, and apical tuft in model with no I_h conductance. Dot symbols along the profiles in **(B–D)** indicate the distance from the soma of compartments along the soma-apical dendrite axis and apical tuft (measured from the beginning of each compartment).

modulation of the soma can be mediated by parvalbumin immunoreactive interneurons (PV) in the hippocampus and cortex (Stark et al., 2013).

CFC Strength: Distal 4 Hz Modulation

The locations of the current injections for the distal 4 Hz modulation scenario are presented in **Figure 4A**, and the profiles of height ratio along the soma-apical dendrite axis for models with I_h conductance densities configured for a soma-apical dendrite exponential gradient, uniform distribution, and zero conductance are presented in **Figures 4B–D**, respectively. (For clarity of presentation, the height ratio profiles in **Figures 4B–D, 5B–D**, and MI profiles in **Figures S1B–D, S2B–D**, were not calculated as continuous functions of distance from the soma, rather they were computed for each compartment which is shown in relation to distance from the soma in part A of each plot. The same format is used in **Figures 6–8**). All three I_h configurations show the same general

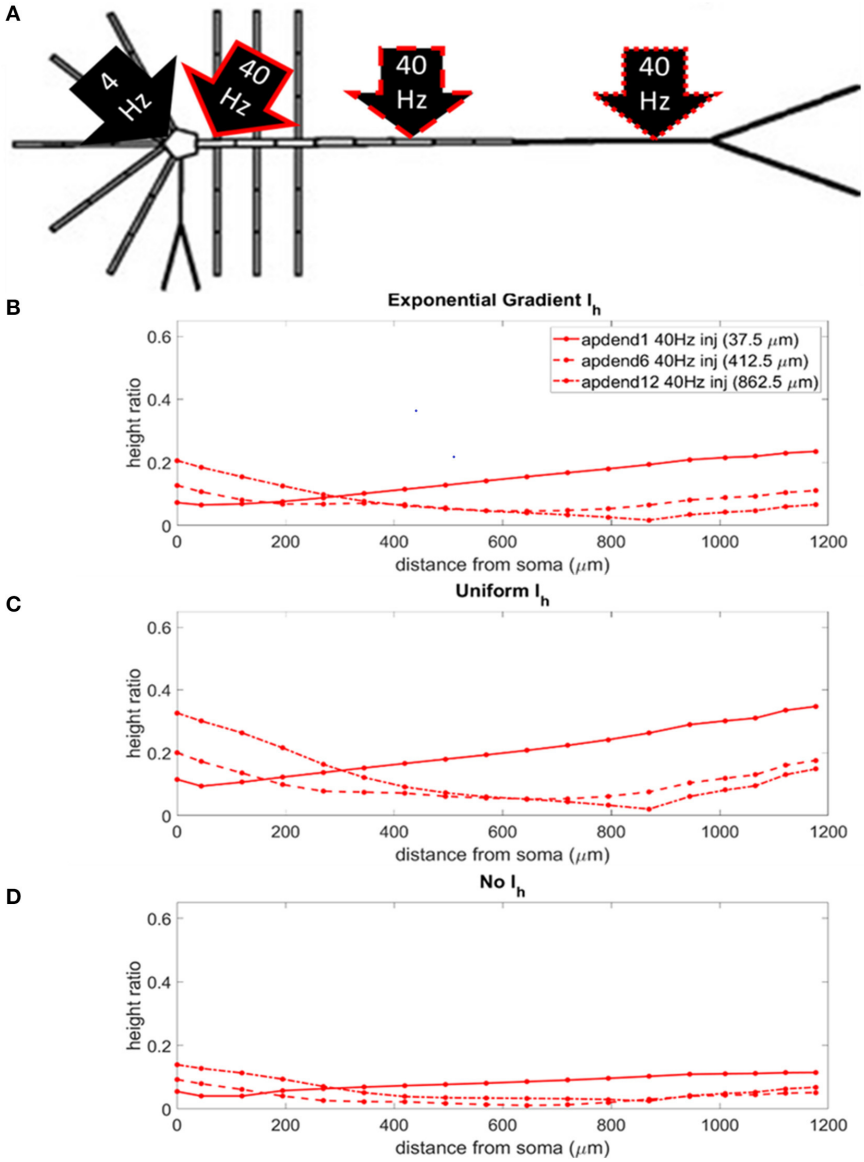

FIGURE 5 | (A) Perisomatic 1.5 nA, 4 Hz modulation (black arrow) with 1.5 nA 40 Hz current injections in base of apical dendrite (black arrow, solid red outline), middle apical dendrite (black arrow, dashed red outline), and distal apical dendrite (black arrow, dot-dashed red outline). **(B)** Height ratio calculated for the soma, apical dendrite, and apical tuft in model with exponential gradient in I_h conductance density along apical dendrite. **(C)** Height ratio calculated for the soma, apical dendrite, and apical tuft in model with uniform I_h conductance density. **(D)** Height ratio calculated for the soma, apical dendrite, and apical tuft in model with no I_h conductance. Dot symbols along the profiles in **(B–D)** indicate the distance from the soma of compartments along the soma-apical dendrite axis and apical tuft (measured from the beginning of each compartment).

pattern: strongest CFC in apical tuft (975–1,200 μm from the soma), with the soma and most of the lower apical dendrite (<700 μm from soma) exhibiting weaker CFC. For all three I_h configurations, the weakest CFC was observed between the soma and middle apical dendrite (0–450 μm from soma) when the 40 Hz current injection occurred in the soma or middle apical compartments. In particular, under these conditions in the model with either exponential gradient in I_h or no I_h, CFC strength almost reduced to zero (height ratios <0.02). On the other hand, when the 40 Hz current injection occurred in the distal apical dendrite (12th apical dendrite compartment, 825 μm from the soma), CFC was strong throughout the soma-apical dendrite axis and tuft.

The strongest CFC occurred in the model with a uniform distribution in I_h conductance density (**Figure 4C**). When the exponential gradient was added to the model (**Figure 4B**), CFC strength was uniformly lower, by almost 50%. When I_h was removed from the model, CFC strength was also lower, with differences in both the shape and relationships among the profiles.

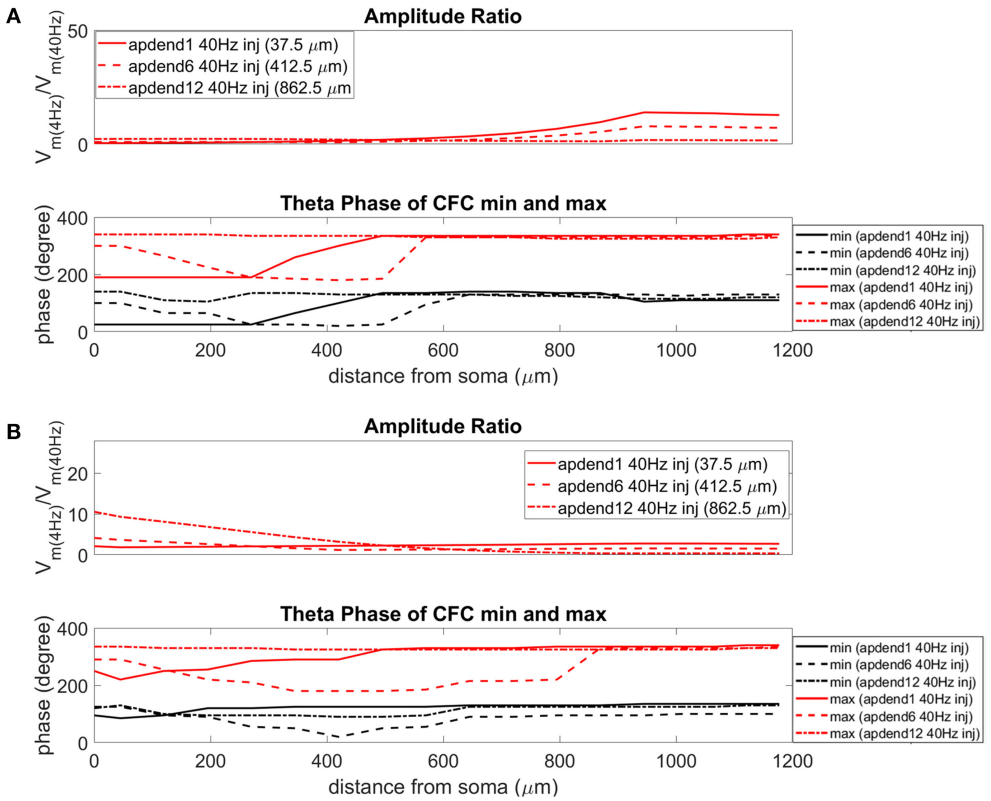

FIGURE 6 | Amplitude ratio and theta phase of CFC for model with exponential gradient in I_h conductance. **(A)** Distal 4 Hz modulation, amplitude ratio (top), and theta phase (bottom) of minimum (black) and maximum (red) amplitude of 40 Hz membrane potential oscillation. Forty Hertz current injected into base of apical dendrite (solid lines), middle of apical dendrite (dashed lines), and distal apical dendrite (dot-dashed lines). **(B)** Perisomatic 4 Hz modulation, amplitude ratio (top) and theta phase (bottom) of minimum (black) and maximum (red) amplitude of 40 Hz membrane potential oscillation. Forty Hertz current injected into base of apical dendrite (solid lines), middle of apical dendrite (dashed lines), and distal apical dendrite (dot-dashed lines).

CFC Strength: Perisomatic 4 Hz Modulation

The locations of the current injections for the perisomatic 4 Hz modulation scenario are presented in **Figure 5A**, and the profiles of height ratio along the soma-apical dendrite axis for models with I_h conductance densities configured for a soma-apical dendrite exponential gradient, uniform distribution, and zero conductance are presented in **Figures 5B–D**, respectively. All three I_h configurations again show a similar pattern for CFC strength. For compartments nearest the soma, CFC was strongest when the 40 Hz current injection occurred in distal apical dendrite, followed by 40 Hz current injection in middle apical dendrite, and finally weakest for 40 Hz current injection at the base of the apical dendrite (37.5 μm from soma). As you move up the apical dendrite toward the apical tuft, crossovers occur where now the strongest CFC was observed for 40 Hz current injection into the base of the apical dendrite, followed by injection into middle apical dendrite, and lastly injection into distal apical dendrite. When I_h is in the model and the 40 Hz current injection occurs in distal apical dendrite, CFC is almost completely eliminated at the site of the injection (12th apical dendrite compartment) (also see **Figures S2B,C**). On the other hand, CFC is almost completely eliminated in the middle apical compartments when I_h is not included in the model and the 40 Hz current injection occurs in the middle of the apical dendrite (see also **Figure S2D**).

Once again, CFC was strongest in the model with uniform I_h conductance density. CFC strength decreased by 30–40% when the exponential gradient in I_h conductance density was added to the model. Unlike in the previous injection scenario when the 4 Hz modulation occurred distally, when I_h was excluded from the model, there was a more noticeable decrease in CFC strength, particularly for the case when the 40 Hz current was injected into the base and middle of the apical dendrite.

AMPLITUDE AND PHASE INFORMATION

Amplitude

The current injection amplitude sensitivity analysis presented in **Figure 3** showed that there is greater dependence of CFC strength on the amplitude of the 4 Hz current injection than on the amplitude of the 40 Hz current injection. Because these two current injections differentially contribute to membrane potential oscillations throughout the model, it should be useful to compare the relative strength of membrane potential oscillations

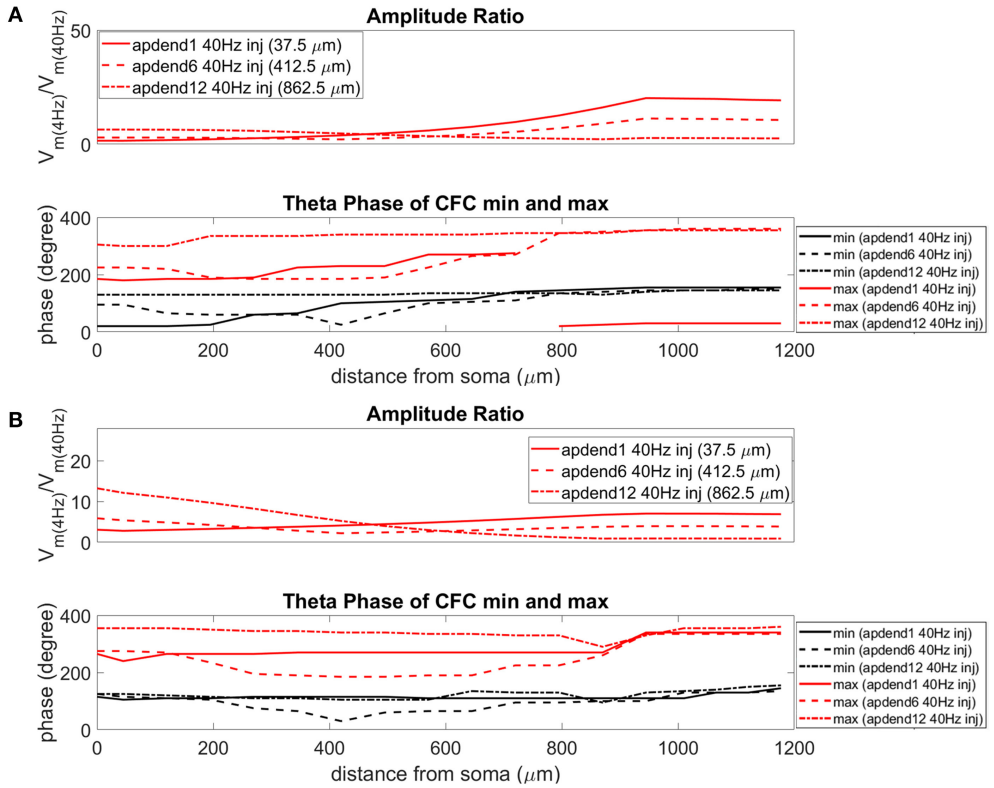

FIGURE 7 | Amplitude ratio and theta phase of CFC for model with uniform I_h conductance. **(A)** Distal 4 Hz modulation, amplitude ratio (top) and theta phase (bottom) of minimum (black) and maximum (red) amplitude of 40 Hz membrane potential oscillation. Forty Hertz current injected into base of apical dendrite (solid lines), middle of apical dendrite (dashed lines), and distal apical dendrite (dot-dashed lines). **(B)** Perisomatic 4 Hz modulation, amplitude ratio (top), and theta phase (bottom) of minimum (black) and maximum (red) amplitude of 40 Hz membrane potential oscillation. Forty Hertz current injected into base of apical dendrite (solid lines), middle of apical dendrite (dashed lines), and distal apical dendrite (dot-dashed lines).

at 4 and 40 Hz to better understand CFC strength for any given compartment. The top panel in **Figures 6A,B** (exponential gradient I_h), **Figures 7A,B** (uniform I_h), and **Figures 8A,B** (no I_h) show the *amplitude ratio* of membrane potential fluctuations at 4 Hz to the membrane potential fluctuations at 40 Hz $[\frac{V_{m(4Hz)}}{V_{m(40Hz)}}]$ along the soma-apical dendrite axis and apical tuft. Because of the low-pass filtering properties of the membrane, this ratio is expected to be higher the greater the distance is between the 4 and 40 Hz current injections. Another condition in which this ratio is expected to be high is when both the 4 and 40 Hz current injection occurred on the same end of the soma-apical dendrite axis and have traveled the length of this axis to the opposite end. In both of these situations, the 40 Hz signal experiences a greater degree of filtering than the 4 Hz signal due to the passive and active membrane properties. To see separately the filtered 4 and 40 Hz components of membrane potential oscillations for the simulation runs of this study, see **Figures S3, S4**.

When the 4 Hz current was injected in distal apical dendrite (**Figures 6A**, **7A**, **8A**, top), *amplitude ratios* were highest in distal apical dendrite and apical tuft when there was distance between the 4 and 40 Hz current injections, i.e., when the 40 Hz current was injected in the base or middle of the apical dendrite. On the other hand, when the 40 Hz current injection also occurred distally, the *amplitude ratio* was highest in the soma yet still lower than the case of 40 Hz current injected in the base or middle of the apical dendrite.

When the 4 Hz current was injected in the soma (**Figures 6B**, **7B**, **8B**, top), *amplitude ratios* were highest in perisomatic compartments when there was distance between the 4 and 40 Hz injections (40 Hz injections in middle and distal apical dendrite). When the 40 Hz injection occurred in the base of the apical dendrite, the amplitude ratio was highest in distal apical regions.

Phase

The bottom panel in **Figures 6A,B** (exponential gradient I_h), **Figures 7A,B** (uniform I_h), and **Figures 8A,B** (no I_h) show the phase of the theta cycle that both the maximum and minimum amplitudes of the 40 Hz oscillation occur along the soma-apical dendrite axis and apical tuft. When I_h is present, minimum amplitudes for the 40 Hz oscillation occurred consistently within the first half of the theta cycle (20–130°) while maximum amplitudes for the 40 Hz oscillation occurred consistently within the second half of the theta cycle (180–340°). In addition, there tends to be a shift toward earlier theta phases that both the minimum and maximum 40 Hz oscillation amplitude occurs. This shift occurs in the lower

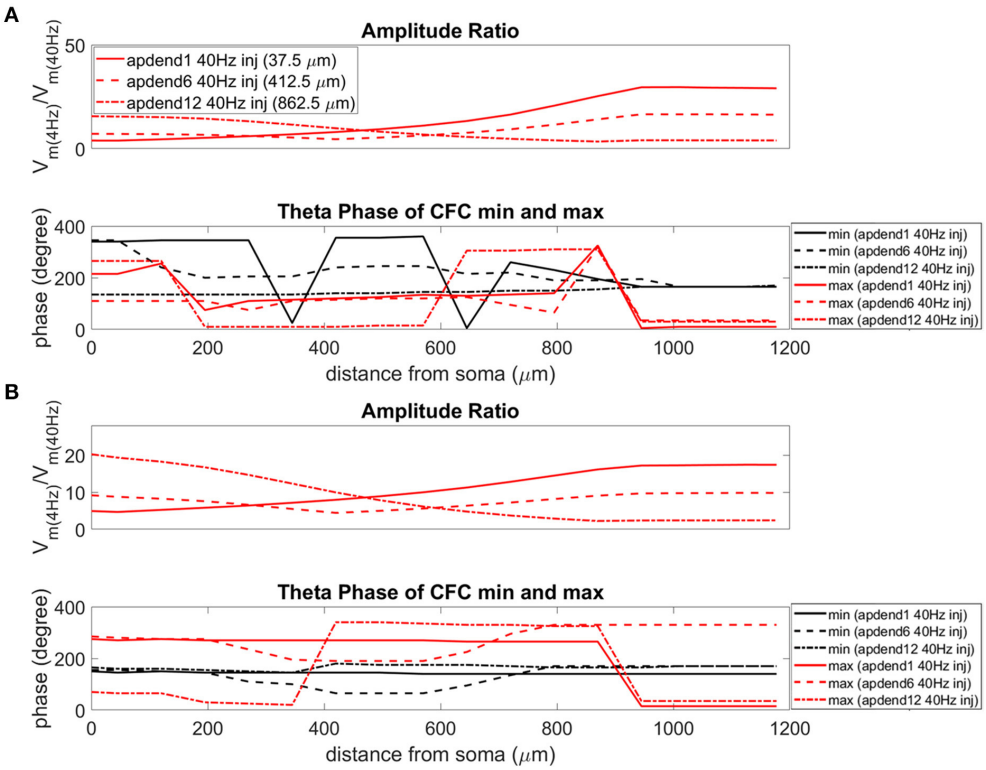

FIGURE 8 | Amplitude ratio and theta phase of CFC for model with no I_h conductance. **(A)** Distal 4 Hz modulation, amplitude ratio (top), and theta phase (bottom) of minimum (black) and maximum (red) amplitude of 40 Hz membrane potential oscillation. Forty Hertz current injected into base of apical dendrite (solid lines), middle of apical dendrite (dashed lines), and distal apical dendrite (dot-dashed lines). **(B)** Perisomatic 4 Hz modulation, amplitude ratio (top), and theta phase (bottom) of minimum (black) and maximum (red) amplitude of 40 Hz membrane potential oscillation. Forty Hertz current injected into base of apical dendrite (solid lines), middle of apical dendrite (dashed lines), and distal apical dendrite (dot-dashed lines).

half of the apical dendrite when the 40 Hz injection occurs there. This effect was the strongest (>100° shift to earlier theta phase) when the 40 Hz current was injected in the middle of the apical dendrite. In addition, this phase shift was very consistent because it occurred no matter the location of the 4 Hz modulation and for both exponential gradient and uniform I_h conductance densities—**Figures 6A,B** bottom and **Figures 7A,B** bottom, respectively. The phase shift that occurred when the 40 Hz current was injected into the base of the apical dendrite was more variable, being drastically reduced when the 4 Hz modulation occurred in distal apical dendrite (**Figure 6A** bottom and **Figure 7A** bottom).

When I_h is removed from the model, the phase relationship between minimum and maximum 40 Hz oscillation amplitude fundamentally changes (**Figures 8A,B** bottom). In the case of distal 4 Hz modulation, the phase relationship almost entirely flips so that the maximum 40 Hz oscillation amplitude often times occurred in the first half of the theta cycle while minimum 40 Hz oscillation amplitude occurs in the second half of the theta cycle. Whether the 4 Hz modulation occurs distally (**Figure 8A** bottom) or perisomatically (**Figure 8B** bottom), the phases of the minimum and maximum 40 Hz oscillation amplitude occur much closer together near the middle of the theta cycle relative to when the I_h conductance is in the model.

DISCUSSION

Implications on the Circuitry Underlying CFC

Distal 4 Hz modulation produced the strongest CFC in the distal regions of the model regardless of the location of the 40 Hz injection. In particular, when I_h was present, whether uniformly or exhibiting an exponential gradient, CFC was strongest for distal 4 Hz modulation and 40 Hz injection into the base and middle of the apical dendrite. On the other hand, when the 4 Hz modulation occurred in the soma, each model saw a significant reduction in the maximum strength of CFC.

The injection scenario of distal 4 Hz modulation and 40 Hz current injection in the base and middle of the apical dendrite is similar to what is typically believed to be a common input pattern to large cortical pyramidal neurons, such as theta or alpha matrix thalamocortical input to distal apical dendrite and tuft, and gamma core thalamocortical input in the perisomatic region (VanRullen and Koch, 2003; Spruston, 2008; Hawkins and Ahmad, 2016). Our results suggest that the processing of this type

TABLE 2 | Input resistance (MΩ) for the compartments receiving the 4 Hz injection (soma and apdend13) for all models used in this study.

	No I_h	Uniform I_h	Exp gradient I_h
perisomatic modulation (4 Hz, soma)	54	34	25
distal modulation(4 Hz, apdend13)	76	52	30

Input resistance obtained by dividing the amplitude of the filtered 4 Hz membrane potential oscillation for a compartment by the amplitude of the 4 Hz current injected into the compartment (1.5 nA).

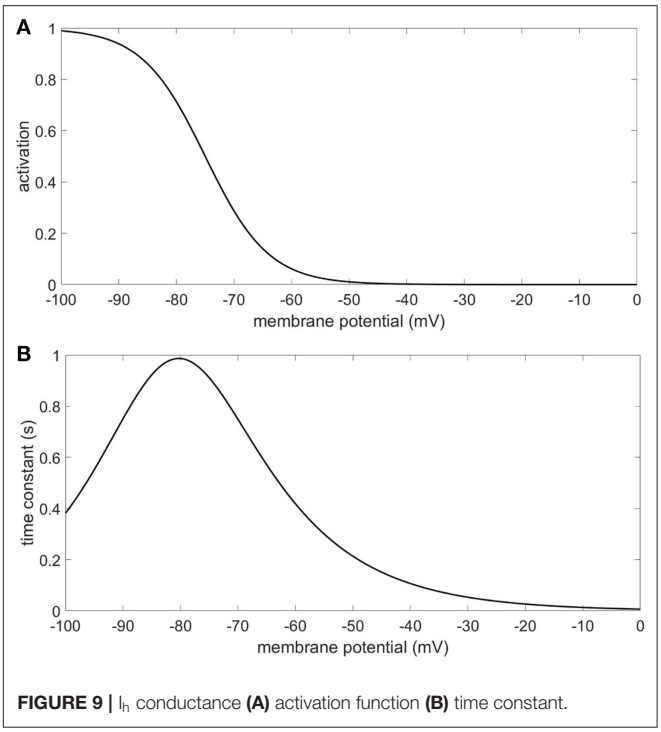

FIGURE 9 | I_h conductance **(A)** activation function **(B)** time constant.

of input to large cortical pyramidal neurons could be a source of CFC observed within the cortex.

The different I_h conductance density distributions used in our models impacts each model's input resistance to the 4 Hz modulation (**Table 2**). Input resistance increases as the I_h conductance density decreases, which happens when the distribution is changed from the exponential gradient to the uniform distribution. Input resistance increases further when I_h is removed. The effect of higher input resistance is larger membrane potential oscillations at 4 Hz in response to the 4 Hz modulation. Therefore, the exponential gradient in I_h more effectively filters the 4 Hz signal, leading to smaller *amplitude ratios* for this model than for the model with uniform I_h. Likewise, removing I_h lead to the highest *amplitude ratios*. As suggested by **Figure 3**, the profiles of *amplitude ratio* for the exponential gradient I_h model and uniform I_h model are a good indicator of CFC strength. In general, the higher the amplitude ratio, the stronger the CFC. However, for the model with no I_h, this trend did not continue. In this case, the highest *amplitude ratios* less faithfully produced the strongest CFC. Furthermore, despite producing the highest amplitude ratios seen in this study, the model with no I_h produced CFC that was comparable in strength to that produced by the exponential gradient I_h model when 4 Hz modulation was distal, and weaker CFC when 4 Hz modulation occurred perisomatically. This suggests that the *amplitude ratio* is only a good predictor of CFC strength when the I_h conductance is present.

Implications for Experimentalists

CFC has been observed for a wide range of spatiotemporal scales of brain activity, such as intracellular, LFP, and EEG recordings (Tort et al., 2010; Hyafil et al., 2015). The CFC profiles that we have observed for our current injection scenarios of distal 4 Hz modulation with perisomatic and middle apical dendrite 40 Hz injection is compatible with CFC results obtained by Sotero et al. (2015) who obtained LFP measurements at 100 μm intervals throughout the cortical depth in rats and performed CFC analysis on the signals. In our simulations, the 40 Hz oscillation generally was maximum during the hyperpolarizing phase of the 4 Hz oscillation and was minimum during the depolarizing phase of the 4 Hz modulation. A similar phase relationship between the 4 and 40 Hz oscillation was observed in Sotero et al. (2015). The results of the simulations in our study and the results of experimental work further suggests that the processing of distal slow input and perisomatic fast input by large cortical pyramidal neurons could be an underlying contributor to the CFC observed by meso- and macro-scale measurements, such as LFP and EEG.

Due to the impact of the I_h conductance on resting membrane potential, a compartment's resting membrane potential depends on where it is in the model and on which model it is in. The range of resting membrane potentials for the compartments in this study was −85 to −56 mV. The 4 Hz modulation starting from resting membrane potential will alternately activate and deactivate the I_h conductance [modeled after the anomalous rectifier in Traub et al. (2003, 2005)] to varying degrees depending on the compartment and I_h configuration (**Figure 9**). We observed that the presence of the I_h conductance in our models not only modulated the strength of CFC, it also imposed order in the phase characteristics of CFC. The phase information in the results of our simulations suggest that the amplitude of the 40 Hz oscillation is minimized when the I_h conductance is deactivating. This occurs during the depolarizing phase of the 4 Hz modulation which raises the membrane potential of the neuron and reduces the activation of the I_h conductance. On the other hand, our results show that the amplitude of the 40 Hz oscillation is maximized when the I_h conductance is activating, which occurs during the hyperpolarizing phase of the 4 Hz modulation. Removing the I_h conductance often resulted in an inverted and much more variable relationship between maximum 40 Hz membrane potential oscillation and minimum 40 Hz membrane potential oscillation, suggesting that the activation and deactivation of the I_h conductance plays a critical role in producing the CFC seen in our simulations. There are other ionic conductances included in our model that can, along with the I_h conductance, contribute to the occurrence of

CFC. However, our results indicate that the I_h conductance can significantly moderate the strength of CFC and impose order on the timing of the modulation experienced by the faster signal.

Connection to Epilepsy Research

Large pyramidal neurons in cortical layer five and the hippocampus are known to have an exponentially-increasing gradient in the conductance density for the I_h conductance along the soma-apical dendrite axis (Hu et al., 2009; Nusser, 2009; Hay et al., 2011). It has been reported that when these neurons lack this exponential gradient, the result is that these neurons may become epileptic sources, at least in part because of increased excitability of these neurons (Brennan et al., 2016). In our study, we have seen that a lack of this exponential gradient is also associated with increased CFC. Therefore, epilepsy may be impacted by not only increased excitability in the underlying neurons, but also increased CFC within these neurons.

CONCLUSION

We take a detailed look at the possible contribution of individual pyramidal neurons to CFC. Large pyramidal neurons, such as neocortical layer five or hippocampal pyramidal neurons, play a central role in the functioning of neocortical and limbic microcircuits. The neuronal membrane of these neurons has a large spatial extent and different parts of the neuron (e.g., proximal vs. distal) receive inputs from different populations of neurons within both local and widespread circuitry. Our study examined the ways in which different configurations of simultaneous fast and slow input are processed by individual pyramidal neurons and interact with each other to produce CFC. Furthermore, we have identified I_h as a current that has a large impact on the occurrence of CFC within pyramidal neurons.

This study can potentially shed light on which configurations of fast and slow input to pyramidal neurons produce the strongest CFC, and where exactly within the neuron CFC is strongest under realistic conditions of input. In addition, this study can illuminate the reasons why there may be differences between CFC strength observed in different regions of the brain and between different populations of neurons based on the configuration of the I_h conductance. This type of analysis may help experimentalists discriminate between CFC that originates from different underlying physiological mechanisms or determine if an exponential gradient in I_h conductance density is present or not.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

AUTHOR CONTRIBUTIONS

All authors contributed equally to experimental design and theoretical considerations. AY helped establish single-neuron model and implemented all simulated current injection capabilities. MF ran all simulations and developed CFC analysis post-processing algorithms. All authors contributed to data interpretation and manuscript preparation.

FUNDING

This research was funded by the U.S. Department of the Army and sponsored by the Combat Capabilities Development Command-Army Research Laboratory (CCDC-ARL). The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of CCDC-ARL or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

ACKNOWLEDGMENTS

We thank the reviewers for their insightful feedback that helped to significantly strengthen the manuscript and draw our attention to an error in the configuration of our model, allowing us to make the necessary correction.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncom.2020.00081/full#supplementary-material

REFERENCES

Brennan, G. P., Baram, T. Z., and Poolos N. P. (2016). Hyperpolarization-activated cyclic nucleotide-gated (HCN) channels in epilepsy. *Cold Spring Harb. Perspect. Med.* 6:a022384. doi: 10.1101/cshperspect.a022384

Bullis, J. B., Terrance, D. J., and Poolos, N. P., (2006). Reversed somatodendritic I_h gradient in a class of rat hippocampal neurons with pyramidal morphology. *J. Physiol.* 9(Pt 2), 431–443. doi: 10.1113/jphysiol.2006.123836

Canolty, R. T., and Knight, R. T., (2010). The functional role of cross-frequency coupling. *Trends Cogn. Sci.* 14, 506–515. doi: 10.1016/j.tics.2010.09.001

Felton, M. A., Yu, A. B., Boothe, D. L., Oie, K. S., and Franaszczuk, P. J. (2018). Resonance analysis as a tool for characterizing functional division of layer 5 pyramidal neurons. *Front. Comput. Neurosci.* 12:29. doi: 10.3389/fncom.2018.00029

Fernández-Ruiz, A., Olivia, A., Nagy, G. A., Maurer, A. P., Berényi, A., and Buzsáki, G. (2017). Entorhinal–CA3 dual-input control of spike timing in the hippocampus by theta-gamma coupling. *Neuron* 93, 1213–1226. doi: 10.1016/j.neuron.2017.02.017

Hawkins, J., and Ahmad, S. (2016). Why neurons have thousands of synapses, a theory of sequence memory in neocortex. *Front. Neural Circuits* 10:23. doi: 10.3389/fncir.2016.00023

Hay, E., Hill, S., Schurmann, F., Markram, H., and Segev, I. (2011). Models of neocortical layer 5b pyramidal cells capturing a wide range of dendritic and perisomatic active properties. *PLoS Comput. Biol.* 7:e1002107. doi: 10.1371/journal.pcbi.1002107

Hu, H., Vervaeke, K., Graham, L. J., and Storm, J. F. (2009). Complementary theta resonance filtering by two spatially segregated mechanisms in CA1 hippocampal pyramidal neurons. *J. Neurosci.* 29, 14472–14483. doi: 10.1523/JNEUROSCI.0187-09.2009

Hyafil, A., Giraud, A. L., Fontolan, L., and Gutkin, B. (2015). Neural cross-frequency coupling: connecting architectures, mechanisms, and functions. *Trends Neurosci.* 38, 725–740. doi: 10.1016/j.tins.2015.09.001

Jensen, O., and Colgin, L. L. (2007). Cross-frequency coupling between neuronal oscillations. *Trends Cogn. Sci.* 11, 267–269. doi: 10.1016/j.tics.2007.05.003

Korotkova, T., Fuchs, E. C., Ponomarenko, A., von Engelhardt, J., and Monyer, H. (2010). NMDA receptor ablation on parvalbumin-positive interneurons impairs hippocampal synchrony, spatial representations, and working memory. *Neuron* 68, 557–569. doi: 10.1016/j.neuron.2010.09.017

Lai, H. C., and Jan, L. Y. (2006). The distribution and targeting of neuronal voltage-gated ion channels. *Nat. Rev. Neurosci.* 7, 548–562. doi: 10.1038/nrn1938

Larkum, M. (2013). A cellular mechanism for cortical associations: an organizing principle for the cerebral cortex. *Trends Neurosci.* 36, 141–151. doi: 10.1016/j.tins.2012.11.006

Lisman, J. E., and Jenson, O. (2013). The theta-gamma neural code. *Neuron* 77, 1002–1016. doi: 10.1016/j.neuron.2013.03.007

Neymotin, S. A., Hilscher, M. M., Moulin, T. C., Skolnick, Y., Lazarewicz, M. T., and Lytton, W. W. (2013). Ih tunes theta/gamma oscillations and cross-frequency coupling in an *in silico* CA3 model. *PLoS ONE* 8:e76285. doi: 10.1371/annotation/00ea7ef3-aafd-4675-a3cc-c5087816b4d6

Nusser, Z. (2009). Variability in the subcellular distribution of ion channels increases neuronal diversity. *Trends Neurosci.* 32, 267–274. doi: 10.1016/j.tins.2009.01.003

Nusser, Z. (2012). Differential subcellular distribution of ion channels and the diversity of neuronal function. *Curr. Opin. Neurbiol.* 22, 366–371. doi: 10.1016/j.conb.2011.10.006

Sotero, R. C., Bortel, A., Naaman, S., Mocanu, V. M., Kropf, P., Villeneuve, M., et al. (2015). Laminar distribution of phase-amplitude coupling of spontaneous current sources and sinks. *Front. Neurosci.* 22:454. doi: 10.3389/fnins.2015.00454

Spruston, N. (2008). Pyramidal neurons: dendritic structure and synaptic integration. *Nat. Rev. Neurosci.* 9, 206–221. doi: 10.1038/nrn2286

Stark, E., Eichler, R., Roux, L., Fujisawa, S., Rotstein, H. G., and Buzsaki, G. (2013). Inhibition-induced theta resonance in cortical circuits. *Neuron* 80, 1263–1276. doi: 10.1016/j.neuron.2013.09.033

Tort, A. B., Komorowski, R. W., Manns, J. R., Kepell, N. J., and Eichenbaum, H. (2009). Theta-gamma coupling increases during the learning of item-context associations. *Proc. Natl. Acad. Sci. U.S.A.* 106, 20942–20947. doi: 10.1073/pnas.0911331106

Tort, A. B. L., Komorowski, R., Eichenbaum, H., and Kopell, N. (2010). Measuring phase-amplitude coupling between neuronal oscillations of different frequencies. *J. Neurophysiol.* 104, 1195–1210. doi: 10.1152/jn.00106.2010

Traub, R. D., Buhl, E. H., Cloveli, T., and Whittington, M. A. (2003). Fast rhythmic bursting can be induced in layer 2/3 cortical neurons by enhancing persistent Na+ conductance or by blocking BK channels. *J. Neurophysiol.* 89, 909–921. doi: 10.1152/jn.00573.2002

Traub, R. D., Contreras, D., Cunningham, M. O., Murray, H., LeBeau, F. E., Roopun, A., et al. (2005). Single-column thalamocortical network model exhibiting gamma oscillations, sleep spindles, and epileptogenetic bursts. *J. Neurophysiol.* 93, 2194–2232. doi: 10.1152/jn.00983.2004

Vaidya, S. P., and Johnston, D. (2013). Temporal synchrony and gamma to theta power conversion in the dendrites of CA1 pyramidal neurons. *Nat. Neurosci.* 16, 1812–1820. doi: 10.1038/nn.3562

VanRullen, R., and Koch, C. (2003). Is perception discrete or continuous? *Trends Cogn. Sci.* 7, 207–213. doi: 10.1016/S1364-6613(03)00095-0

Wuff, P., Ponomarenko, A. A., Bartos, M., Korotkova, T. M., Fuchs, E. C., Bahner, F., et al. (2009). Hippocampal theta rhythm and its coupling with gamma oscillations require fast inhibition onto parvalbumin-positive interneurons. *Proc. Natl. Acad. Sci. U.S.A.* 106, 3561–3566. doi: 10.1073/pnas.0813176106

Zhang, X., Min, X., Xu, X., Zheng, M., and Guo, L. (2016). ZD7288, a selective hyperpolarization-activated cyclic nucleotide-gated channel blocker, inhibits hippocampal synaptic plasticity. *Neural Regen. Res.* 11, 779–786. doi: 10.4103/1673-5374.182705

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.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 30 September 2020
doi: 10.3389/fncom.2020.558477

Identifiability of a Binomial Synapse

Camille Gontier[1]* *and Jean-Pascal Pfister*[1,2]

[1] Department of Physiology, University of Bern, Bern, Switzerland, [2] Institute of Neuroinformatics and Neuroscience Center Zurich, University of Zurich/ETH Zurich, Zurich, Switzerland

Synapses are highly stochastic transmission units. A classical model describing this stochastic transmission is called the binomial model, and its underlying parameters can be estimated from postsynaptic responses to evoked stimuli. The accuracy of parameter estimates obtained via such a model-based approach depends on the identifiability of the model. A model is said to be structurally identifiable if its parameters can be uniquely inferred from the distribution of its outputs. However, this theoretical property does not necessarily imply practical identifiability. For instance, if the number of observations is low or if the recording noise is high, the model's parameters can only be loosely estimated. Structural identifiability, which is an intrinsic property of a model, has been widely characterized; but practical identifiability, which is a property of both the model and the experimental protocol, is usually only qualitatively assessed. Here, we propose a formal definition for the practical identifiability domain of a statistical model. For a given experimental protocol, this domain corresponds to the set of parameters for which the model is correctly identified as the ground truth compared to a simpler alternative model. Considering a model selection problem instead of a parameter inference problem allows to derive a non-arbitrary criterion for practical identifiability. We apply our definition to the study of neurotransmitter release at a chemical synapse. Our contribution to the analysis of synaptic stochasticity is three-fold: firstly, we propose a quantitative criterion for the practical identifiability of a statistical model, and compute the identifiability domains of different variants of the binomial release model (uni or multi-quantal, with or without short-term plasticity); secondly, we extend the Bayesian Information Criterion (BIC), a classically used tool for model selection, to models with correlated data (which is the case for most models of chemical synapses); finally, we show that our approach allows to perform data free model selection, i.e., to verify if a model used to fit data was indeed identifiable even without access to the data, but having only access to the fitted parameters.

Keywords: model selection, practical identifiability, structural identifiability, binomial, synapse

OPEN ACCESS

Edited by:
Renaud Blaise Jolivet, Université de Genève, Switzerland

Reviewed by:
Ryota Kobayashi, University of Tokyo, Japan
Lubomir Kostal, Institute of Physiology CAS, Czechia
Suhita Nadkarni, Indian Institute of Science Education and Research, India

***Correspondence:**
Camille Gontier
camille.gontier@unibe.ch

Received: *02 May 2020*
Accepted: *18 August 2020*
Published: *30 September 2020*

Citation:
Gontier C and Pfister J-P (2020) Identifiability of a Binomial Synapse. Front. Comput. Neurosci. 14:558477. doi: 10.3389/fncom.2020.558477

1. INTRODUCTION

Model selection is highly relevant to neuroscience, as neurons, dendrites, and synapses can be represented by models with different levels of complexity and abstraction. When it comes to fitting recorded data, predicting the output of a system to a given stimulus, or making sense of an observed phenomenon, several possible models can be used: this raises the question of what makes a good model. Finding the correct model is a crucial issue in studying the brain.

Firstly, a good model needs to be sufficiently complex to account for observed data, while being simple enough to generalize to future observations. Competing models are typically compared based on their ability to fit an observed data set, while being penalized for their complexity (or number of free parameters) to avoid overfitting. Different model selection tools (Bayesian Information Criterion, Akaike Information Criterion,...) are classically used to determine which model is the best one to fit a data set (Daw et al., 2011).

Secondly, models also differ in their nature, and can be classified as phenomenological, normative, or biophysical. On the one hand, purely phenomenological models are useful for relating the output of a system to its input, and can provide a computationally efficient way to make prediction. However, as they are solely based on the empirical relation between the input and the output of the system, and not on its inner biological principles, they lack interpretability. On the other hand, normative and biophysical models can be computationally challenging to fit on data, but are more realistic. In a normative approach, the output of a system is computed from an objective function which models its high-level functions and principles. As opposed to this top-down approach, biophysical models aim at precisely describing the low-level biological components of the system. An interesting property of these biophysical models is that their parameters correspond to real physical quantities: when the parameters of a system cannot be measured directly, they can be estimated by fitting a corresponding biophysical model on recorded output data of the system, a procedure known as model-based inference. By computing the likelihood of the data as a function of the parameters, it is possible to follow a maximum-likelihood approach to obtain a point estimate of the parameters (Barri et al., 2016), or to compute the full posterior distribution over them (Bird et al., 2016).

Such a parameter inference requires that the model used is identifiable. Structural (i.e., model-based) identifiability is a property of the model, regardless of experimental results. In a structurally identifiable system, the dimension of the output is sufficiently high with respect to the dimension of the parameters vector to uniquely define it: the parameters can be non-ambiguously inferred if the complete distribution of the output is known. Structural identifiability has been widely studied in many fields of physics and biology (Raue et al., 2009, 2011; Komorowski et al., 2011; Koyama, 2012; Hines et al., 2014), and different criteria exist to assess the structural identifiability of a model (Massonis and Villaverde, 2020).

This theoretical property is not equivalent to practical (i.e., experiment-based) identifiability, which is a property of both the model and the experimental protocol: a model which is structurally identifiable might lead to a poor practical identifiability of parameters if data points are noisy or scarce. The accuracy of model-based methods for inferring the values of parameters depends on the experimental protocol used to record the data, as observations need to be sufficiently informative to allow a correct estimation of the parameters. Contrary to structural identifiability, a quantitative criterion is lacking for practical identifiability, which is usually only qualitatively assessed. Non-practical identifiability refers to regimes in which parameters can only be loosely estimated; but one would need to define what does "loose" mean. Such a definition could be intrinsic to the model: a model could be considered as practically identifiable given a certain experimental protocol if the expected variance of its parameters' estimate is below a threshold. But this threshold would need to be arbitrarily defined. Here, we propose an extrinsic yet non-arbitrary definition of practical identifiability, by transforming a model identifiability problem into a model selection problem.

A model is said to be practically identifiable when its parameters can be correctly inferred given a certain experimental protocol. But, as explained previously, different possible models can be fitted on a data set. Recorded data need to be sufficiently informative not only to give a correct estimate of the parameters of a model, but also to select the correct model (i.e., the model from which they have been generated). We argue that a model is practically identifiable if and only if it is also correctly identified as the model providing the best fit to the data. For a given experimental protocol, we define the practical identifiability domain of a statistical model as the set of parameters for which the model is correctly identified as the ground truth compared to a simpler alternative submodel.

Our proposed definition of practical identifiability can be applied to any setting where submodels or a nested family of models can be defined. Here, we apply it to the particular problem of estimating the parameters of a chemical synapse. A classical biophysical model used to describe the stochastic release of neurotransmitter at chemical synapses is called the binomial model (Katz, 1969), for which different variants of increasing complexity (in term of the number of free parameters) can be considered.

Different model-based approaches have been proposed (Bykowska et al., 2019) for obtaining an accurate estimate of the parameters describing a synapse (namely, its number of independent release sites, their release probability upon the arrival of a presynaptic spike, the quantum of current elicited by one release event, etc.) These parameters cannot be measured directly, but can be inferred using excitatory postsynaptic currents (EPSCs[1]) recorded on the post-synaptic side and elicited by experimental stimulation of the presynaptic cell. By measuring their values before and after a stimulation protocol, it is possible to study the mechanisms and loci of synaptic plasticity (Costa et al., 2015, 2017a,b) and homeostasis (Davis and Müller, 2015; Wentzel et al., 2018). On a more theoretical level, a correct inference of synaptic parameters is necessary to study the computational role of synaptic stochasticity (Levy and Baxter, 2002; Guo and Li, 2012). Finally, an accurate inference of synaptic parameters would allow to clarify the role of synaptic transmission in different diseases (Van Spronsen and Hoogenraad, 2010), such as mental retardation (Pfeiffer and Huber, 2009), schizophrenia (Stephan et al., 2006), Parkinson's disease (Calabresi et al., 2006), autism (Südhof, 2008), Alzheimer's disease (Selkoe, 2002), compulsive

[1]It is also possible to perform model-based inference of synaptic parameters based on post-synaptic spike trains instead of EPSCs, as in Ghanbari et al. (2017, 2020)

behavior (Welch et al., 2007), and addiction (Kauer and Malenka, 2007).

Our contribution to the analysis of synaptic stochasticity is three-fold. Firstly, we propose a definition for the practical identifiability of a model of synaptic transmission, and compute the identifiability domains of different variants of the binomial release model. Besides, we observe that model selection criteria are classically derived by assuming that recorded data are not correlated, which does not hold for most models of chemical synapse. We extend the Bayesian Information Criterion (BIC), a classically used tool for model selection, to models with correlated data. Finally, a proper description of the model selection step is often missing in studies where a model based-approach is used to infer synaptic parameters. We show that our approach allows to perform *data free* model selection, i.e., to verify if a model used to fit data was indeed identifiable even if a proper model selection step had not been performed.

2. METHODS

2.1. Binomial Models of Neurotransmitter Release

2.1.1. The Classical Binomial Model

The quantal nature of synaptic transmission was first unveiled in Del Castillo and Katz (1954), in which the authors observed that the postsynaptic responses to presynaptic stimulations were all multiples of a small unit of current. They explained how the total response is built up of several of these units, or quanta, each of them arising from a single presynaptic release event. Upon the arrival of an action potential in the presynaptic terminal, vesicles are released with a given probability p. The binomial model (Katz, 1969) assumes that there are N independent release sites and that for each site the release probability p is identical. Therefore, the number of released vesicles k_i after spike i is distributed according to a binomial distribution. This model further assumes that each vesicle release gives rise to a quantal current q, such that the overall excitatory postsynaptic current is given by $e_i = qk_i + \epsilon$, where ϵ models a measurement noise typically drawn from a normal distribution with variance σ^2. Under the binomial model described by its parameters N, p, q, and σ, the distribution of EPSCs is given by

$$p(e_i) = \sum_{k_i=0}^{N} p(e_i|k_i)p(k_i)$$

where k_i follows a binomial distribution with parameters N and p, and e_i conditioned on k_i follows a normal distribution with mean qk_i and variance σ^2. Postsynaptic responses are characterized by their mean Npq and their variance $q^2Np(1-p)+\sigma^2$. A first feature of synaptic transmission is thus its stochasticity. Due to different sources of noise, such as probabilistic vesicles release or recording noise, postsynaptic recordings exhibit trial-to-trial variability.

2.1.2. Full Model of Synaptic Transmission

Although this simple binomial model accounts for synaptic stochasticity, it does not allow to model its dynamics: postsynaptic responses do not only depend on the parameters of the synapse, but also on its previous activity. On the one hand, successive presynaptic stimulations within a short time interval will lead to a depletion of the readily-releasable vesicle pool, and hence to reduced successive postsynaptic responses, a phenomenon known as short-term depression. This can be modeled by assuming that the number of available vesicles at time i is $n_i \leq N$ (while the simplified binomial model described above assumes that all vesicles are readily releasable, and hence $n_i = N$). On the other hand, successive stimulations will gradually increase the presynaptic calcium concentration, and hence the release probability, which is called short-term facilitation.

Short-term depression and facilitation can be modeled using the Tsodyks-Markram model (Tsodyks et al., 1998; Costa et al., 2015). It consists in two ordinary differential equations, which model the proportion of available vesicles r_i and the release probability u_i at time i. r_i is reduced by an amount $u_i r_i$ after each presynaptic spike, and recovers back to 1 with a depression time constant τ_D between each spike. Similarly, u_i is increased by an amount $p(1-u_i)$, and decays back to p (its baseline value) with a facilitation time constant τ_F. Different values of the parameters p, τ_D, and τ_F allow to represent different synaptic dynamics (either depression, facilitation, or no plasticity at all).

However, such a deterministic approach to short-term plasticity only allows to model averages, and neglects correlations between successive postsynaptic responses. In recent studies (Barri et al., 2016; Bird et al., 2016), models of synapses incorporating both short-term plasticity and binomial models of vesicles release and refill have been proposed. In these models, the release probability u_i evolves according to the equation of the Tsodyks-Markram model, while each vesicle refills with a probability $1 - \exp(-\Delta_{t_i}/\tau_D)$, where Δ_{t_i} is the time interval between two successive presynaptic stimulations. This approach allows to represent both the stochasticity and the dynamics of neurotransmitter release, and to compute the likelihood of a set of recorded data $\mathcal{D}$ given the parameters θ and the presynaptic stimulation protocol Ψ.

We consider a model of chemical synapse which encompasses both short-term depression (STD) and facilitation (STF) (Barri et al., 2016; Bird et al., 2016). Its parameters are (**Figure 1A**):

- N: the number of independent release sites [-]
- p: their initial release probability [-]
- σ: the recording noise. It encompasses both the noise coming from the experimental apparatus (thermal noise of the amplifier, electric line noise, etc.) and from the recordings per se (such as fluctuations in the membrane potential of the cell) [A]
- q: the quantum of current elicited by one release event [A][2]
- τ_D: the time constant of vesicles refilling, and hence of short-term depression [s]
- τ_F: the time constant of Ca^{2+} dynamics, and hence of short-term facilitation [s]

[2]The outputs of a model of chemical synapse can be either electric postsynaptic currents (EPSC) or potentials (EPSP). In the latter case, σ and q will be expressed in [V] instead of [A].

which defines a vector $\theta = (N, p, q, \sigma, \tau_D, \tau_F)$. A probability conditioned on a parametrization θ is written p_θ.

$\mathbf{n} = \{n_i\}_{1\leq i\leq T}$, $\mathbf{k} = \{k_i\}_{1\leq i\leq T}$, and $\mathcal{D} = \{e_i\}_{1\leq i\leq T}$ represent respectively the number of available vesicles at the moment of spike i, the number of vesicles released after spike i, and the i-th recorded EPSC (**Figure 1B**). The experimental protocol $\Psi = \{t_1, t_2, \ldots, t_T\}$ encompasses the times of presynaptic stimulation: the time of the i-th spike is written t_i and $\Delta t_i = t_i - t_{i-1}$. For simplicity, we will drop the dependency on Ψ from the notations of probabilities.

The probability of recording $\mathcal{D}$ is computed as the marginal of the joint distribution of the observations $\mathcal{D}$ and the hidden variables $\mathbf{n}$ and $\mathbf{k}$:

$$p_\theta(\mathcal{D}, \mathbf{n}, \mathbf{k}) = p_\theta(e_1|k_1)p_\theta(k_1|n_1)p_\theta(n_1) \prod_{i=2}^{T} p_\theta(e_i|k_i)p_\theta(k_i|n_i)p_\theta(n_i|n_{i-1}, k_{i-1}) \quad (1)$$

where $p_\theta(e_i|k_i)$ is the emission probability, i.e., the probability to record e_i knowing that k_i vesicles released neurotransmitter and assuming a normally distributed recording noise[3]:

$$p_\theta(e_i|k_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(e_i - qk_i)^2}{2\sigma^2}\right) \quad (2)$$

$p_\theta(k_i|n_i)$ is the binomial distribution and represents the probability that, given n_i available vesicles, k_i of them will indeed release neurotransmitter:

$$p_\theta(k_i|n_i) = \binom{n_i}{k_i} u_i^{k_i}(1 - u_i)^{n_i - k_i} \quad (3)$$

where the release probability u_i evolves as

$$u_i = p + u_{i-1}(1 - p)\exp\left(-\frac{\Delta t_i}{\tau_F}\right) \quad (4)$$

with $u_1 = p$. $p_\theta(n_i|n_{i-1}, k_{i-1})$ represents the process of vesicles refilling. During the time interval Δt_i, each empty vesicle can refill with a probability I_i:

$$p_\theta(n_i|n_{i-1}, k_{i-1}) = \binom{N - n_{i-1} + k_{i-1}}{n_i - n_{i-1} + k_{i-1}} I_i^{n_i - n_{i-1} + k_{i-1}}(1 - I_i)^{N - n_i} \quad (5)$$

with

$$I_i = 1 - \exp\left(-\frac{\Delta t_i}{\tau_D}\right) \quad (6)$$

It is usually assumed that, at the beginning of the experiment, all release sites are filled, and hence that $n_1 = N$ (Barri et al., 2016; Bird et al., 2016).

[3]Other distributions can also be used for the emission probability. Barri et al. (2016) assumed an inverse Gaussian to account for the observed right-skewness of mEPSP (Bekkers et al., 1990; Bhumbra and Beato, 2013)

2.2. Models, Submodels, and Nested Families

Definition 2.1. Model. For a given data set $\mathcal{D}$ and experimental protocol Ψ, a model $\mathcal{M}$ is defined as a triplet $\mathcal{M} = \{\Theta, \pi, \mathcal{L}\}$ where Θ is the space of parameters $\theta \in \Theta$, π is the prior for the parameters $\pi(\theta) = p(\theta|\mathcal{M})$, and $\mathcal{L}$ is the likelihood of the parameters $\mathcal{L}(\theta|\mathcal{D}) = p(\mathcal{D}|\theta, \mathcal{M}, \Psi)$.

Examples: Different models can be derived from Equations (1) to (6). We consider four models of decreasing complexity:

Model $\mathcal{M}_3$ is the full model with both STD and STF. Its six parameters are N, p, q, σ, τ_D, and τ_F, and hence $\Theta_3 = \mathbb{N}^* \times [0, 1] \times (\mathbb{R}_+)^4$. Its likelihood function $\mathcal{L}_3$ is obtained by marginalizing out the hidden variables $\mathbf{n}$ and $\mathbf{k}$:

$$\mathcal{L}_3(\theta|\mathcal{D}) = \sum_{\mathbf{n},\mathbf{k}} p_\theta(\mathcal{D}, \mathbf{n}, \mathbf{k}) \quad (7)$$

where $p_\theta(\mathcal{D}, \mathbf{n}, \mathbf{k})$ is given by Equation (1).

Model $\mathcal{M}_2$ has only short-term depression (and no short-term facilitation). Its five parameters are N, p, q, σ, and τ_D, and hence $\Theta_2 = \mathbb{N}^* \times [0, 1] \times (\mathbb{R}_+)^3$. Its likelihood $\mathcal{L}_2$ can be derived from (7) by further assuming that u_i is a constant equal to p.

Model $\mathcal{M}_1$ shows no short-term plasticity at all, and can be derived from model $\mathcal{M}_2$ by further assuming that I_i (defined in 6) is a constant equal to 1 (and hence $n_i = N$). Its four parameters are N, p, q, and σ, and hence $\Theta_1 = \mathbb{N}^* \times [0, 1] \times (\mathbb{R}_+)^2$. In this setting, data points are independent and (7) becomes

$$\mathcal{L}_1(\theta|\mathcal{D}) = \prod_{i=1}^{T} \left(\sum_{k_i=0}^{N} p_\theta(e_i|k_i)p_\theta(k_i)\right) \quad (8)$$

with $p_\theta(k_i) = \binom{N}{k_i} p^{k_i}(1 - p)^{N - k_i}$ being the binomial distribution;

Model $\mathcal{M}_0$ is a Gaussian model, in which EPSCs are simply generated from a normal distribution parameterized by its mean and variance. Its two parameters are μ and σ^2, and hence $\Theta_0 = \mathbb{R} \times \mathbb{R}_+$. Its likelihood $\mathcal{L}_0$ is simply

$$\mathcal{L}_0(\theta|\mathcal{D}) = \prod_{i=1}^{T} p_\theta(e_i) \quad (9)$$

with $p_\theta(e_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(e_i - \mu)^2}{2\sigma^2}\right)$.

To ensure the completeness of the definition of the models, we will assume for each parameter θ a uniform prior between two values $\theta_{\min}$ and $\theta_{\max}$ (Bird et al., 2016). Note however that the approximate identifiability domain defined in (17) does not depend on the prior.

Definition 2.2. Submodels. Although ubiquitous in statistics (as in the likelihood-ratio test or Pearson's chi-squared test), the notion of submodels (or nested models) is rarely formally defined in the literature (Edwards and MacCallum, 2012). It is usually said that $\mathcal{M}_0$ is a submodel of $\mathcal{M}_1$ (or is nested within $\mathcal{M}_1$)

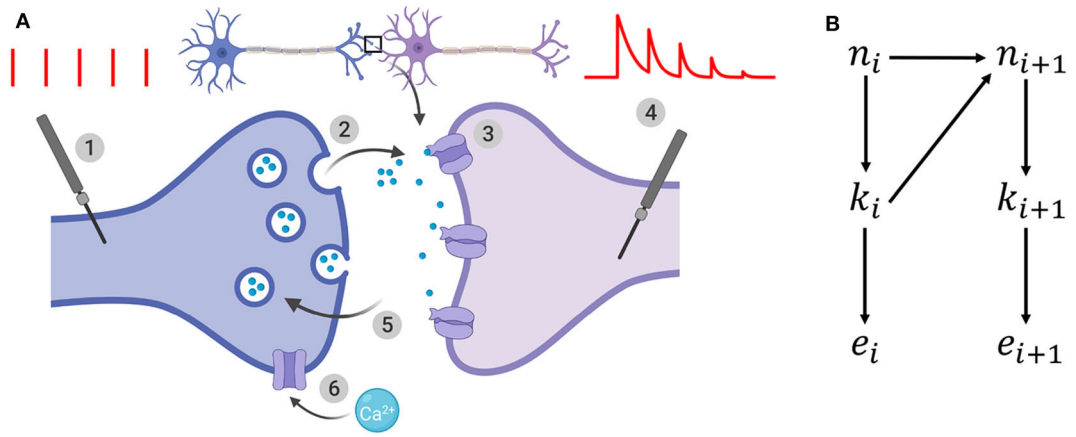

FIGURE 1 | (A) Illustration of the binomial model. (1): the presynaptic button is artificially stimulated. Red vertical bars show 5 presynaptic spikes with a constant interspike interval. (2): these evoked stimuli lead to neurotransmitter release. After spike i, k_i vesicles (out of the n_i vesicles from the readily-releasable vesicle pool) release their neurotransmitter with a probability u_i. (3): neurotransmitter bind to receptors and elicit a postsynaptic current. A single release event triggers a quantal response of amplitude q. (4): the recorded postsynaptic response after spike i is the sum of the effects of the k_i release events. EPSCs correspond to the amplitude of each peak of the postsynaptic response to a presynaptic spike. (5): out of the N release sites, only n_i are in the readily-releasable vesicle pool at the moment of spike i. After releasing, vesicles recover with a time constant τ_D which determines short-term depression. (6): in the same time, each spike increases the calcium concentration in the presynaptic button, and hence increases the release probability u_i. This short-term facilitation is characterized by a time constant τ_F. **(B)** Generative model for the dynamical binomial model where n_i is the number of ready releasable vesicles and k_i is the number of released vesicles at time i. e_i is the EPSC amplitude at time t_i.

if $\mathcal{M}_0$ can be obtained by constraining the parameters of $\mathcal{M}_1$ (Gottman, 1995). We propose the following formal definition, that encompasses the space of parameters, their priors, and their likelihood.

$\mathcal{M}_0 = \{\Theta_0, \pi_0, \mathcal{L}_0\}$ is said to be a submodel of $\mathcal{M}_1 = \{\Theta_1, \pi_1, \mathcal{L}_1\}$ if

1. $\Theta_0 \subset \Theta_1$ (i.e. the parameters of $\mathcal{M}_0$ also appear in $\mathcal{M}_1$)
2. $\pi_0(\theta_0) = \int_{\Theta_1 \setminus \Theta_0} \pi_1(\theta_0, \tilde{\theta}) d\tilde{\theta}, \quad \forall \theta_0 \in \Theta_0$ (i.e. $\mathcal{M}_0$ and $\mathcal{M}_1$ share the same priors for the parameters they have in common)
3. $\forall \theta_0 \in \Theta_0, \exists \tilde{\theta}$ s.t. $p(\mathcal{D}|\theta_0, \mathcal{M}_0) = p(\mathcal{D}|(\theta_0, \tilde{\theta}), \mathcal{M}_1)$ with $(\theta_0, \tilde{\theta}) \in \Theta_1$ (i.e., $\mathcal{M}_0$ can be retrieved from $\mathcal{M}_1$ by constraining its parameters).

We use the notation $\mathcal{M}_0 \preceq \mathcal{M}_1$.

Examples: The model $\mathcal{M}_2$ with only short-term depression is a submodel of $\mathcal{M}_3$ (which accounts for both depression and facilitation). Indeed, they have the parameters N, p, q, σ, and τ_D in common, and $\mathcal{M}_2$ can be retrieved from $\mathcal{M}_3$ by constraining $\tau_F \to 0$. Similarly, the model without STP $\mathcal{M}_1$ is a submodel of $\mathcal{M}_2$ where $\tau_D \to 0$; and the uni-quantal model $\mathcal{M}_0$ is a submodel of the multi-quantal model $\mathcal{M}_1$ where $p = 1$ and $\mu = Nq$.

We propose the following definitions to characterize the nestedness of a family of models:

Definition 2.3. Nested family. $\mathcal{F} = \{\mathcal{M}_0, \mathcal{M}_1, \ldots, \mathcal{M}_n\}$ is said to be a nested family if

$$\mathcal{M}_i \preceq \mathcal{M}_j, \ \forall 0 \leq i \leq j \leq n$$

2.3. Structural Identifiability

Definition 2.4. Structural identifiability domain. Consistently with Raue et al. (2009) and Massonis and Villaverde (2020), let the structural identifiability domain Θ_S of a model $\mathcal{M} = \{\Theta, \pi, \mathcal{L}\}$ be defined as:

$$\Theta_S = \{\theta \in \Theta \mid \forall \theta' \in \Theta, \ \theta \neq \theta' \iff p(\mathcal{D}|\theta, \mathcal{M}) \neq p(\mathcal{D}|\theta', \mathcal{M})\} \tag{10}$$

Similarly, $\mathcal{M}$ is said to be structurally identifiable if $\Theta = \Theta_S$.

If θ is in the structural identifiability domain of $\mathcal{M}$, it can be uniquely identified from $p(\mathcal{D}|\theta, \mathcal{M})$. For instance, a Gaussian distribution of mean μ and variance σ^2 is uniquely defined by its parameters $\theta = (\mu, \sigma^2)$. Its structural identifiability domain is thus $\Theta_S = \mathbb{R} \times \mathbb{R}^+$. Similarly, if $N \neq 0$, $p \neq 0$, $p \neq 1$, and $q \neq 0$, the probability density of EPSCs under the binomial model without short-term plasticity $\mathcal{M}_1$ is structurally identifiable if we restrict Θ_1 to $\mathbb{N}^* \times]0, 1[\times (\mathbb{R}_+^*)^2$ (**Figure 2**).

2.4. Informative Domain

In some regimes, parameters may not be precisely inferred from observations, even though the model is otherwise structurally identifiable. Indeed, in practice we usually only have access to a finite number of (possibly noisy) observations. *Practical* identifiability thus differs from the *structural* identifiability defined in section 2.3.

A definition for the practical identifiability of a parameter has previously been proposed in Raue et al. (2009), along with an approach for detecting practical non-identifiabilities based on the profile likelihood (Venzon and Moolgavkar, 1988; Murphy and Van der Vaart, 2000). The authors first define the likelihood-based confidence intervals for the estimator $\hat{\theta}_i$ of the i-th

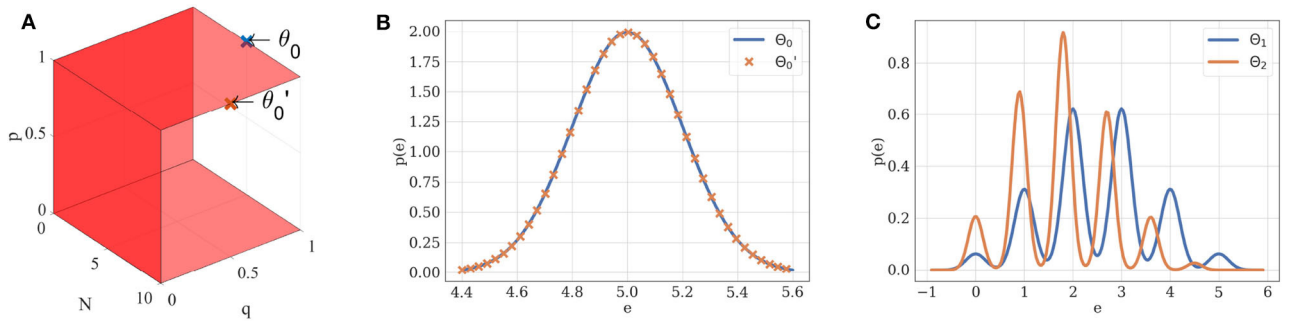

FIGURE 2 | (A) Structural identifiability domain of the binomial model $\mathcal{M}_1$. This model, parameterized by $\theta = (N, p, q, \sigma)$, is structurally identifiable if and only if $N \neq 0$, $p \neq 0$, $p \neq 1$, and $q \neq 0$. These conditions are represented by the red hyperplanes in the $N \times p \times q$ domain. **(B)** Two different sets of parameters θ_0 and θ_0' may lead to the same distribution of observations if taken out of the structural identifiability domain. If $p = 1$, the distribution of EPSCs under model $\mathcal{M}_1$ follows a Gaussian law of variance σ^2 and mean Nq: different combinations of N and q can thus ambiguously describe it. Blue distribution: $N = 5$, $p = 1$, $q = 1$, $\sigma = 0.2$. Orange distribution: $N = 10$, $p = 1$, $q = 0.5$, $\sigma = 0.2$. **(C)** Two different sets of parameters θ_1 and θ_2 will lead to different distributions when taken within the structural identifiability domain of $\mathcal{M}_1$: its distribution is uniquely defined by its parameters. Blue distribution: $N = 5$, $p = 0.5$, $q = 1$, $\sigma = 0.2$. Orange distribution: $N = 5$, $p = 0.4$, $q = 0.9$, $\sigma = 0.15$.

parameter of a model $\mathcal{M}$:

$$C_{i,\Delta} = \{\theta_i \mid \mathcal{L}(\hat{\theta}_i|\mathcal{D}) - \mathcal{L}(\theta_i|\mathcal{D}) < \Delta\}$$

where

$$\mathcal{L}(\theta_i|\mathcal{D}) = \max_{\theta_{j\neq i}} \mathcal{L}(\theta|\mathcal{D})$$

for a given threshold Δ. Then, they propose the following definition: *A parameter estimate $\hat{\theta}_i$ is practically non-identifiable, if the likelihood-based confidence region is infinitely extended in increasing and/or decreasing direction of θ_i, although the likelihood has a unique minimum for this parameter*, meaning that the decrease in likelihood compared to the optimal parameters estimate stays below the threshold Δ in direction of θ_i. When plotting the likelihood as a function of the parameters, practical non-identifiability can be seen as an infinitely extended flat valley, in which the decrease in likelihood stays below Δ. The authors also describe an algorithm for computing the profile likelihood and hence detecting such practical non-identifiabilities: *Structural non-identifiable parameters are characterized by a flat profile likelihood. The profile likelihood of a practically non-identifiable parameter has a minimum, but is not excessing a threshold Δ for increasing and/or decreasing values of θ_i* (see Figure 3 in Raue et al., 2009).

An important limitation of this definition is to be data-dependent: it only holds for a specific set of recorded data $\mathcal{D}$. Indeed, likelihood-based confidence intervals, and hence practical identifiability, are defined with respect to a certain data set $\mathcal{D}$, and may thus vary for different realizations of the experiment. However, an identifiability criterion can be made data-independent by averaging it over all possible realizations of $\mathcal{D}$, i.e., by computing its expectation with respect to the distribution $p(\mathcal{D}|\theta^*, \mathcal{M}, \Psi)$. Such an averaged criterion would correspond to the *a priori* expected identifiability before a specific $\mathcal{D}$ is recorded.

Practical information about θ is a function of the experimental protocol Ψ: for a given Ψ, the informative domain $\Theta_I(\Psi)$ of a model $\mathcal{M}$ could be defined based on the variance of the estimator. For instance, in a Bayesian setting, the domain $\Theta_I(\Psi)$ could be the set of parameters for which the expected informativeness of the posterior distribution of the parameters (measured as the Kullback-Leibler divergence between the posterior and the prior) is above a threshold Δ:

$$\Theta_I(\Psi) = \{\theta^* \in \Theta \mid \langle D_{KL}(p(\theta|\mathcal{D}, \mathcal{M}, \Psi) \,||\, p(\theta|\mathcal{M}))\rangle_{p(\mathcal{D}|\theta^*, \mathcal{M}, \Psi)} \geq \Delta\} \tag{11}$$

Although data-independent, this definition suffers from the same limitation as the one proposed in Raue et al. (2009): it requires to set a specific threshold Δ. Instead of defining an arbitrary criterion Δ on the possible precision of parameters estimate, we will derive our definition from a model selection argument.

2.5. Model Selection

In model selection, the plausibility of two competing models $\mathcal{M} = \{\Theta, \pi, \mathcal{L}\}$ and $\mathcal{M}' = \{\Theta', \pi', \mathcal{L}'\}$ based on observations $\mathcal{D}$ can be assessed using the Bayes Factor (Kass and Raftery, 1995):

$$B_{\mathcal{M},\mathcal{M}'}(\mathcal{D}) = \frac{p(\mathcal{D}|\mathcal{M})}{p(\mathcal{D}|\mathcal{M}')} = \frac{\int_\Theta \mathcal{L}(\theta|\mathcal{D})\pi(\theta)d\theta}{\int_{\Theta'} \mathcal{L}'(\theta|\mathcal{D})\pi'(\theta)d\theta} \tag{12}$$

If the Bayes Factor is superior to 1, then the evidence for $\mathcal{M}$ is higher than the evidence for $\mathcal{M}'$. It is worth pointing out that the Bayes Factor will not only favor models which provide a good fit to the data, but also includes a tendency to favor simpler models, a natural form of Occam's Razor (Jefferys and Berger, 1991; MacKay and MacKay, 2003). Indeed, a complex model (i.e., a model with many independent parameters or with a broader prior for its parameters) will be able to explain a larger set of possible observed data than a simple model; but this comes at the price of spreading its likelihood over a larger set of possible outcomes. Hence, if two models fit the observed data equally well, the simpler one will be favored.

2.6. Proposed Definition of Practical Model Identifiability

To compute the identifiability domain of any model $\mathcal{M}$ compared to another model $\mathcal{M}'$, we introduce the Average Log

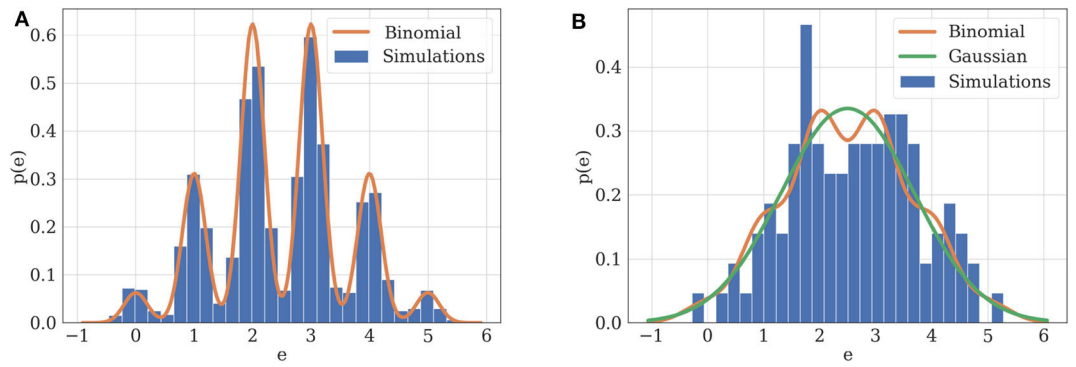

FIGURE 3 | (A) Illustration of practical identifiability. Orange: theoretical distribution for model $\mathcal{M}_1$ parameterized with $N = 5$, $p = 0.5$, $q = 1$, $\sigma = 0.2$. Blue: histogram of 2,000 simulated EPSCs generated using the same parameters. Parameters can be precisely inferred from the observations, which fit their theoretical distribution. **(B)** Illustration of practical non-identifiability. Blue: histogram of 100 simulated EPSCs with $N = 5$, $p = 0.5$, $q = 1$, $\sigma = 0.4$. Due to the small number of data points and high recording noise σ, the binomial parameters can only be loosely estimated, which is characterized by the fact that a Gaussian distribution (green) will provide a better fit to the data than a binomial distribution (orange).

Bayes Factor:

$$\mathcal{B}_{\mathcal{M},\mathcal{M}'}(\theta^*, \Psi) = \langle \log B_{\mathcal{M},\mathcal{M}'}(\mathcal{D}) \rangle_{p(\mathcal{D}|\theta^*,\mathcal{M},\Psi)} \tag{13}$$

For a given parameter θ^* and protocol Ψ, model $\mathcal{M}$ is said to be practically identifiable compared to $\mathcal{M}'$ if $\mathcal{B}_{\mathcal{M},\mathcal{M}'}(\theta^*, \Psi) \geq 0$. Intuitively, the identifiability domain of $\mathcal{M}$ compared to $\mathcal{M}'$ corresponds to all the settings (parameters and protocols) for which, on average, data generated from the ground truth $\mathcal{M}$ will be better explained by $\mathcal{M}$ than by $\mathcal{M}'$.

In contrast to the definition in Raue et al. (2009), our proposed definition does not require to set a (possibly arbitrary) threshold Δ. Instead, it is derived from a model selection criterion. We argue that the parameters of a model $\mathcal{M}$ are practically identifiable if $\mathcal{M}$ is itself practically identifiable. In some settings (as for the nested models of chemical synapse described in section 2.1), a family of submodels might naturally arise, while the choice of a threshold Δ would be arbitrary.

Another interest of our approach is to be data-independent, while the definition proposed in Raue et al. (2009) only holds for a specific set of recorded data $\mathcal{D}$. Indeed, we define practical identifiability as a data-independent and intrinsic property of the model $\mathcal{M}$ and experimental protocol Ψ. As the log-Bayes Factor in (13) is averaged over all possible realizations of $\mathcal{D}$, it corresponds to the *a priori* expected identifiability before $\mathcal{D}$ is recorded. Our approach thus allows to define practical identifiability *domains*:

Definition 2.5. Practical identifiability domain. Consider a model $\mathcal{M} = \{\Theta, \pi, \mathcal{L}\}$ and a submodel $\mathcal{M}' = \{\Theta', \pi', \mathcal{L}'\}$ of $\mathcal{M}$. For a given experimental protocol Ψ, the practical identifiability domain $\Theta_P(\Psi)$ of $\mathcal{M}$ is the set of parameters θ^* for which it is identifiable compared to its submodel:

$$\Theta_P(\Psi) = \{\theta^* \in \Theta \mid \mathcal{B}_{\mathcal{M},\mathcal{M}'}(\theta^*, \Psi) \geq 0\} \tag{14}$$

Note that in the limit where the priors π and π' are highly peaked (i.e., $\pi(\theta) = \delta(\theta - \bar{\theta})$ and $\pi'(\theta) = \delta(\theta - \bar{\theta}')$), the condition $\mathcal{B}_{\mathcal{M},\mathcal{M}'}(\theta^*, \Psi) \geq 0$ is always satisfied due to Gibbs' inequality. In this case we have $\Theta_P(\Psi) = \Theta, \forall \Psi$. However, generically the condition $\mathcal{B}_{\mathcal{M},\mathcal{M}'}(\theta^*, \Psi) \geq 0$ is not always satisfied since $p(\mathcal{D}|\mathcal{M})$ is not equal to $p(\mathcal{D}|\theta^*, \mathcal{M})$. The latter is the probability of observing $\mathcal{D}$ given $\mathcal{M}$ and a certain parametrization θ^*, while the former is the marginal likelihood over all parameters (12).

Two examples can illustrate this correspondence between model selection and parameter inference. Consider first the case of data recorded from $\mathcal{M}_1$. If the experimental protocol is not sufficiently informative (i.e., if data are scarce or noisy), not only will the inference of synaptic parameters be poor, but a Gaussian distribution will also provide a better fit than a binomial release model to the data. Indeed, as $[e_i|k_i] \sim \mathcal{N}(qk_i, \sigma)$, in the absence of recording noise (i.e., if $\sigma = 0$), the distribution of EPSCs is a series of Dirac delta functions located at each multiple of the quantal size qk for $k \in \{0, 1, \ldots, N\}$. In this ideal case, q is clearly identifiable (**Figure 3A**). However, upon addition of a recording noise of amplitude σ, EPSCs are normally distributed around qk for $k \in \{0, 1, \ldots, N\}$, and the peaks on the histogram corresponding to each multiple of the quantal size might overlap if σ is sufficiently high with respect to q (**Figure 3B**).

Similarly, we can consider the example of a synapse which shows short-term depression (STD) with a time constant τ_D (model $\mathcal{M}_2$). If the presynaptic cell is stimulated with an inter-spike intervals longer than τ_D, no depression will be visible in the recorded data, and the true model with STD will not be identifiable from a simpler binomial model without STD. In the same time, it will impossible to correctly infer the value of τ_D.

Our proposed definition of practical identifiability and of the identifiability domain of a model extend the landscaping technique introduced in Navarro et al. (2004) as well as the framework for testing identifiability of Bayesian models introduced in Acerbi et al. (2014). Especially, comparing the expected supports $\langle \log p(\mathcal{D}|\mathcal{M}) \rangle_{p(\mathcal{D}|\theta^*,\mathcal{M},\Psi)}$ and $\langle \log p(\mathcal{D}|\mathcal{M}') \rangle_{p(\mathcal{D}|\theta^*,\mathcal{M},\Psi)}$ of $\mathcal{M}$ and $\mathcal{M}'$ (given that values

are averaged over $\langle\cdot\rangle_{p(\mathcal{D}|\theta^*,\mathcal{M},\Psi)}$) allows us to define a quantitative criterion for identifiability.

The model evidence $p(\mathcal{D}|\mathcal{M})$ in (12) is often intractable in practice for complex models, as it requires to integrate marginals for each parameter. Different methods have been proposed to approximate it: MCMC computations (Weinberg et al., 2012), Savage-Dickey method (Wagenmakers et al., 2010), supermodels (Mootoovaloo et al., 2016). A practical and time-efficient approximation of the model evidence is given by the Bayesian Information Criterion $\mathrm{BIC}_{\mathcal{M}}(\mathcal{D})$ (Schwarz et al., 1978):

$$\mathrm{BIC}_{\mathcal{M}}(\mathcal{D}) = -2\log \mathrm{p}(\mathcal{D}|\hat{\theta},\mathcal{M}) + \mathrm{k}_{\mathcal{M}}\log(\mathrm{T}) \approx -2\log \mathrm{p}(\mathcal{D}|\mathcal{M}) \tag{15}$$

where $\hat{\theta} = \arg\max_\theta \mathcal{L}(\theta|\mathcal{D})$ is the maximum likelihood estimator (MLE) of $\mathcal{L}(\theta|\mathcal{D})$, $k_{\mathcal{M}} = \dim(\Theta)$ is the number of independent parameters of $\mathcal{M}$, and $T = |\Psi|$ is the number of data points in $\mathcal{D}$. A detailed derivation is provided in **Supplementary Material**. The BIC is the sum of two terms: a likelihood term $-2\log p(\mathcal{D}|\hat{\theta},\mathcal{M})$ which represents the ability of the model $\mathcal{M}$ to explain $\mathcal{D}$, and a penalty term $k_{\mathcal{M}}\log(T)$ which favors simpler models, as explained in section 2.5.

The BIC is commonly used as an approximation of the model evidence $p(\mathcal{D}|\mathcal{M})$ in model selection: the model with the lowest BIC is preferred over the others. The main advantage of using the BIC is to transform a complex integration problem (i.e., the computation of $p(\mathcal{D}|\mathcal{M})$) into a simpler optimization problem (i.e. the computation of $\hat{\theta}$). Besides, it allows to perform model selection without the need to specify a prior for the parameters, and is thus a popular tool for model selection (Daw et al., 2011).

As stated in **Supplementary Material**, the approximation $\mathrm{BIC}_{\mathcal{M}}(\mathcal{D}) \approx -2\log \mathrm{p}(\mathcal{D}|\mathcal{M})$ is only valid under the hypothesis that data points are independent and identically distributed (i.i.d.), which is not the case for models with short-term plasticity. If data are correlated, we are left with the following approximation, which does not simplify in the general case:

$$-2\log p(\mathcal{D}|\mathcal{M}) \approx -2\log p(\mathcal{D}|\hat{\theta},\mathcal{M}) + \log(|H(\hat{\theta})|) \tag{16}$$

where $H(\hat{\theta})$ is the Hessian matrix of $-\log p(\mathcal{D}|\theta,\mathcal{M})$ in Equation (15).

We emphasize that the classical definition of the BIC (15) should not be used if observations are correlated. Here, for models in which output are not independent, we use the approximation given by Equation (16), in which the term $k_{\mathcal{M}}\log(T)$ in the BIC is replaced by $\log(|H(\hat{\theta})|)$. In some settings, the computation of the Hessian matrix can be challenging. However, MCMC methods can be used to approximate $H(\hat{\theta})$, even without an explicit expression for the gradient of the function (Spall, 2005). In our case, a numerical method for computing $|H(\hat{\theta})|$ is detailed in the **Supplementary Material**.

Using approximation (15) in definition (14) yields the following approximation for the practical identifiability domain in case the model evidence $p(\mathcal{D}|\mathcal{M})$ in (12) is intractable:

$$\tilde{\Theta}_P(\Psi) = \{\theta^* \in \Theta \mid \langle \mathrm{BIC}_{\mathcal{M}}(\mathcal{D})\rangle_{\mathrm{p}(\mathcal{D}|\theta^*,\mathcal{M},\Psi)} \leq \langle \mathrm{BIC}_{\mathcal{M}'}(\mathcal{D})\rangle_{\mathrm{p}(\mathcal{D}|\theta^*,\mathcal{M},\Psi)}\} \tag{17}$$

3. RESULTS

3.1. Identifiability Domain of the Binomial Model Without Short-Term Plasticity

We study here the conditions under which a binomial model without short-term plasticity $\mathcal{M}_1$ can be correctly identified from a Gaussian model having the same mean and variance ($\mathcal{M}_0$). In order for the binomial model to be identifiable from a Gaussian quantum-less distribution, the recording noise needs to be sufficiently low compared to q for the peaks on the histogram of recorded EPSC to be identified. We will thus plot the identifiability domain as a function of the recording noise of amplitude σ for a fixed q. The identifiability domain corresponds to the points θ in the parameters space Θ_1 for which the average BIC of $\mathcal{M}_1$ over all possible outputs of $\mathcal{M}_1$ parameterized with θ is lower than the average BIC of $\mathcal{M}_0$.

Per se, the identifiability domain depends on all the parameters of $\mathcal{M}_1$, as well as on the experimental protocol. For simplicity and in order to obtain a plot in 2 dimensions, we will only plot it as a function of p and σ while holding other variables to a fixed value. For a given experimental setup Ψ (which encompasses only the number of recorded data points T, the inter-spike intervals playing no role in these models), the following Markov-Chain-Monte-Carlo (MCMC) procedure is implemented:

1. A set of values p^* and σ^* are chosen from the space of possible values for p and σ;
2. Using p^* and σ^*, 400 independent data sets $(\mathcal{D}_i)_{1\leq i\leq 400}$ are generated from $\mathcal{M}_1$. Each data set consists in T EPSCs;
3. For each $\mathcal{D}_i$, the BIC of both models are computed; these values are averaged over i to compute an average BIC and identifiability is assessed if $\mathcal{M}_1$ is preferred over $\mathcal{M}_0$, which corresponds to the black dots in **Figure 4A**.

The procedure of plotting a complete identifiability domain can be quite time-consuming. Indeed, it requires to span the entire space of parameters; for each vector θ^*, to generate a large number of independent data sets $(\mathcal{D}_i)$; and for each of these data sets, to compute the maximum likelihood estimator $\hat{\theta}$ using the Expectation-Maximization algorithm (Barri et al., 2016). Details on the computation of $\hat{\theta}$ are available in **Supplementary Material**.

However, as both models $\mathcal{M}_0$ and $\mathcal{M}_1$ generate i.i.d. data, and by making the approximation $\hat{\theta} \approx \theta^*$ (i.e., by assuming that the maximum likelihood estimator $\hat{\theta}$ will be close to the true value θ^* from which data were generated), the condition that model $\mathcal{M}_1$ is identifiable (17) can be approximated as follows:

$$-2T\int p(e|\theta^*,\mathcal{M}_1)\log p(e|\theta^*,\mathcal{M}_1)de + k_{\mathcal{M}_1}\log(T) \leq -2T\int p(e|\theta^*,\mathcal{M}_1)\log p(e|\hat{\theta}_{\mathcal{M}_0},\mathcal{M}_0)de + k_{\mathcal{M}_0}\log(T) \tag{18}$$

where $\hat{\theta}_{\mathcal{M}_0} = (\mu,\sigma^2)$ represents the mean $\mu \approx N^*p^*q^*$ and the variance $\sigma^2 \approx N^*p^*(1-p^*)q^{*2} + \sigma^{*2}$ of the data generated from $\mathcal{M}_1$.

The condition specified by inequality (18) can be checked for any point θ^* without the need to generate a large number of independent data sets nor to compute the estimator $\hat{\theta}$.

Solving (18) numerically for σ allows to draw the border of the identifiability domain of $\mathcal{M}_1$, represented as solid lines in **Figures 4A,B**.

Several points are worth highlighting. Firstly, **Figure 4A** shows a good agreement between the results of the MCMC simulations (black dots) and those from the semi-analytical method (18) (blue line). Secondly, as expected, **Figures 4A,B** illustrate that the identifiability domain increases with the number of data points T: intuitively, a larger data set facilitates the correct identification of a complex model. Besides, irrespective of the values of T and σ, for $p = 0$ and $p = 1$ the model $\mathcal{M}_1$ is structurally indistinguishable from a Gaussian distribution (see **Figure 2A**). Finally, the maximum noise σ which makes the binomial model $\mathcal{M}_1$ indistinguishable from a Gaussian distribution $\mathcal{M}_0$ is larger for extreme values of p (close to 0.9 or 0.1) than for $p = 0.5$. Indeed, in the latter case, the distributions of EPSC will be symmetric (as in the upper panel of **Figure 4C**), and hence just a little increase in recording noise will be enough to cover the inter-peak intervals and make the distribution Gaussian-shaped. In the former case, the distribution will be highly skewed, and thus difficult to approximate with a normal distribution.

The same approach can then be extended to more complicated models, by defining their identifiability domains as the part of the parameters plane where their average BIC will be lower than the BIC of a simpler one.

3.2. Identifiability Domain of the Binomial Model With Short Term Depression

We study here the conditions under which a binomial model with short-term depression ($\mathcal{M}_2$) can be correctly identified from a model without short-term plasticity ($\mathcal{M}_1$). In a first example, we assume that the presynaptic cell is stimulated at a constant inter-spike interval (ISI), which needs to be sufficiently short with respect to the time constant τ_D to make depression visible. We thus plot the identifiability domain as a function of both p and τ_D. We use the same method as in 3.1: For each set of parameters p^* and τ_D^*, 400 independent data sets are generated from $\mathcal{M}_2$. Both models $\mathcal{M}_2$ and $\mathcal{M}_1$ are fitted on them, and black dots in **Figure 5A** correspond to the parameters for which the average BIC of $\mathcal{M}_2$ is lower than the average BIC of $\mathcal{M}_1$.

As expected, we verify that the identifiability of $\mathcal{M}_2$ is only possible when τ_D is sufficiently long with respect to the inter-spike interval. Besides, if the release probability p is low, correlations between recordings will be weak and the effect of short-term depression will not be detectable. A major difference between models $\mathcal{M}_1$ and $\mathcal{M}_2$ is that, in the latter, observations $\{e_i\}_{1\leq i\leq T}$ are not i.i.d.. The value of the i-th recorded EPSC is a function of the number of available and released vesicles n_i and k_i, which in turn depend on their previous values and on the ISI Δt_i. This has two main consequences. Firstly, using the same approximation as in (18) would lead to a biased estimate of the identifiability domain. Secondly, the classical definition of the BIC (15) should not be used since observations are correlated. Rather, we use Equation (16) to compare the evidence for $\mathcal{M}_1$ and $\mathcal{M}_2$ for a given data set.

Plotting the identifiability domain of a model also allows to investigate how the identifiability depends on the experimental protocol. For model $\mathcal{M}_1$, we already saw that the identifiability domain increases with the number of data points T (see **Figures 4B,D**): a larger data set is more informative and allows for more reliable inference. In this case, T is the only experimental variable, as observations $\{e_i\}_{1\leq i\leq T}$ are i.i.d. On the other hand, the identifiability domain of $\mathcal{M}_2$ will depend not only on the number of data points, but also on the stimulation protocol. We compare the constant stimulation protocol (T data points with a constant inter-spike interval $ISI = 0.05s$) of **Figure 5A** with a more realistic stimulation protocol in **Figure 5B**. In electrophysiological recordings, synaptic transmission is classically studied by stimulating the presynaptic cell with short regular train of spikes at a given frequency, followed by a recovery spike. This protocol is then repeated several times (Costa et al., 2013; Barri et al., 2016; Bird et al., 2016). Such periodic trains are more informative than a constant stimulation protocol, as they allow to probe a broader range of temporal dynamics.

In **Figure 5B**, we use 20 repetitions of a train of 4 spikes at 20Hz ($ISI = 0.05s$), followed by a recovery spike $0.5s$ later. This protocol entails the same number of data points $T = 100$ as the constant one, but allows to identify STD for a broader range of depression time constants (namely, for $\tau_D < 0.3s$). On the other hand, since there are fewer successive stimulations within a short time interval than in the constant protocol, depression can only be identified when the release probability p is sufficiently high to induce vesicle pool depletion.

3.3. Data Free Model Selection

In model-based inference of synaptic parameters, a crucial step related to the estimation of the parameters is model selection, which is usually performed in several steps:

1. Data $\mathcal{D}$ are acquired from a synapse using protocol Ψ;
2. A nested family of $n+1$ possible models $\mathcal{F} = \{\mathcal{M}_0, \mathcal{M}_1, ..., \mathcal{M}_n\}$ is defined;
3. Each of these models is fitted on $\mathcal{D}$ to obtain $n+1$ MLE $\hat{\theta}_0, \hat{\theta}_1, ..., \hat{\theta}_n$;
4. A model selection criterion (Bayes Factor, BIC, AIC...) is computed to quantify and rank the fitness of each model on $\mathcal{D}$;
5. If $\mathcal{M}_i$ is the selected model, then its MLE $\hat{\theta}_i$ is selected as the inference of synaptic parameters.

However, in many studies (Barri et al., 2016; Bird et al., 2016; Ghanbari et al., 2017), such a model selection step is not described. In this section, we investigate the possibility, having only access to the inferred values $\hat{\theta}$ of the parameters and to the description of the experimental protocol Ψ, to verify that the model used to infer $\hat{\theta}$ was indeed practically identifiable (i.e., to verify if a simpler model would have given a better fit to the data).

We use the notation $\mathcal{D}$ for a set of data generated from a model $\mathcal{M}$ parameterized with θ^*, and $\hat{\theta}$ the inferred parameters obtained by fitting the parameters θ of model $\mathcal{M}$ on $\mathcal{D}$. If θ^* is within the practical identifiability domain of $\mathcal{M}$ as we defined it, it is then possible to correctly infer it from $\mathcal{D}$, and hence $\hat{\theta} \approx \theta^*$

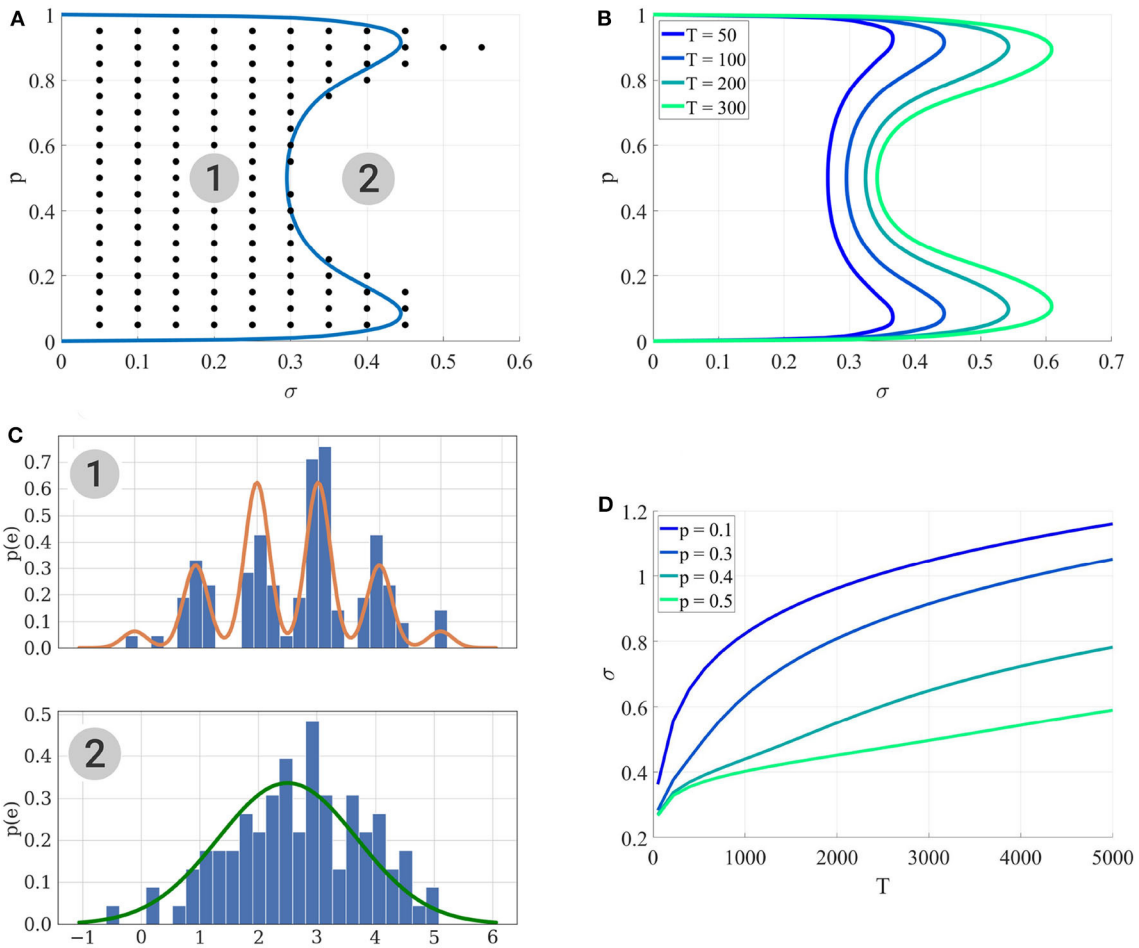

FIGURE 4 | (A) Identifiability domain of $\mathcal{M}_1$ as a function of p and σ. Blue line: domain of identifiability from Equation (18). On the left part of the blue line, the recording noise σ is sufficiently low to identify $\mathcal{M}_1$. Black dots: values (p^*, σ^*) for which the average BIC of 400 data sets results in the correct identifiability of $\mathcal{M}_1$. Results obtained for $N = 5$, $q = 1$, and $T = 100$. **(B)** Identifiability domain of the binomial model $\mathcal{M}_1$ compared to a Gaussian distribution $\mathcal{M}_0$, computed from (18), for different values of T. **(C)** To visualize the effect of σ on the data, this panel shows histograms of data generated from $\sigma = 0.2$ (1) and $\sigma = 0.4$ (2), alongside with their theoretical distribution from Equation (8) (orange line in the upper panel) or when a Gaussian distribution is fitted on them (green line in the lower panel). In the identifiability domain (1), quantal peaks are clearly visible. Outside of the identifiability domain (2), the binomial distribution becomes Gaussian-shaped. **(D)** Another visualization of the identifiability domains displayed in **(A,B)**. For different values of p, the maximum recording noise σ (i.e., the boundary of the identifiability domain) is plotted as a function of the number of data points T. The identifiability domain increases with T: intuitively, a larger data set facilitates the correct identification of a complex model.

will also be within the identifiability domain of $\mathcal{M}$. Reciprocally, if $\hat{\theta}$ is not in the identifiability domain of $\mathcal{M}$, then a submodel would have provided a better fit to the data $\mathcal{D}$ than $\mathcal{M}$. Is it thus possible to verify if $\mathcal{M}$ overfits the data simply by verifying if $\hat{\theta}$ is in its identifiability domain, without having access to the data.

This is illustrated in **Figures 6A,B**, where $\mathcal{M}_1$ is fitted on data generated from its submodel $\mathcal{M}_0$. For six different values of θ_0^* (**Figure 6A**), the inferred parameters are out of the identifiability domain of $\mathcal{M}_1$ (**Figure 6B**), showing that data are indeed better explained by $\mathcal{M}_0$ than by $\mathcal{M}_1$.

3.3.1. First Example: Application to the Data From Katz et al. (1954)

We first apply our *data free* model selection method to the seminal 1954 paper from Del Castillo and Katz (1954), in which the quantal nature of neurotransmitter release is identified for the first time. In order to observe mEPSP, they artificially reduced the release probability p by lowering the external calcium concentration. Although the quantal components of postsynaptic potentials are clearly visible and thoroughly analyzed, it would be interesting to verify, using our proposed model identifiability analysis method, that the binomial model (i.e., a multi-quantal distribution) indeed provides a better fit to the data than a simpler Gaussian model (i.e., a uni-quantal distribution).

Data (Fatt and Katz, 1952) consist in 328 EPSPs recorded at the neuro-muscular junction (NMJ) of a frog muscle. Fitting the binomial model and running the Expectation-Maximization algorithm on them yields $\hat{N} = 42$, $\hat{p} = 0.013$, $\hat{q} = 0.875\ mV$, and $\hat{\sigma} = 0.15\ mV$ (and hence $\frac{\hat{\sigma}}{\hat{q}} \approx 17\%$). For this particular example, we have not only access to the inferred parameters $\hat{\theta}$,

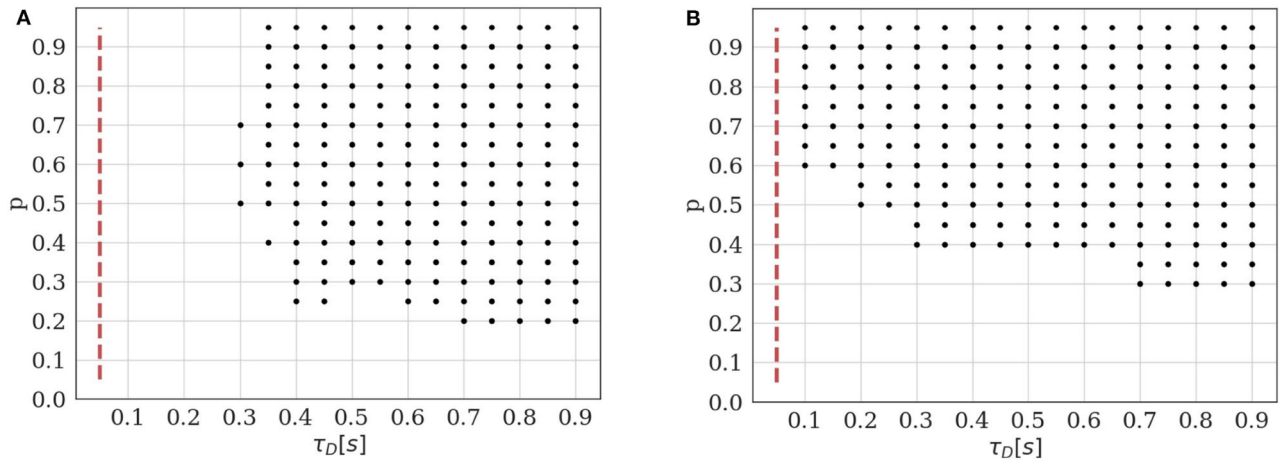

FIGURE 5 | Identifiability domain of the binomial model with short-term depression $\mathcal{M}_2$ as a function of p and τ_D. Black dots correspond to the parameters for which the average BIC of 400 data sets results in the correct identifiability of the depressed model. Results obtained for $N = 5$, $q = 1$, $\sigma = 0.2$, and $T = 100$. **(A)** Constant stimulation protocol with an inter-spike interval $ISI = 0.05s$ (red dotted line). **(B)** Stimulation protocol consisting in 20 repetitions of the same spike train: 4 spikes with an inter-spike interval $ISI = 0.05s$ followed by a recovery spike $0.5s$ later.

but also to the data: it is thus possible to directly compare the BIC of a Gaussian ($\text{BIC}_{\mathcal{M}_0} = 764.95$) and of a binomial ($\text{BIC}_{\mathcal{M}_1} = 470.37$) distributions, which indeed confirms that data are better explained by the binomial quantal model.

However, even without the data, we can verify that the point in the parameter-protocol space specified by $\hat{\theta}$ (the inferred values of the parameters) and Ψ (the number of data points $T = 328$) is indeed within the identifiability domain of the binomial model $\mathcal{M}_1$ compared to $\mathcal{M}_0$ (see **Figure 6C**), thus confirming the multi-quantal nature of the recordings.

3.3.2. Second Example: Application to the Data From Barri et al. (2016)

We then apply our method to the results presented in the 2016 paper from Barri et al. (2016), in which the complete binomial model (with STD and STF) is fitted on recordings from layer 5 pyramidal neurons. They use a slight variation of the binomial release model with short term plasticity described by Equations (1)–(6), in which the emission probability does not follow a Gaussian, but an inverse Gaussian distribution:

$$p_\theta(e_i|k_i) = \frac{q^{3/2}k_i}{\sqrt{2\pi\sigma^2 e_i^3}} \exp\left(-\frac{q(e_i - qk_i)^2}{2\sigma^2 e_i}\right) \tag{19}$$

To verify that the data would not have been better fitted by a simpler model (and hence, that the published estimates of synaptic parameters are reliable), 100 synthetic data sets were generated from the complete binomial model using the stimulation protocol and the inferred values of the parameters described in (Barri et al., 2016):

- 20 repetitions of the same stimulation protocol consisting in 8 presynaptic spikes at 20Hz followed by a recovery spike 500 ms later;
- $N^* = 17$, $p^* = 0.27$, $q^* = 0.18\ mV$, $\sigma^* = 0.06\ mV$, $\tau_D^* = 202\ ms$, and $\tau_F^* = 449\ ms$.

$\mathcal{M}_0$, $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$ were then fitted on the generated data. Average values of their respective BIC are presented in **Figure 7**, and confirm the identifiability of the model used in the study.

4. DISCUSSION

Obtaining an accurate estimate of the parameters of a system from noisy and scarce observations is a crucial problem in neuroscience. Especially, different methods have been proposed for estimating the parameters describing a synapse (namely, its number of independent release sites, their release probability upon the arrival of a presynaptic spike, the quantum of current elicited by one release event, the time constants of depression and facilitation, etc...). Inferring their values allows to analyze the locus of synaptic plasticity and homeostasis; to study possibly synapse-related diseases; and more generally to investigate learning, memory, and neural dynamics, which are mediated by synaptic transmission.

It is usually impossible to measure directly these parameters. However, they can be estimated by fitting a biophysical model of synapse on currents recorded on the post-synaptic side and elicited by experimental stimulation of the presynaptic cell. This approach for estimating the parameters of a system is referred to as model-based inference. As different competing models may be used to describe the system and explain its output, model-based inference of parameters thus raises the question of what makes a good model.

Prior to any data recording, a required property for competing models is identifiability. Although structural identifiability has been widely studied, no quantitative criterion exists for practical identifiability, which is usually only qualitatively assessed. Here, we propose a definition for the practical identifiability of a

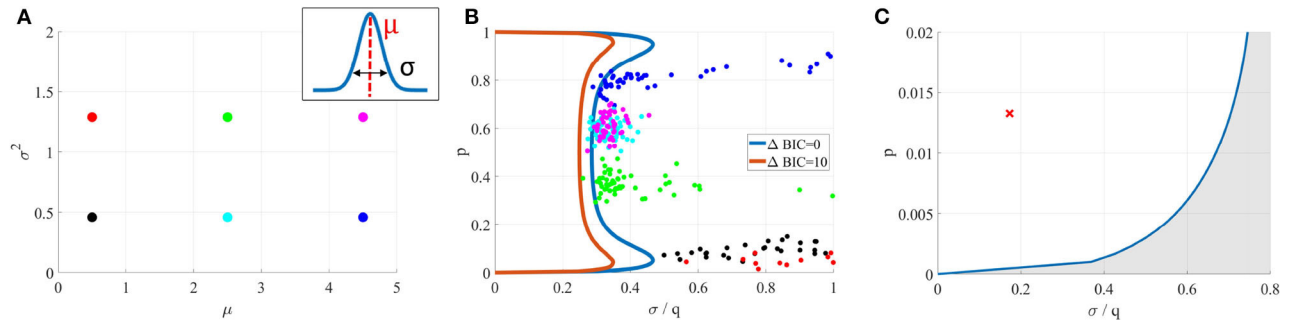

FIGURE 6 | (A) Data sets were generated from a Gaussian model $\mathcal{M}_0$ (for six different means μ and variances σ^2). **(B)** A binomial model $\mathcal{M}_1$ was fitted on them. In each case, inferred parameters (colored dots) are out of the identifiability domain of $\mathcal{M}_1$. It is thus possible to verify if a model used to fit data was indeed identifiable, without having access to the data and only using inferred parameters. **(C)** Blue line: identifiability domain of the binomial model compared to a Gaussian distribution, for $N = 42$ and $T = 328$, computed from (18). The red cross corresponds to the parameters inferred from (Del Castillo and Katz, 1954), and is indeed within the identifiability domain.

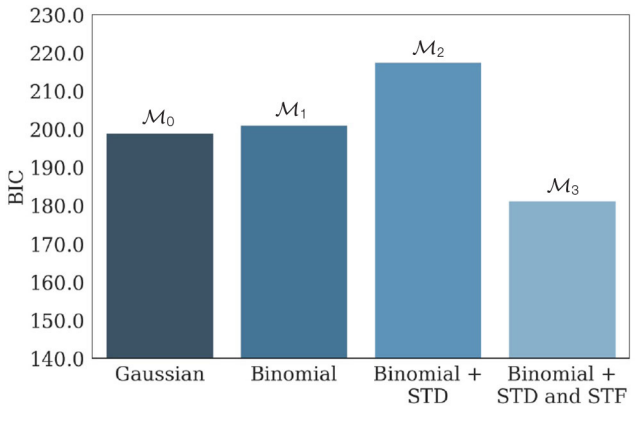

FIGURE 7 | Average BIC of $\mathcal{M}_0$, $\mathcal{M}_1$, $\mathcal{M}_2$, and $\mathcal{M}_3$ when fitted on 100 independent data sets generated from $\mathcal{M}_3$ parameterized with $N^* = 17$, $p^* = 0.27$, $q^* = 0.18\,mV$, $\sigma^* = 0.06\,mV$, $\tau_D^* = 202\,ms$, and $\tau_F^* = 449\,ms$. $\mathcal{M}_3$ has the lowest average BIC compared to its submodels, showing that the parameters used to generate the data are indeed within the identifiability domain of $\mathcal{M}_3$. As a consequence, we can infer that $\mathcal{M}_3$ indeed provided the best fit to the data compared to its submodels, and that inferred parameters presented in Barri et al. (2016) are reliable. The facilitating nature of the synapse is illustrated by the fact that the BIC of $\mathcal{M}_2$ (the model with only STD and no STF) is substantially larger than the one of $\mathcal{M}_3$.

model, based on its expected support given the distribution of the data. We define the practical identifiability domain of a statistical model as the set of parameters for which the model is correctly identified as the ground truth compared to a simpler alternative submodel, and we study the identifiability domains of different models of synaptic release. In the process, we propose an extension of the Bayesian Information Criterion (BIC) for models with correlated data. The BIC is a widely used tool for model selection, but it is derived by assuming that the outputs of the system are mutually independent, which is not the case for models of chemical synapse. Finally, we show that our approach allows to perform *data free* model selection, i.e., to verify the identifiability of a model without having access to the data.

The definition of practical identifiability we introduced here differs from the influential contribution of Raue et al. (2009) in two ways. Firstly, our definition is data-independent: it does not only hold for a specific set of recorded data $\mathcal{D}$. Indeed, we define practical identifiability as an intrinsic property of the model $\mathcal{M}$ and experimental protocol Ψ. We actually define the *a priori* expected identifiability before a specific $\mathcal{D}$ is recorded, which allows to study how identifiability is affected by different experimental protocols. Secondly, since our definition is derived from a model-selection argument, it does not require to select a possibly arbitrary threshold on the practical identifiability of parameters. Rather, it is defined with respect to a particular submodel. Although the choice of the submodel might itself be arbitrary, we argue that nested models and families naturally arise in commonly used statistical techniques, such as polynomial regression (Edwards and MacCallum, 2012), or Generalized Linear Models (GLM) (Pillow et al., 2008). Especially, the widespread use of phenomenological models in neuroscience (Kobayashi et al., 2009; Melanson et al., 2014; Wang et al., 2016; Levenstein et al., 2020) makes the use of nested families and submodels relevant.

Another limitation of our approach is its practical implementation. As mentioned, the model evidence $p(\mathcal{D}|\mathcal{M})$, on which our definition is based, is often intractable in practice for complex models, and needs to be estimated. For practical purpose, we used the Bayesian Information Criterion (BIC) to compute the identifiability domains of our different models of synapse. However, we acknowledge that the BIC only provides a valid approximation of the model evidence when the number of samples is sufficiently large. A future step would be to study the robustness of our approach to different computations of the model evidence or to other approximations, such as the Akaike Information Criterion (AIC) (Burnham and Anderson, 2004).

Our identifiability domains are similar to the approach adopted in Koyama (2012), in which the authors study under which regime of rate fluctuation are the temporal variations of a neuron firing rate correctly identified. Spike trains are generated from a model of spiking neuron with a fluctuating

firing rate (complex model); but under a certain value of rate fluctuation, this model becomes indistinguishable from a model of spiking neuron with a constant rate (simple model). Plotting the identifiability of the fluctuating-rate model as a function of the amplitude of rate fluctuation allows them to identify which distribution of inter-spike intervals has the broader identifiability domain (and thus maximizes the efficiency of rate fluctuation transmission).

In model-based inference and parameter estimation, one is often interested in obtaining theoretical bounds on the achievable error performance. Such theoretical bounds allow to assess *a priori* the possibility to correctly infer the parameters. A well-known theoretical result is the Cramér-Rao bound (Van Trees, 2004; Van Trees and Bell, 2007), which provides a lower bound on the variance of the parameter estimator. This bound, which depends on the model, its parameters, and the experimental protocol, may actually be too loose in practice, and does not account for the threshold effect described in Kostal et al. (2015). In many cases, as the number of data points increases, the estimate error displays a threshold-like transition, from a region of low performance to a region of high performance where the Cramér-Rao bound is attained. Our definition of practical identifiability also discriminates between regions of low information (for small signal-to-noise ratios and sample size) and high accuracy, provides a quantitative criterion to discriminate them, and can be extended to the case of non-i.i.d. data. An interesting future step would be to verify how the boundaries of our proposed identifiability domains compare with the transition threshold described in Kostal et al. (2015).

An interesting topic would be to study the practical identifiability domain as the number of observations T goes to infinity. In this asymptotic case, practical non-identifiability means that the model cannot be identified, even with an infinite amount of data. We can conjecture that practical identifiability is equivalent to structural identifiability in this asymptotic case, as hinted by **Figure 4**: the identifiability domain increases with T. A future step would be to verify if the practical identifiability domain of a model is included in its structural identifiability domain, and how it behaves when the number of observations T goes to infinity.

We applied our analysis to four variants of the binomial model, of increasing complexity: a Gaussian model (i.e., a uni-quantal distribution); a binomial model without short-term plasticity; a binomial model with only short-term depression; and a binomial model with both short-term depression and facilitation. A future step would be to extend our analysis to further generalizations of the binomial model, in order to account for parameters heterogeneity. Especially, the binomial model assumes that the release probability and the quantal amplitude are identical for each release site. It is however possible to hypothesize that there are several pools of vesicles, each having different parameters (for instance a fast depleting pool and a slow depleting pool). There will be regimes in which those sub-pools can be detected and other in which the noise is too high or the experimental protocol not informative enough to identify them, which can be quantified using our definition of identifiability. Another possible generalization of the binomial model is to assume that the postsynaptic response to one vesicle release is not fixed, but follows for instance a Gamma distribution (Bhumbra and Beato, 2013) to account for variability in vesicles size and neurotransmitter content.

Model selection is not only a first step in model-based inference of synaptic parameters (as it is necessary to have a reliable estimates of the parameters), but also a tool to study the mechanisms of neurotransmitter release at a chemical synapse. An alternative hypothesis (e.g., *"this synapse shows short-term plasticity"*) can be compared to a null hypothesis (*"this synapse does not show short-term plasticity"*) by computing how well the complex model (i.e. with short-term plasticity) explains the behavior of the synapse compared to the simple model (i.e., without short-term plasticity). Testing models of growing complexity allows to study the nature of the synapse and to identify mechanisms of neurotransmitter release. But the possibility to correctly select the model that corresponds to the true behavior of the synapse will depend on its parameters and on the experimental protocol used to record data: there are regimes in which the specific features of a model do not appear in the data. Such regimes correspond to the identifiability domain of the model, and studying them allows to draw conclusions on the nature of the synapse.

As stated previously, the problem of inferring parameters from noisy and scarce observations is not restricted to synaptic parameters estimation, but is a crucial question in neuroscience. Our proposed methodology could also be applied to models of single neurons (Koch, 2004; Jolivet et al., 2008; Gerstner and Naud, 2009; Mensi et al., 2012), neural population dynamics (René et al., 2020), or calcium-driven vesicles fusion (Schneggenburger and Neher, 2000; Lou et al., 2005; Sun et al., 2007).

On a broader scale, instead of seeing parametric non-identifiability as a statistical problem, we could consider it as a biophysical feature. The total synaptic strength between two cells is a function of both presynaptic (N, p) and postsynaptic (q) parameters. Different combinations of these parameters could lead to the same average postsynaptic response: a presynaptic modification of the number of release sites N can be compensated by an inverse modification of the postsynaptic number of receptors affecting q. This combined effect of presynaptic and postsynaptic plasticity has been shown to enable reliable and flexible learning (Costa et al., 2015) and homeostatic modulation (Davis and Müller, 2015). More generally, the question of degeneracy, defined as the ability of different elements to perform the same function, could be addressed within the framework of identifiability analysis (Drion et al., 2015; Rathour and Narayanan, 2019).

Finally, our proposed definition of model identifiability is paving the way toward Optimal Experiment Design (OED) for model selection and parameter inference. The information conveyed by the data about the ground truth model and its

parameters depends on the experimental protocol: number of recorded data points, stimulation frequency, etc. The goal of OED is to optimize the experimental protocol in order to maximize the possibility to discriminate between competing models (Vanlier et al., 2014; Balietti et al., 2018) and the precision of the inference of their parameters. An OED for inferring the parameters of a given model maximizes the mutual information between the data and the parameters $\mathcal{I}(\mathcal{D}, \theta)$ (Huan and Marzouk, 2013). This quantity turns out to be equal to the expected gain in information about θ (defined as the Kullback-Leibler divergence between its prior and its posterior), on which our proposed definition of the informative domain (11) is based. Similarly, maximizing the Average Log Bayes Factor (13) is equivalent to maximizing the discriminability between the two models $\mathcal{M}$ and $\mathcal{M}'$, and hence finding an OED for model selection. As a thorough theoretical preliminary analysis of the properties of the competing models is a first step prior to model selection and parameter inference (Asprey and Macchietto, 2000), we believe that our theoretical contribution to model analysis will contribute to the development of OED techniques for synaptic transmission study.

DATA AVAILABILITY STATEMENT

Matlab and Python files are available in the following GitHub repository: https://github.com/camillegontier/identifiability_binomial.git.

AUTHOR CONTRIBUTIONS

CG and J-PP wrote the paper and derived the equations. CG performed the numerical simulations, made the figures, and wrote the MATLAB code and the Python code. J-PP conceived of the presented idea and supervised the project. All authors contributed to the article and approved the submitted version.

FUNDING

The work presented in this paper was supported by the Swiss National Science Foundation under grant number 31003A_175644 entitled Bayesian Synapse.

ACKNOWLEDGMENTS

We thank Simone Carlo Surace and Mihai A. Petrovici for the fruitful discussions. Figures were created using BioRender. Calculations were performed on UBELIX (http://www.id.unibe.ch/hpc), the HPC cluster at the University of Bern.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncom.2020.558477/full#supplementary-material

REFERENCES

Acerbi, L., Ma, W. J., and Vijayakumar, S. (2014). "A framework for testing identifiability of Bayesian models of perception," in *Advances in Neural Information Processing Systems* (Montréal, QC), 1026–1034.

Asprey, S., and Macchietto, S. (2000). Statistical tools for optimal dynamic model building. *Comput. Chem. Eng.* 24, 1261–1267. doi: 10.1016/S0098-1354(00)00328-8

Balietti, S., Klein, B., and Riedl, C. (2018). Fast model-selection through adapting design of experiments maximizing information gain. *arXiv preprint arXiv:1807.07024*.

Barri, A., Wang, Y., Hansel, D., and Mongillo, G. (2016). Quantifying repetitive transmission at chemical synapses: a generative-model approach. *eNeuro* 3, 1–21. doi: 10.1523/ENEURO.0113-15.2016

Bekkers, J., Richerson, G., and Stevens, C. (1990). Origin of variability in quantal size in cultured hippocampal neurons and hippocampal slices. *Proc. Natl. Acad. Sci. U.S.A.* 87, 5359–5362. doi: 10.1073/pnas.87.14.5359

Bhumbra, G. S., and Beato, M. (2013). Reliable evaluation of the quantal determinants of synaptic efficacy using Bayesian analysis. *J. Neurophysiol.* 109, 603–620. doi: 10.1152/jn.00528.2012

Bird, A. D., Wall, M. J., and Richardson, M. J. (2016). Bayesian inference of synaptic quantal parameters from correlated vesicle release. *Front. Comput. Neurosci.* 10:116. doi: 10.3389/fncom.2016.00116

Burnham, K. P., and Anderson, D. R. (2004). Multimodel inference: understanding AIC and BIC in model selection. *Sociol. Methods Res.* 33, 261–304. doi: 10.1177/0049124104268644

Bykowska, O. S., Gontier, C., Sax, A.-L., Jia, D. W., Llera-Montero, M., Bird, A. D., et al. (2019). Model-based inference of synaptic transmission. *Front. Synapt. Neurosci.* 11:21. doi: 10.3389/fnsyn.2019.00021

Calabresi, P., Picconi, B., Parnetti, L., and Di Filippo, M. (2006). A convergent model for cognitive dysfunctions in Parkinson's disease: the critical dopamine-acetylcholine synaptic balance. *Lancet Neurol.* 5, 974–983. doi: 10.1016/S1474-4422(06)70600-7

Costa, R. P., Froemke, R. C., Sjöström, P. J., and van Rossum, M. C. (2015). Unified pre-and postsynaptic long-term plasticity enables reliable and flexible learning. *Elife* 4:e09457. doi: 10.7554/eLife.11988

Costa, R. P., Mizusaki, B. E., Sjöström, P. J., and van Rossum, M. C. (2017a). Functional consequences of pre-and postsynaptic expression of synaptic plasticity. *Philos. Trans. R. Soc. B Biol. Sci.* 372:20160153. doi: 10.1098/rstb.2016.0153

Costa, R. P., Padamsey, Z., D'Amour, J. A., Emptage, N. J., Froemke, R. C., and Vogels, T. P. (2017b). Synaptic transmission optimization predicts expression loci of long-term plasticity. *Neuron* 96, 177–189. doi: 10.1016/j.neuron.2017.09.021

Costa, R. P., Sjostrom, P. J., and Van Rossum, M. C. (2013). Probabilistic inference of short-term synaptic plasticity in neocortical microcircuits. *Front. Comput. Neurosci.* 7:75. doi: 10.3389/fncom.2013.00075

Davis, G. W., and Müller, M. (2015). Homeostatic control of presynaptic neurotransmitter release. *Annu. Rev. Physiol.* 77, 251–270. doi: 10.1146/annurev-physiol-021014-071740

Daw, N. D. (2011). "Trial-by-trial data analysis using computational models," in *Decision Making, Affect, and Learning: Attention and Performance XXIII*, eds M. R. Delgado, E. A. Phelps, and T. W. Robbins (Oxford: Oxford University Press), 3–38. doi: 10.1093/acprof:oso/9780199600434.003.0001

Del Castillo, J., and Katz, B. (1954). Quantal components of the end-plate potential. *J. Physiol.* 124, 560–573. doi: 10.1113/jphysiol.1954.sp005129

Drion, G., O'Leary, T., and Marder, E. (2015). Ion channel degeneracy enables robust and tunable neuronal firing rates. *Proc. Natl. Acad. Sci. U.S.A.* 112, E5361–E5370. doi: 10.1073/pnas.1516400112

Edwards, M. C., and MacCallum, R. C. (2012). *Current Topics in the Theory and Application of Latent Variable Models*. New York, NY: Routledge. doi: 10.4324/9780203813409

Fatt, P., and Katz, B. (1952). Spontaneous subthreshold activity at motor nerve endings. *J. Physiol.* 117, 109–128.

Gerstner, W., and Naud, R. (2009). How good are neuron models? *Science* 326, 379–380. doi: 10.1126/science.1181936

Ghanbari, A., Malyshev, A., Volgushev, M., and Stevenson, I. H. (2017). Estimating short-term synaptic plasticity from pre-and postsynaptic spiking. *PLoS Comput. Biol.* 13:e1005738. doi: 10.1371/journal.pcbi.1005738

Ghanbari, A., Ren, N., Keine, C., Stoelzel, C., Englitz, B., Swadlow, H. A., et al. (2020). Modeling the short-term dynamics of in vivo excitatory spike transmission. *J. Neurosci.* 40, 4185–4202. doi: 10.1523/jneurosci.1482-19.2020

Gottman, J. M. (1995). *The Analysis of Change*. Mahwah, NJ: Psychology Press.

Guo, D., and Li, C. (2012). Stochastic resonance in Hodgkin-Huxley neuron induced by unreliable synaptic transmission. *J. Theor. Biol.* 308, 105–114. doi: 10.1016/j.jtbi.2012.05.034

Hines, K. E., Middendorf, T. R., and Aldrich, R. W. (2014). Determination of parameter identifiability in nonlinear biophysical models: a Bayesian approach. *J. Gen. Physiol.* 143, 401–416. doi: 10.1085/jgp.201311116

Huan, X., and Marzouk, Y. M. (2013). Simulation-based optimal Bayesian experimental design for nonlinear systems. *J. Comput. Phys.* 232, 288–317. doi: 10.1016/j.jcp.2012.08.013

Jefferys, W. H., and Berger, J. O. (1991). *Sharpening Ockham's Razor on a Bayesian Strop.* Technical Report, Department of Statistics, Purdue University.

Jolivet, R., Kobayashi, R., Rauch, A., Naud, R., Shinomoto, S., and Gerstner, W. (2008). A benchmark test for a quantitative assessment of simple neuron models. *J. Neurosci. Methods* 169, 417–424. doi: 10.1016/j.jneumeth.2007.11.006

Kass, R. E., and Raftery, A. E. (1995). Bayes factors. *J. Am. Stat. Assoc.* 90, 773–795. doi: 10.1080/01621459.1995.10476572

Katz, B. (1969). *The Release of Neural Transmitter Substances*. Liverpool: Liverpool University Press.

Kauer, J. A., and Malenka, R. C. (2007). Synaptic plasticity and addiction. *Nat. Rev. Neurosci.* 8, 844–858. doi: 10.1038/nrn2234

Kobayashi, R., Tsubo, Y., and Shinomoto, S. (2009). Made-to-order spiking neuron model equipped with a multi-timescale adaptive threshold. *Front. Comput. Neurosci.* 3:9. doi: 10.3389/neuro.10.009.2009

Koch, C. (2004). *Biophysics of Computation: Information Processing in Single Neurons*. Oxford: Oxford University Press.

Komorowski, M., Costa, M. J., Rand, D. A., and Stumpf, M. P. (2011). Sensitivity, robustness, and identifiability in stochastic chemical kinetics models. *Proc. Natl. Acad. Sci. U.S.A.* 108, 8645–8650. doi: 10.1073/pnas.1015814108

Kostal, L., Lansky, P., and Pilarski, S. (2015). Performance breakdown in optimal stimulus decoding. *J. Neural Eng.* 12:036012. doi: 10.1088/1741-2560/12/3/036012

Koyama, S. (2012). "Coding efficiency and detectability of rate fluctuations with non-poisson neuronal firing," in *Advances in Neural Information Processing Systems* (Stateline), 180–188.

Levenstein, D., Alvarez, V. A., Amarasingham, A., Azab, H., Gerkin, R. C., Hasenstaub, A., et al. (2020). On the role of theory and modeling in neuroscience. *arXiv preprint arXiv:2003.13825*.

Levy, W. B., and Baxter, R. A. (2002). Energy-efficient neuronal computation via quantal synaptic failures. *J. Neurosci.* 22, 4746–4755. doi: 10.1523/JNEUROSCI.22-11-04746.2002

Lou, X., Scheuss, V., and Schneggenburger, R. (2005). Allosteric modulation of the presynaptic Ca2+ sensor for vesicle fusion. *Nature* 435, 497–501. doi: 10.1038/nature03568

MacKay, D. J., and MacKay, D. J. (2003). *Information Theory, Inference and Learning Algorithms*. Cambridge: Cambridge University Press.

Massonis, G., and Villaverde, A. F. (2020). Finding and breaking lie symmetries: implications for structural identifiability and observability in biological modelling. *Symmetry* 12:469. doi: 10.3390/sym12030469

Melanson, A., Mejias, J. F., Jun, J. J., Maler, L., and Longtin, A. (2014). A phenomenological model for self-initiated movement in electric fish. *BMC Neurosci.* 15:P112. doi: 10.1186/1471-2202-15-S1-P112

Mensi, S., Naud, R., Pozzorini, C., Avermann, M., Petersen, C. C., and Gerstner, W. (2012). Parameter extraction and classification of three cortical neuron types reveals two distinct adaptation mechanisms. *J. Neurophysiol.* 107, 1756–1775. doi: 10.1152/jn.00408.2011

Mootoovaloo, A., Bassett, B. A., and Kunz, M. (2016). Bayes factors via savage-dickey supermodels. *arXiv preprint arXiv:1609.02186*.

Murphy, S. A., and Van der Vaart, A. W. (2000). On profile likelihood. *J. Am. Stat. Assoc.* 95, 449–465. doi: 10.1080/01621459.2000.10474219

Navarro, D. J., Pitt, M. A., and Myung, I. J. (2004). Assessing the distinguishability of models and the informativeness of data. *Cogn. Psychol.* 49, 47–84. doi: 10.1016/j.cogpsych.2003.11.001

Pfeiffer, B. E., and Huber, K. M. (2009). The state of synapses in fragile X syndrome. *Neuroscientist* 15, 549–567. doi: 10.1177/1073858409333075

Pillow, J. W., Shlens, J., Paninski, L., Sher, A., Litke, A. M., Chichilnisky, E., et al. (2008). Spatio-temporal correlations and visual signalling in a complete neuronal population. *Nature* 454, 995–999. doi: 10.1038/nature07140

Rathour, R. K., and Narayanan, R. (2019). Degeneracy in hippocampal physiology and plasticity. *Hippocampus* 29, 980–1022. doi: 10.1002/hipo.23139

Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., et al. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* 25, 1923–1929. doi: 10.1093/bioinformatics/btp358

Raue, A., Kreutz, C., Maiwald, T., Klingmuller, U., and Timmer, J. (2011). Addressing parameter identifiability by model-based experimentation. *IET Syst. Biol.* 5, 120–130. doi: 10.1049/iet-syb.2010.0061

René, A., Longtin, A., and Macke, J. H. (2020). Inference of a mesoscopic population model from population spike trains. *Neural Comput.* 1–51. doi: 10.1162/neco_a_01292

Schneggenburger, R., and Neher, E. (2000). Intracellular calcium dependence of transmitter release rates at a fast central synapse. *Nature* 406, 889–893. doi: 10.1038/35022702

Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461–464. doi: 10.1214/aos/1176344136

Selkoe, D. J. (2002). Alzheimer's disease is a synaptic failure. *Science* 298, 789–791. doi: 10.1126/science.1074069

Spall, J. C. (2005). Monte Carlo computation of the fisher information matrix in nonstandard settings. *J. Comput. Graph. Stat.* 14, 889–909. doi: 10.1198/106186005X78800

Stephan, K. E., Baldeweg, T., and Friston, K. J. (2006). Synaptic plasticity and dysconnection in schizophrenia. *Biol. Psychiatry* 59, 929–939. doi: 10.1016/j.biopsych.2005.10.005

Südhof, T. C. (2008). Neuroligins and neurexins link synaptic function to cognitive disease. *Nature* 455, 903–911. doi: 10.1038/nature07456

Sun, J., Pang, Z. P., Qin, D., Fahim, A. T., Adachi, R., and Südhof, T. C. (2007). A dual-Ca2+-sensor model for neurotransmitter release in a central synapse. *Nature* 450, 676–682. doi: 10.1038/nature06308

Tsodyks, M., Pawelzik, K., and Markram, H. (1998). Neural networks with dynamic synapses. *Neural Comput.* 10, 821–835. doi: 10.1162/089976698300017502

Van Spronsen, M., and Hoogenraad, C. C. (2010). Synapse pathology in psychiatric and neurologic disease. *Curr. Neurol. Neurosci. Rep.* 10, 207–214. doi: 10.1007/s11910-010-0104-8

Van Trees, H. L. (2004). *Detection, Estimation, and Modulation Theory, Part I: Detection, Estimation, and Linear Modulation Theory*. New York, NY: John Wiley & Sons.

Van Trees, H. L., and Bell, K. L. (2007). Bayesian bounds for parameter estimation and nonlinear filtering/tracking. *AMC* 10:12. doi: 10.1109/9780470544198

Vanlier, J., Tiemann, C. A., Hilbers, P. A., and van Riel, N. A. (2014). Optimal experiment design for model selection in biochemical networks. *BMC Syst. Biol.* 8:20. doi: 10.1186/1752-0509-8-20

Venzon, D., and Moolgavkar, S. (1988). A method for computing profile-likelihood-based confidence intervals. *J. R. Stat. Soc.* 37, 87–94. doi: 10.2307/2347496

Wagenmakers, E.-J., Lodewyckx, T., Kuriyal, H., and Grasman, R. (2010). Bayesian hypothesis testing for psychologists: a tutorial on the savage-dickey method. *Cogn. Psychol.* 60, 158–189. doi: 10.1016/j.cogpsych.2009.12.001

Wang, T., Yin, L., Zou, X., Shu, Y., Rasch, M. J., and Wu, S. (2016). A phenomenological synapse model for asynchronous neurotransmitter release. *Front. Comput. Neurosci.* 9:153. doi: 10.3389/fncom.2015.00153

Weinberg, M. D. (2012). Computing the Bayes factor from a Markov chain Monte Carlo simulation of the posterior distribution. *Bayesian Anal.* 7, 737–770. doi: 10.1214/12-BA725

Welch, J. M., Lu, J., Rodriguiz, R. M., Trotta, N. C., Peca, J., Ding, J.-D., et al. (2007). Cortico-striatal synaptic defects and OCD-like behaviours in SAPAP3-mutant mice. *Nature* 448, 894–900. doi: 10.1038/nature06104

Wentzel, C., Delvendahl, I., Sydlik, S., Georgiev, O., and Müller, M. (2018). Dysbindin links presynaptic proteasome function to homeostatic recruitment of low release probability vesicles. *Nat. Commun.* 9, 1–16. doi: 10.1038/s41467-017-02494-0

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.

frontiers in Computational Neuroscience

HYPOTHESIS AND THEORY
published: 11 January 2021
doi: 10.3389/fncom.2020.593166

From Topological Analyses to Functional Modeling: The Case of Hippocampus

*Yuri Dabaghian**

Department of Neurology, The University of Texas McGovern Medical School, Houston, TX, United States

Topological data analyses are widely used for describing and conceptualizing large volumes of neurobiological data, e.g., for quantifying spiking outputs of large neuronal ensembles and thus understanding the functions of the corresponding networks. Below we discuss an approach in which convergent topological analyses produce insights into how information may be processed in mammalian hippocampus—a brain part that plays a key role in learning and memory. The resulting functional model provides a unifying framework for integrating spiking data at different timescales and following the course of spatial learning at different levels of spatiotemporal granularity. This approach allows accounting for contributions from various physiological phenomena into spatial cognition—the neuronal spiking statistics, the effects of spiking synchronization by different brain waves, the roles played by synaptic efficacies and so forth. In particular, it is possible to demonstrate that networks with plastic and transient synaptic architectures can encode stable cognitive maps, revealing the characteristic timescales of memory processing.

Keywords: spatial learning, hippocampus, topological methods, place cells, theoretical model

OPEN ACCESS

Edited by:
Tatyana Sharpee,
Salk Institute for Biological Studies,
United States

Reviewed by:
Masami Tatsuno,
University of Lethbridge, Canada
Daya Shankar Gupta,
Camden County College,
United States

***Correspondence:**
Yuri Dabaghian
yuri.a.dabaghian@uth.tmc.edu

Received: *09 August 2020*
Accepted: *02 December 2020*
Published: *11 January 2021*

Citation:
Dabaghian Y (2021) From Topological Analyses to Functional Modeling: The Case of Hippocampus. Front. Comput. Neurosci. 14:593166. doi: 10.3389/fncom.2020.593166

1. INTRODUCTION

Spatial cognition in mammals is based on an internal representation of their environments—a cognitive map—used for spatial planning, navigating paths, finding shortcuts, remembering the location of the home nest, food sources and so forth. A central role in producing these maps is played by the hippocampal neurons famous for their spatially tuned spiking activity. In rats, these neurons, known as "place cells," fire in specific domains of the navigated environment—their respective "place fields" (O'Keefe and Nadel, 1978; Moser et al., 2008). Thus, the spatial layout of the place fields in a given environment $\mathcal{E}$—a place field map $M_{\mathcal{E}}$—defines the temporal order in which place cells fire during animal's moves (Schmidt and Redish, 2013; Agarwal et al., 2014), and can therefore be viewed as a geometric "proxy" of the animal's cognitive map.

Experiments in "morphing" $2D$ environments demonstrate that place field maps are flexible: if the environment is deformed, then the place fields may change their shapes, sizes and locations, while preserving mutual overlaps, adjacencies, containments, etc. (Gothard et al., 1996; Leutgeb et al., 2005; Touretzky et al., 2005; Wills et al., 2005; Dabaghian et al., 2014; Bellmund et al., 2020; Place and Nitz, 2020). Hence the sequences in which the place cells fire during animal's navigation remain largely invariant within a certain range of geometric transformations, which suggests that the hippocampus provides a qualitative, topological representation of space—more akin to a subway map than a than to a topographical city street map (Alvernhe et al., 2012; Dabaghian et al., 2014; Wu and Foster, 2014).

The mechanisms that produce cognitive maps and the computational principles by which the brain converts patterns of neuronal firing into global representations of external space remain vague. Broadly, it is believed that the information provided by the individual place cells is somehow combined into single coherent whole. However, this "fusion" should not be viewed as a naïve aggregation of the smaller "pieces," because the signals provided by the individual neurons have no intrinsic spatial attributes; rather, spatial properties are *emergent*, i.e., appearing at a neuronal ensemble level (Wilson and McNaughton, 1993; Pouget et al. , 2000; Postle, 2006).

A computational framework developed in Dabaghian et al. (2012), Arai et al. (2014), Hoffman et al. (2016), Basso et al. (2016), Babichev et al. (2016a,b), Babichev and Dabaghian (2018), Dabaghian (2019), and Dabaghian (2016) helps to understand these phenomena by integrating the activity of the individual neurons into a large-scale map of the environment and to study the dynamics of its appearance, using algebraic topology techniques. Below we review some basic ideas and key concepts used in this framework, and discuss how they may apply to hippocampal physiology and cognitive realm. We then outline several examples that demonstrate how various characteristics of individual cells and synapses can be incorporated into the model and what effect these "microscopic" parameters produce at a "macroscale," i.e., in the map that they jointly encode.

2. TOPOLOGICAL MODEL

2.1. Alexandrov-Čech's Theorem

The topological nature of the cognitive map suggests that the information transmitted via place cell spiking should be amenable to topological analyses. For example, a place field map can be viewed as a cover of the environment $\mathcal{E}$ by the place fields υ_i, $\mathcal{E} = \cup_i \upsilon_i$, which, according to the Alexandrov-Čech's theorem (Alexandroff, 1928; Čech, 1932), encodes the topological shape of $\mathcal{E}$. To evaluate its specific characteristics, one constructs the nerve of the cover—an abstract simplicial complex $\mathcal{N}$ whose simplexes, $\nu_{i_0,i_1,\ldots,i_k} = [\upsilon_{i_0}, \upsilon_{i_1}, \ldots \upsilon_{i_k}]$, correspond to non-empty overlaps between the place fields, $\upsilon_{i_0} \cap \upsilon_{i_1} \cap \ldots \cap \upsilon_{i_k} \neq \emptyset$. If these overlaps are contractible, then $\mathcal{N}$ has the same topological shape as $\mathcal{E}$, i.e., the same number of components, holes, tunnels, etc. (Hatcher, 2002). An implication of this construction is that if the place fields cover the environment sufficiently densely, then their overlaps encode the topology of $\mathcal{E}$, which provides a link between the place cells' spiking pattern and the topology of the represented space (De Silva and Ghrist, 2007; Curto and Itskov, 2008; Chen et al. , 2012; Dabaghian et al., 2012; Kang et al., 2020).

2.2. Temporal Coactivity Complex

From the physiological perspective, the arguments based on the analyses of place fields provide only an indirect description of the information processing in the brain. In reality, the hippocampus and the downstream brain regions do not have access to the shapes and the locations of the place fields, which are but artificial constructs used by experimentalists to visualize their data. In the brain, the information is transmitted via neuronal spiking activity: if the animal enters a location where several place fields overlap, then there is a *probability* that the corresponding place cells will produce spike trains that overlap *temporally* (Curto and Itskov, 2008; Dabaghian et al., 2012). Such coactivities may be interpreted *intrinsically* by the downstream brain areas, and integrated into a global map of the ambient space. Thus, a proper description of place cell (co)activity requires a *temporal* analog of the nerve complex, built using temporal relationships between spike trains—which is, in fact, straightforward. Indeed, since the place field overlaps represent place cells' coactivities, one can construct a "coactivity complex" $\mathcal{T}$ whose simplexes correspond to combinations of active place cells, $\sigma = [c_{i_0}, c_{i_1}, \ldots, c_{i_k}]$. It was shown in De Silva and Ghrist (2007), Curto and Itskov (2008), and Dabaghian et al. (2012) that if such a complex is sufficiently complete (i.e., if it incorporates a sufficient number of the coactivity events) then its structure is similar to the structure of the spatially-derived nerve complex $\mathcal{N}$, e.g., $\mathcal{T}$ correctly captures the topology of the physical environment. Note however, that structural similarity between $\mathcal{N}$ and $\mathcal{T}$ (representability of $\mathcal{T}$) is a non-trivial point with profound mathematical implications (Tancer, 2013).

2.3. Simplicial Schemas of Cognitive Maps

Both complexes $\mathcal{N}$ and $\mathcal{T}$ provide a contextual framework for representing spatial information encoded by the place cells (Babichev et al., 2016b). For example, a sequence of place fields traversed during the rat's moves over a particular trajectory γ and the place cell combinations ignited along this trajectory can be represented, respectively, by a "nerve path" $\Gamma_{\mathcal{N}} = \{\nu_1, \nu_2, \ldots, \nu_k\}$—a chain of nerve-simplexes in $\mathcal{N}$, or by a "coactivity path" $\Gamma_{\mathcal{T}} = \{\sigma_1, \sigma_2, \ldots, \sigma_k\}$—a chain of the coactivity-simplexes in $\mathcal{T}$ (see also Babichev and Dabaghian , 2018). These simplicial paths qualitatively represent the shape of the physical trajectories: a closed simplicial path represents a closed physical route; a non-contractible simplicial path corresponds to a class of the physical paths that enclose unreachable or yet unexplored parts of the environment; two topologically equivalent simplicial paths $\Gamma_1 \sim \Gamma_2$ represent physical paths γ_1 and γ_2 that can be deformed into one another and so forth (Brown et al., 1998; Jensen and Lisman, 2000; Guger et al., 2011; Dabaghian, 2016). By the Alexandrov-Čech's theorem, the net pool of the simplicial paths can thus be used to describe the topological connectivity of the environment $\mathcal{E}$ via homological characteristics of $\mathcal{N}$ and $\mathcal{T}$.

2.4. The Large-Scale Topology of the Cognitive Map

$\mathcal{C}(\mathcal{E})$, as represented by a coactivity complex, can be described at different levels. A particularly concise description of a topological shape is given in terms of its topological loops (surfaces identified up to topological equivalence) in different dimensions, i.e., by its Betti numbers b_n, $n = 0, 1, \ldots$ (Alexandrov, 1965; Hatcher, 2002). For example, the number of inequivalent topological loops that can be contracted to a zero-dimensional ($0D$) vertex, $b_0(\mathcal{T})$, corresponds to the number of the connected components in $\mathcal{T}$; the number of loops that contract to a one-dimensional ($1D$) chain of links, $b_1(\mathcal{T})$, defines the number of holes and so forth (Alexandrov, 1965; Hatcher, 2002). The full list of

the Betti numbers of a space or a complex X is known as its topological barcode, $\mathfrak{b}(\mathcal{T}) = (b_0(\mathcal{T}), b_1(\mathcal{T}), b_2(\mathcal{T}), \ldots)$, which captures the topological identity of $\mathcal{T}$ (Zomorodian, 2005; Zomorodian and Carlsson, 2005; Ghrist, 2008; Carlsson, 2009, 2013; Edelsbrunner and Harer, 2010). For example, the barcode $\mathfrak{b} = (1, 1, 0, \ldots)$ corresponds to a topological annulus, the barcode $\mathfrak{b} = (1, 0, 1, 0, \ldots)$—to a two-dimensional ($2D$) sphere S^2, the barcode $\mathfrak{b} = (1, 2, 1, 0, \ldots)$—to a torus T^2 and so forth (Edelsbrunner and Harer, 2010). Thus, by comparing the barcode of the coactivity complex $\mathfrak{b}(\mathcal{T})$ to the barcode of the environment $\mathfrak{b}(\mathcal{E})$ one can establish whether their topological shapes may match, i.e., whether the coactivity complex provides a faithful representation of the environment at a given moment t.

2.5. A Model of Spatial Learning

A key difference between the complexes $\mathcal{N}$ and $\mathcal{T}$ is that the topological shape of $\mathcal{N}$ is fully defined by the structure of the place field map, whereas the shape of $\mathcal{T}$ unfolds in time at the rate with which the spike trains are produced. At every given moment of time, the coactivity complex $\mathcal{T}$ represents connections between the place fields that the animal had time to "probe": as the animal begins to explore a new environment, $\mathcal{T}$ is small, fragmented and may contain gaps that represent lacunae in the animal's internal map of the navigated space, rather than physical obstacles or inaccessible spatial domains. As the animal continues to navigate, more combinations of coactive place cells contribute connectivity information, the coactivity complex grows, $\mathcal{T}(t) \subseteq \mathcal{T}(t')$, $t < t'$, and acquires more details, converging to a stable shape that captures the physical structure of the surroundings.

Mathematically, $\mathcal{T}$ can thus be viewed as a *filtered* complex, with the filtration defined by the times of the simplexes' first appearance, t_σ (Dabaghian et al., 2012). Methods of the Persistent Homology theory allow describing the dynamics of the topological loops in $\mathcal{T}$, e.g., evaluating the minimal time $T_{\min}$ after which the topological structure of $\mathcal{T}$ matches the topology of the environment, $b_n(\mathcal{T}) = b_n(\mathcal{E})$ (De Silva and Ghrist, 2007; Curto and Itskov, 2008; Dabaghian et al., 2012). Biologically, this value provides a low-bound theoretical estimate for the time required to learn a novel topological map from place cell outputs (**Figure 1A**) (Dabaghian et al., 2012; Arai et al., 2014; Babichev et al., 2016a,b, 2018; Basso et al., 2016; Hoffman et al., 2016; Dabaghian, 2019).

2.6. Facing the Biological Realm

The physiological viability of these algebraic-topological constructions depends on the parameters of neuronal firing activity: just as there must be a sufficient number of place fields covering a space in order to produce a topologically correct nerve complex $\mathcal{N}$, certain conditions must be met by the place cell spiking profiles in order to produce an operational coactivity complex $\mathcal{T}$. For example, there should be enough cofiring of place cells with sufficient spatial specificity of spiking; the encoded relationships should not be washed out by noise; the model should make realistic predictions, e.g., produce viable learning periods in different environments, etc. Given that biological systems are highly variable (Fenton and Muller, 1998), these criteria may or may not be met by the physiological place cell ensembles, or vice versa, the model may single out a certain "operational" scope of parameters that may not match the biological range. In the following, we discuss this and other correspondences between the topological model and hippocampal physiology. We demonstrate that, first, the model can incorporate a vast scope of physiologically relevant characteristics of spike times, spiking statistics, their modulations by the "brain waves," efficacies of synaptic connections, architectures of the neuronal networks, etc., all of which correlate with dynamics of spatial learning. Second, the model allows converting this information *consistently* into coherent, biologically viable descriptions of a wide scope of neurophysiological phenomena. It becomes possible to systematically deduce functional properties of the system following not just the only empirical observations or experimental line of reasoning that currently dominate neurophysiological literature, but also the models' own, intrinsic logic.

2.7. Parameterization

To cope with the complexity of the cognitive map's construction, the model is built hierarchically: its main components implement most prominent physiological phenomena, and more subtle effects are incorporated as modifications of the skeletal structures. In the following, we will proceed in steps, by selecting a specific phenomenon, embedding it into the model using a minimal set of tools, outlining the results and discussing biological implications.

To simplify the approach, we will describe neuronal spiking in terms of Poisson firing rates, which, in case of the place cells, can be approximated by Gaussian functions of rat's coordinates with the amplitude f_i (the ith place cell's maximal firing rate), and the width s_i (the size of the corresponding place field) (Barbieri et al., 2004; Dabaghian et al., 2012). For an ensemble of N place cells, the N values s_i and f_i can be viewed as instantiations of two random variables drawn from their respective distributions with certain modes (s and f correspondingly) and standard deviations, σ_s and σ_f. To avoid overly broad or overly narrow distributions we impose additional conditions $\sigma_s = bs$ and $\sigma_f = af$ with the coefficients a and b selected so match the experimental statistics (Brunel et al., 2004; Barbour et al., 2007; Buzsáki and Mizuseki, 2014). As a result, each specific place cell ensemble can be indexed by a triplet of parameters, (s, f, N).

Second block of parameters characterizes animal's behavior, e.g., speeds and trajectory shapes, which are computationally intractable. We assume a practical approach to this problem and simulate non-preferential exploratory spatial behavior, with no artificial moving patterns or favoring of one segment of the environment over another, with typical experimentally observed speed ranges. Such approach allows reproducing a natural flow of spiking data and estimating how long it takes to integrate it into a topological map. The statistical alternatives for the model are produced by randomizing place field maps over a fixed trajectory rather than by sampling over different trajectories, which is practically much more efficient.

It should be emphasized however, that these and all the subsequent simplifications should not be viewed as limitations

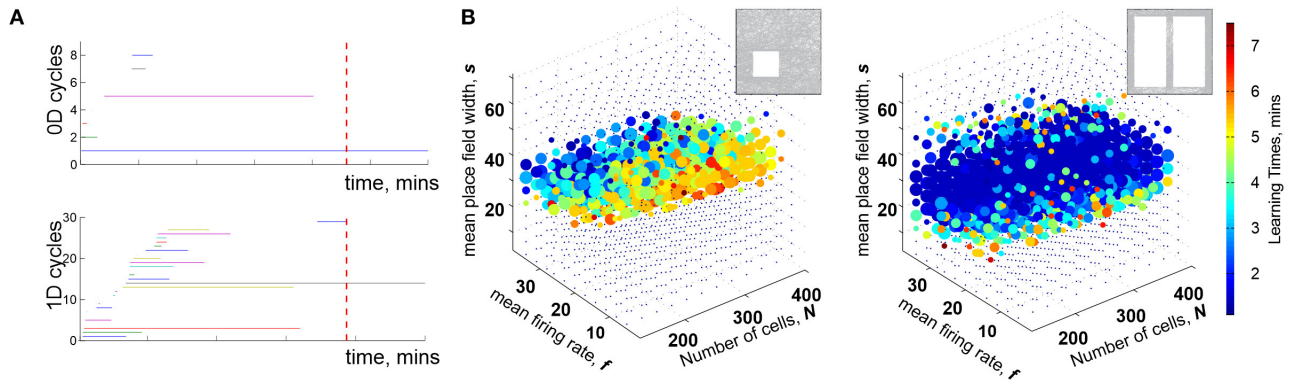

FIGURE 1 | Topological description of spatial learning. **(A)** As the rat begins to explore an environment, the simplicial complex $\mathcal{T}(t)$ contains many 0*D* cycles that mark contractible spatial domains represented by small groups of cofiring place cells. Additionally, there exists many 1*D* cycles that represent transient holes in $\mathcal{T}(t)$. As exploration continues, spurious cycles disappear, leaving behind only a few persisting ones, which express stable topological information. **(B)** Each point in the parameters space with coordinates (s, f, N) represents a particular place cell ensemble. The colors of the dots represent the mean learning time T_{min}; the larger the dot, the higher the rate of capturing the correct topological information for the corresponding (s, f, N). The ensembles that can produce a correct map occupy a particular domain of the parametric space—the *Learning Region*, $\mathcal{L}$, where learning is fastest and most accurate; near the boundary, map forms over times and exhibits higher error rates. Outside of $\mathcal{L}$ learning fails. Importantly, the parameter values that correspond to $\mathcal{L}$ happen to parallel experimentally derived values, which indicates a biological relevance of the model. The smaller $\mathcal{L}$ on the left panel corresponds to a 2*D* arena with a hole, about 1.5 × 1.5 m in size, and the larger $\mathcal{L}$ on the right corresponds to a quasi-linear environment (top right corner of each panel). The more complex the environment, the more tuned the neural ensembles have to be to learn the space.

of the approach but only as approximations used for simplifying specific computations. The model would also work with more detailed information, e.g., using more precisely estimated spike times or behavioral parameters, physiologically recorded or generated via accurate network models, detailed synaptic transmission mechanisms, etc.

3. OVERVIEW OF THE RESULTS

3.1. The Learning Region

For a particular set of values (s, f, N), a trajectory traversing a place field map $M_{\mathcal{E}}$ produces a certain time-dependent coactivity complex $\mathcal{T}(t)$. One may inquire whether, and for which ensembles, such a complex acquires the correct topological shape and how long this process may take. As it turns out, the coactivity complexes produced by generic place field maps can assume correct topological shapes, $b_k(\mathcal{T}) = b_k(\mathcal{E}), k \geq 0$, in a biologically feasible period—if the spiking parameters fall into a specific domain in the parameter space that we refer to as the *learning region*, $\mathcal{L}$ (**Figure 1B**). It is important to note that although the exact structure of $\mathcal{T}(t)$ depends profusely on the details of the map $M_{\mathcal{E}}$ (Babichev and Dabaghian , 2018), most large-scale characteristics of $\mathcal{T}(t)$, e.g., its Betti numbers, are largely $M_{\mathcal{E}}$-independent. This leads to the model's first predictive outcome, namely to the observation that the mean spiking parameters (s, f, N) may be used to identify a particular hippocampal "state" with a certain learning capacity,

$$T_{\min} = T_{\min}(s, f, N). \tag{1}$$

The second key observation is that the placement of the learning region in the parameter space matches the biological range of spiking characteristics derived from electrophysiologically recorded data (Dabaghian et al., 2012). *A priori*, this correspondence is not guaranteed: the region $\mathcal{L}$ that emerges from the "homological" computations could have appeared anywhere in the parameter space. However, the fact that the "operational" domain of the topological model appears to match the biological domain, suggests that the topological approach captures actual aspects of the neurophysiological computations taking place in the hippocampal network. In particular, it indicates that the physiological neurons can indeed encode a topological map of space in a biologically feasible time. On the other hand, boundedness of $\mathcal{L}$ also shows that spatial selectivity of firing does not, by itself, guarantee a reliable mapping of the environment, despite a widespread belief among neuroscientists to the contrary.

Third, the size and the shape of $\mathcal{L}$ reflect the scope of the biological variability that the hippocampus can afford in a given environment: the larger the learning region $\mathcal{L}$, the more stable the map (**Figure 1B**). Indeed, the model implies that the hippocampus can change its operating state inside $\mathcal{L}$ without compromising the integrity of the topological map: if one parameter begins to move outside the learning region, then a successful spatial learning can still occur, provided that compensatory changes of other parameters can keep the neuronal ensemble inside $\mathcal{L}$. This observation allows reasoning about the effects of certain diseases [e.g., Alzheimer's Cacucci et al., 2008; Cohen et al., 2013] or environmental toxins [e.g., ethanol Matthews et al., 1996; White and Best, 2000, cannabinoids Robbe and Buzsáki, 2009] that produce more diffuse place fields, lower place cell firing rates, smaller numbers of active cells and thus may disrupt spatial learning by shifting system's parameters beyond the perimeter of the learning region.

Fourth, the structure of the learning region may also vary with the geometry of the environment, the laboriousness of

navigation: the greater the task's complexity, the narrower the range that can sustain learning—as suggested by experimental studies (Nithianantharajah and Hannan, 2006; Fenton et al., 2008; Eckert and Abraham, 2010). Thus, despite the topological nature of the information processing, the place cells are not "agnostic" about the scale and the shape of the navigated space. In fact, it can be shown that maps of large spaces can be assembled from the maps of their parts, e.g., if a domain $\mathcal{E}$ is split into two subdomains $\mathcal{E}_1$ and $\mathcal{E}_2$ that meet but do not overlap, then one can compute the individual learning times $T_{\min}(\mathcal{E}_1)$ and $T_{\min}(\mathcal{E}_2)$ using only the spikes fired within each subdomain. The sum of these learning times is similar to total time spent by rat in the entire arena, $T_{\min}(\mathcal{E}) \approx T_{\min}(\mathcal{E}_1) + T_{\min}(\mathcal{E}_2)$, with statistically insignificant differences (Arai et al., 2014). Mathematically, this result may be viewed as an adaptation of the Mayer-Vietoris theorem that states that if a space $\mathcal{E}$ is split into pieces $\mathcal{E}_1$ and $\mathcal{E}_2$ that overlap over a domain with vanishing homologies, $H_q(\mathcal{E}_1 \cap \mathcal{E}_2) = 0$, then the homologies of the whole space are given by the direct sum of the homologies of the components, $H_q(\mathcal{E}) = H_q(\mathcal{E}_1) \oplus H_q(\mathcal{E}_2)$ (Hatcher, 2002). In case of the coactivity complexes, simulations demonstrate that persistent loops that represent topological obstacles in two complementary domains combine into the set of the persistent loops that represent the whole space, providing a novel perspective on the learning process.

These outcomes of the model correspond well with our subjective learning experiences: the complexity of the task and the size of the navigated environment influence learning time; difficult tasks are accomplished at or just beyond the limits of our capacity; disease or intoxication can reveal limits in our spatial cognition that would normally be compensated for, and so forth.

3.2. Coactivity Window

The results discussed above are based on topological analyses of spiking data produced by large populations of coactive place cells; but what defines neuronal coactivity in the first place? At a phenomenological level, an instance of coactivity may be characterized by the length of the period allocated for detecting the spikes fired by two or more cells. Experimental studies suggest that the "physiological" width w of the coactivity window ranges between tens to hundreds of milliseconds, with the standard estimate $w \sim 200$ ms (Ang et al., 2005; Huhn et al., 2005; Maurer et al., 2006; Mizuseki et al., 2009). The topological model allows addressing this question theoretically: one can ask, e.g., what range of window sizes *could* allow constructing topological maps and would these values match the biological range of coactivity periods? One can also inquire, given a particular width w, whether the dynamics of $\mathcal{T}(t)$ depends on a specific arrangement of the coactivity intervals along the time axis and how sensitive the results may be with respect to the windows' variations from one instance of coactivity to another. In biological terms: can the noise and/or variability of coactivity readouts affect the animal's learning capacity?

As it turns out, the answer to the latter two questions is negative: the statistics of place cell coactivity and hence the structure of the coactivity complex do not exhibit strong dependence on either the coactivity windows' random temporal shifts or on the window sizes' "jitter" (both for up to 50% of the mean w). On the one hand, this justifies using a single parameter w for studying the dependence of the coactivity complex' structure on the window width. On the other hand, it is clear that learning dynamics should depend on the systematic changes of w: if the coactivity window is too narrow, then the spike trains produced by the place cells will often "miss" one another, so that the map will either fail or take a long time to emerge. However, if w is too wide, then the place cells with disconnected place fields will contribute spurious links that may compromise the map's structure.

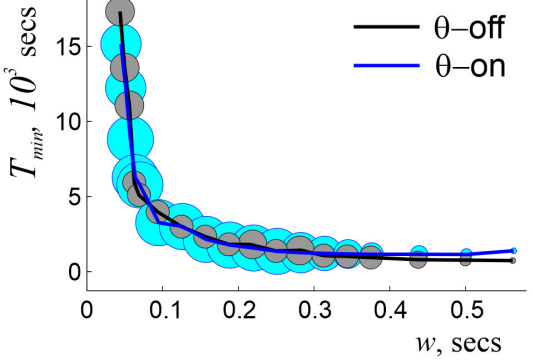

FIGURE 2 | Dependence of learning time on window width, with (blue line) and without (black line) θ-modulation of spiking activity. The radius of the circles indicates the percentage of times when topological learning is successful. In both cases, the coactivity complexes with physically correct topological shapes start to form at about $w_o = 0.2$ θ-periods, when the learning time is long (hours) and sensitive to the variations of w, and fail at $w \sim 4.5$ θ-periods, when learning becomes unreliable. At $w_s \sim 1.5$ θ-periods the dependence $T_{\min}(w)$ plateaus, marking the domain of stable w, which continues to $w_s \sim 3$ θ-periods. As w grows further, the success rate diminishes, and for $w_s > 5.5$ θ-periods topological learning fails. Here $s = 23$ cm, $f = 28$ Hz, and $N = 350$ cells.

Simulations show that indeed, an accurate topological map emerges within a well-defined range of ws, $w_o \approx 25 \leq w \leq w_c \approx 1,250$ ms, beyond which the maps have vanishing convergence rates (i.e., maps rarely or never produces the correct Betti numbers). In-between, the learning time follows a power law dependence, $T_{\min}(w) \sim w^{-\alpha}$, with $\alpha \approx 1.2$ starting at high values [$T_{\min}(w_o) \approx 5$ h] that rapidly decrease with growing w (**Figure 2**). The "operational" range of ws is even smaller since the biological dependence $T_{\min}(w)$ should be not only finite, but also *stable*, i.e., it should not be hypersensitive to variations of w or exhibit low convergence rates. In the model, such a range of ws lays approximately between 125 and 250 ms (**Figure 2**), which matches the domain implicated in experimental studies. Thus, the model once again allows deriving the physiologically observed values—in this case the operational widths of the coactivity windows—from purely theoretical considerations.

3.3. The Brain Waves

The temporal organization of the spike trains is strongly influenced by the oscillating extracellular electrical fields—the

brain waves, that control the temporal architecture of the spiking activity and the parcellation of the information flow in the brain (Buzsáki and Draguhn, 2004). In particular, the θ-wave (4–12 Hz) and the γ-waves (40–80 Hz), are known to modulate the place cells' activity at several timescales and affect spatial leaning (Buzsáki, 2002; Hasselmo et al., 2002; Colgin and Moser, 2010). However, it remains unclear at what level, and through what mechanisms, do these waves exert their influence. Most theoretical analyses address the effect of θ- and γ-rhythms on individual cells' spiking (Lisman and Idiart, 1995; Jensen and Lisman, 2000; Hasselmo et al., 2002). In contrast, the topological model allows addressing this question at the ensemble level, by tracing how the θ- and the γ-modulation of spike trains changes the dynamics of the corresponding coactivity complexes, e.g., the speed of their convergence toward correct topological shape, the statistics of topological defects exhibited during this process and so forth. Let us discuss a few examples.

(i) *θ-phase precession* is a key mechanism by which the θ-wave controls place cell's spiking: as a rat moves through a place field, the corresponding place cell spikes near a certain preferred θ-phase that progressively diminishes for each new θ-cycle (Buzsáki, 2005; Huxter et al., 2008) (**Figure 3A**). As discussed in Jensen and Lisman (1996) and Skaggs et al. (1996), this phenomenon helps to recapitulate the temporal sequence of the rat's positions in space during each θ-period and it is therefore widely believed to enhance learning (Buzsáki, 2002, 2005).

Simulations show that indeed, θ-precession significantly enlarges the learning region, making otherwise poorly performing ensembles much more capable of learning. Without θ-precession, the learning region $\mathcal{L}$ is small and sparse, and vice versa, certain place cell ensembles that in absence of θ lay beyond the learning region, become functional with the addition of θ-precession (**Figure 3B**). Moreover, θ-precession increases the probability of the correct outcome for ensembles that occasionally fail to form an accurate map, which suggests that θ-precession may not just correlate with, but actually enforce spatial learning (Arai et al., 2014).

In terms of the coactivity complex' structure, θ-enhancement of learning is manifested through shortened durations of the spurious $1D$ cycles, while initially increasing their number. In other words, θ-modulation suppresses spurious defects in the cognitive map at the price of creating more transient errors at the initial stages of the navigation. Curiously, simulations also show that learning times are relatively insensitive to the details of the θ-wave structure: the presence of a spike-modulating θ-rhythm by itself is more important than a specific wave shape (Arai et al., 2014).

As for the interplay with the coactivity parameters, the stabilization of the $T_{\min}(w)$ dependence is achieved at approximately the same range of ws as without the θ-precession, at $w \sim 1-2\,\theta$-cycles (Arai et al., 2014) (**Figure 2**). Such recurrent matches between the preferred coactivity timescale and the θ-timescale suggest that the interplay between neuronal spiking and the parameters of animal's behavior (e.g., speed) required for optimal processing of topological information may actually define the temporal domain of neuronal synchronization in the rat's hippocampal network. Thus, θ-modulated coactivity complexes provide a self-consistent description of the hippocampal network's function at the θ-timescale, predicting *inter alia* an optimal integration window for reading out the information and the temporal domain of synchronization.

(ii) *γ-modulation of spiking*. As w shrinks beyond the range predicted for the independently θ-precessing place cells ($w < w_o$), spatial learning fails. Interestingly, this happens precisely at the timescale where complementary mechanisms of spike synchronization, driven by the second key component of the hippocampal brain waves—the γ-oscillations—are taking over (Colgin et al., 2009; Buzsáki and Wang, 2012). This raises question about whether an additional γ-synchronization of spiking could improve the predicted properties of the cognitive map, i.e., produce topologically correct coactivity complexes.

Physiologically, γ-wave represents fast oscillations of the inhibitory post-synaptic potentials. As its amplitude $A_\gamma(t)$ drops at a certain location, the surrounding cells with high membrane potential spike (Lisman, 2005; Jia and Kohn, 2011; Nikoli et al., 2013). As a result, each γ-trough defines the preferred θ-phase of several cells, i.e., marks an ignition of a particular place cell combination, represented by a coactivity simplex. Computationally, coupling spike times with the γ-wave can be achieved by modulating neuronal firing rates with a Boltzmann factor $e^{-A_\gamma(t)/\tau_i}$. The parameter τ_i can be interpreted as an effective "temperature" that controls the temporal spread of spikes around the ith γ-trough: for large mean $\tau = \langle\tau_i\rangle$, the spikes are "hot," i.e., spread diffusely near the γ-troughs and for small τ they "freeze" at them. In particular, the case in which the spike trains are uncorrelated with the γ-troughs corresponds to the limiting case of an "infinitely hot" hippocampus ($\tau = \infty$, e.g., the pure θ-modulated cells discussed above). Meanwhile, the "physiological" effective temperature that describes the characteristic huddling of spikes within a γ-period observed in the experiments (Colgin et al., 2009; Colgin and Moser, 2010) is comparable to the mean γ-amplitude, $\tau \approx \overline{A_\gamma}$.

The net effect of the γ-modulation on the coactivity complexes is as follows: as the effective temperature drops and the temporal spread of the spikes near the γ-troughs shrinks, the coactivity complexes produce fewer, faster-contracting spurious loops. In particular, at the "physiological" effective temperatures, γ-synchronized cognitive map can robustly capture the topology of the environment by integrating place cell coactivity at the γ-timescale, i.e., yield *finite* learning times at $w < w_o$s, which provides a direct demonstration of the importance of the γ-synchronization at the systemic level.

This result may shed light on the well-known correlation between successful learning and retrieval with the increase of the γ-amplitude in raised attention states (Moretti et al., 2009; Vugt et al., 2010; Lundqvist et al., 2011; Trimper et al., 2014). In particular, it helps understanding why suppression of the γ-waves induced, e.g., by psychoactive drugs (Whittington et al., 2000a,b), such as cocaine (Dilgen et al., 2013; McCracken and Grace, 2013), or arising due to neurodegeneration or aging (Vreugdenhil and Toescu, 2005; Lu et al., 2011), usually correlates with learning impairments—according to the model,

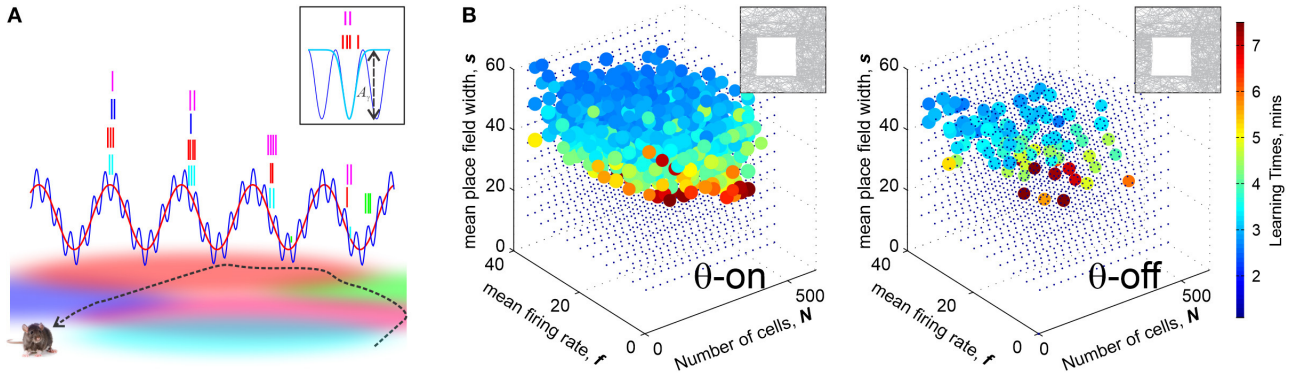

FIGURE 3 | Brain waves enhances learning. **(A)** θ-precession and γ-synchronization modulate place cell spiking activity: Spike times precess with the θ-rhythm ($\approx$ 8 Hz), schematically shown by the red wave: as the rat traverses a place field, the corresponding place cell discharges at a progressively earlier phase in each new θ-cycle. The preferred θ-phases correspond to γ-cycles ($\approx$ 60 Hz). The blue wave shows the net $\theta + \gamma$ amplitude. Boxed image: the spikes, shown by tickmarks colored according to the place fields traversed by the animal's trajectory, cluster over the γ-troughs, yielding dynamical cell assemblies. The spike time probabilities are modulated by a Boltzmann factor $e^{-A_\gamma(t)/\tau}$, where A_γ the amplitude of a trough and τ is an "effective temperature." **(B)** Learning regions with θ-precession (left) and without it (right). In the latter case, the size and the density of $\mathcal{L}$ diminishes, indicating that θ-oscillations enhance place cells' ability of to encode spatial maps, making them more resilient in the face of the spiking rate or population size changes. Computations are made for a 1 × 1 m environment shown in the top left corner.

all these phenomena suppress map formation—or retrieval—at the γ timescale. On the constructive side, the model suggests a new characteristics of the γ-synchronized spiking activity—the effective γ-temperature of spiking—that may be studied empirically and explained via neuronal mechanisms.

3.4. Ramifications of Coactivity Complexes

The predictions derived from the constructions discussed above are not universal. For example, a direct application of the model to the case of the bats navigating 3*D* caves (Ulanovsky and Moss, 2007; Yartsev and Ulanovsky, 2013) often produces dysfunctional coactivity complexes, with hundreds of persistent spurious loops—even for the experimentally observed parameters of spiking activity (Hoffman et al., 2016). On the one hand, this failure can be explained by the relatively high speeds of the bat's movements (over 2 m/s), which allows producing spurious coactivities between place cells with non-overlapping place fields (Hoffman et al., 2016). On the other hand, it also suggests that the very idea that place cells operate by responding to certain spatial domains (currently dominating in the field) may be only a simplified interpretation of their spiking mechanism, suitable for low speeds and basic environments. The model points out that deriving topological maps from such "passive responses" may, at higher speeds, generate mismatches between the spatial pattern of the prearranged place fields and the temporal pattern of the corresponding place cells' coactivities. In other words, the model suggests that the raw pool of place cell spiking data requires *editing*—a surprising conclusion because it appeals to reasoning beyond the model's original setup. In effect, it suggests that the hippocampal network should be *wired* to highlight some place cell coactivities and suppress others, even though no explicit references to the networks' structure were made in the original Alexandrov-Čech construction.

Curiously, this line of arguments addresses to a well-known neurophysiological phenomenon, namely the fact that place cells tend to form operative units known as *cell assemblies*—functionally interconnected groups of neurons that drive their respective "readout" neurons in the downstream networks (Harris et al., 2003; Harris, 2005; Jackson and Redish, 2007; O'Neill et al. , 2008; Buzsaki, 2010). The spiking response of the latter actualizes connectivity relationships between the regions encoded by the individual place cells: if a specific instance of place cell coactivity does not elicit a response of a readout neuron, then the corresponding connectivity information does not contribute to the hippocampal map (Buzsaki, 2010; Babichev et al., 2016b). A cell assembly network of a specific architecture can thus control processing of the information supplied by the place cell spiking activity and the overall connectivity structure of the cognitive maps (**Figure 4A**).

3.4.1. Clique Coactivity Complexes

A simple model a place cell assembly network can be built by constructing a coactivity graph $\mathcal{G}$, whose vertexes v_i correspond to place cells c_i and the links, $\varsigma_{i_0 i_1} = [v_{i_0}, v_{i_1}]$ represent the connections (functional or physiological) between pairs of coactive cells (Burgess and O'Keefe, 1996; Muller et al., 1996). The place cell assemblies then correspond to fully interconnected subgraphs of $\mathcal{G}$, i.e., to its maximal cliques $\varsigma = [c_{i_0}, c_{i_1}, \ldots, c_{i_n}]$. As a combinatorial objects, cliques are identical to the simplexes span by the same sets of vertexes; hence the collection of $\mathcal{G}$-cliques produces a complex (Jonsson, 2008) that may serve as a schematic representation of either the cell assembly network or the cognitive map encoded by it (Babichev et al., 2016a).

Simulations show that such complexes, denoted below as $\mathcal{T}_\varsigma$, are structurally very similar to the original coactivity complexes derived from the higher-order place cell coactivities, which we will denote as $\mathcal{T}_\sigma$. However functionally, $\mathcal{T}_\varsigma$s often perform much

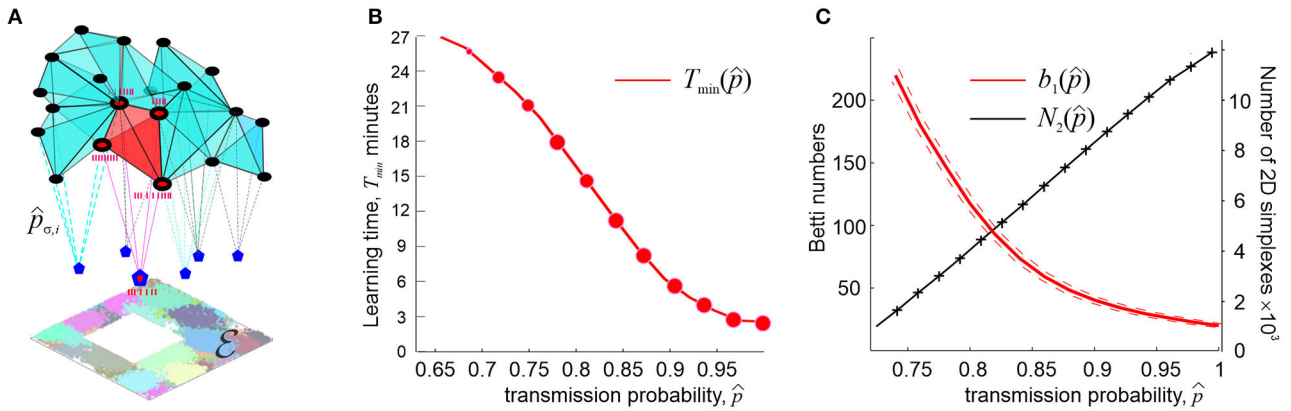

FIGURE 4 | Synaptic efficacies and cell assembly complexes. **(A)** Cell assemblies are functionally interconnected network of place cells (black dots) are modeled as cliques of the coactivity graph $\mathcal{G}$. Spikes from kth pair of coactive place cells in an assembly σ are transmitted to a readout neuron (blue pentagons) with probability $p_{\sigma,k} < 1$ (an ignited cell assembly is shown in red). The net structure of the cell assembly network is represented by the corresponding cell assembly complex $\mathcal{T}_{CA}$, which captures the topology of underlying environment $\mathcal{E}$. **(B)** The number of coactivity links shrinks with the diminishing spike transmission probability at a power rate (black line), whereas the number of spurious topological loops in $\mathcal{T}_{CA}$ proliferates exponentially. **(C)** As synapses weaken, the learning time $T_{\min}$ grows at a power rate. The size of the data points represents the percentage of the outcomes with the correct Betti numbers $[b_{0,1}(\mathcal{T}_{CA}) = b_{0,1}(\mathcal{E}) = 1]$. Computations are performed using an ensemble of $N = 400$ neurons with a mean firing rate of $f = 28$ Hz and mean place field size 30 cm.

better than $\mathcal{T}_\sigma$s, e.g., they exhibit a much smaller number of shorter-living spurious loops, more robust learning times, etc. (Babichev et al., 2016a; Basso et al., 2016; Hoffman et al., 2016). The explanation for this effect is simple: the lowest order, *pairwise* place cell coactivities are captured easier and more reliably than the higher-order coactivity events (Katz et al., 2007; Brette, 2012). An additional advantage is offered by a structural flexibility of the clique coactivity complexes, since it is possible to assemble its individual cliques $\varsigma \in \mathcal{T}_\varsigma$ by *accumulating* low order coactivities over time, rather than by *detecting* higher-order coactivity events. For example, in order to identify a third-order coactivity clique, $\varsigma = [c_{i_0}, c_{i_1}, c_{i_2}]$, one can first detect the coactive pair $[c_{i_0}, c_{i_1}]$, then $[c_{i_1}, c_{i_2}]$ and then $[c_{i_0}, c_{i_2}]$, over an extended integration window ϖ, whereas in order to produce a coactivity simplex $\sigma = [c_{i_0}, c_{i_1}, c_{i_2}]$, all three cells must become active within a single coactivity window w.

From the physiological perspective, the clique construction can be used to model a wide scope of physiological phenomena, e.g., for testing whether the readout neurons may operate as "coincidence detectors" that respond to nearly simultaneous activity of the pre-synaptic cells [for short integration windows $\varpi \sim w$ Katz et al., 2007; Brette, 2012] or as "integrators" of the spiking inputs [for $\varpi \gg w$ König et al., 1996; Magee, 2000; London and Häusser, 2005; Spruston, 2008; Ratté et al., 2015], along with the intermediate and/or mixed cases. The original approach based on the Alexandrov-Čech's construction corroborates with the first scenario: indeed, the nerve complex $\mathcal{N}$ is derived from the spatial overlaps between the regions, which mark the domains of nearly simultaneous place cell coactivity. The architecture of the clique coactivity complex suggests an alternative approach that significantly broadens the models' capacity to represent synaptic computations.

Simulations show that, in fact, the connections within most cliques of $\mathcal{G}$ activate nearly simultaneously, i.e., most simplexes of $\mathcal{T}_\sigma$ are also present in $\mathcal{T}_\varsigma$. Nevertheless, there exists a small population of cliques that are never observed as simultaneous coactivity events and require assembling over extended periods (Hoffman et al., 2016). As a result, clique coactivity complexes $\mathcal{T}_\varsigma$ are typically larger and produce much fewer spurious topological loops that rapidly disappear with learning. In particular, such complexes produce correct topological maps of $3D$ spaces for the experimentally observed parameters of the spiking activity (Hoffman et al., 2016), suggesting that the readout neurons in bats' (para)hippocampal areas should function as integrators of synaptic inputs (with estimated spike integration period of about 4 min), rather than detectors of place cells' coactivity—a prediction that may potentially be verified experimentally.

Another curious difference between the rats' and the bats' cognitive map construction mechanism is that less than 4% of the bat's place cells exhibit significant θ-modulated firing (Yartsev and Ulanovsky, 2013), which implies that θ-precession in these animals may not play the same role as in rats. Indeed, simulating bat's movements with and without θ-precession reveals that in the θ-off case, the ensembles of place cells acquire correct maps faster than in the θ-on cases, producing fewer topological loops both in the simplicial and in the clique coactivity complexes (Hoffman et al., 2016). To explain these results, one can consider the effect of θ-precession from two perspectives: on the one hand, it synchronizes place cells and hence increases their coactivity rate, which may help learning (Buzsáki, 2002; Harris et al., 2002; Lee et al., 2004; Geisler et al., 2010; Jezek et al., 2011). On the other hand, it can be viewed as a constraint that reduces the probability of the cells' coactivity and hence decimates the pool of coactivity events (Skaggs et al., 1996). In relatively slow moving rats, when the coactivity events are reliably captured, the first effect dominates, contributing a steady influx of grouped spikes to downstream neurons. In rapidly moving bats however, when the network

struggles to capture the coactivities, the constraint imposed by phase precession acts more as an impediment and slows down spatial learning process.

3.5. Cell Assembly Complex

Question arises, whether the coactivity complexes may be implemented, in some capacity, in physiological networks or vice versa, whether it is possible to construct complexes that capture the organization of the cell assembly networks. Simulations show that the original set of coactive place cell combinations is very large: the numbers of d-dimensional simplexes in $\mathcal{T}_\varsigma$, N_d, scale proportionally to the binomial coefficients C_N^{d+1}. More specifically, it can be shown that the ratios $\eta_d = N_d/C_N^{d+1}$ depend primarily on the mean place field sizes and the firing rates and not on the number of cells within the ensemble, N (Arai et al., 2014). In contrast, the number of cells that may potentially serve as readout neurons is similar to the number of place cells N, which implies that only a small fraction of coactive place cell groups can form assemblies (Shepherd, 2004; Buzsaki, 2010). This raises the question: is it possible to identify a sufficiently small but functionally complete set of place cell combinations—putative cell assemblies—using simple selection rules?

In model's terms, the task of identifying a subpopulation of coactive place cell combinations corresponds to selecting a "cell assembly subcomplex" $\mathcal{T}_{CA}$ of $\mathcal{T}_\varsigma$, according to some biologically motivated criteria. First, the total number of the maximal simplexes in $\mathcal{T}_{CA}$ should be comparable to the number of its vertexes (i.e., of active cells), $N_{\max}(\mathcal{T}_{CA}) \approx N(\mathcal{T}_{CA})$, but the latter should not differ significantly from the original number of place cells, $N(\mathcal{T}_{CA}) \approx N(\mathcal{T}_\varsigma)$. Second, only a few cell assemblies (selected cliques) should be active at a given location, to avoid redundancy of the place cell code. Conversely, the periods during which all place cell assemblies are inactive should be short, so that the rat's movements should not go unnoticed by the hippocampal network. Third, the larger is the number of cells shared by consecutively igniting cell assemblies (i.e., by the adjacent simplexes in a simplicial path), the more contiguous is the representation of the rat's moves. Hence the contiguity between the simplexes in $\mathcal{T}_{CA}$ should not decrease compared to $\mathcal{T}_\varsigma$. Lastly, $\mathcal{T}_{CA}$ should correctly capture the topological shape of the environment (Babichev et al., 2016a).

As it turns out, it is possible to carry out the required construction by selecting the most prominent combinations of coactive place cells—the ones that appear most frequently. This selection principle is motivated by the Hebbian "fire together wire together" neuronal plasticity mechanisms: frequently appearing combinations have a higher chance of being wired into the network (Neves et al., 2008). Specifically, one can construct the desired clique complexes by identifying the connections the coactivity graphs $\mathcal{G}(\xi)$ that activate at a rate exceeding a certain threshold ξ. Alternatively, one can select, for every cell c_i, its n_0 neighbor-cells that are most frequently coactive with c_i, which yields another family of coactivity graphs, $\mathcal{G}(n_0)$. Computations show that the first family, $\mathcal{G}(\xi)$, exhibits certain random graph properties while the second family, $\mathcal{G}(n_0)$, demonstrates scale-free properties (Barabási and Albert, 1999; Albert and Barabási, 2002), characteristic of the hippocampal network (Bonifazi et al., 2009; Li et al., 2010). However, both families of "restricted" coactivity graphs allow constructing operational cognitive map models, for a viable set of ξs and n_0s.

As expected, the size and the dimensionality of the corresponding clique complexes, $\mathcal{T}_\varsigma(\xi)$ and $\mathcal{T}_\varsigma(n_0)$, decrease with the growing threshold ξ or diminishing n_0. In addition, their maximal simplexes become more contiguous and their number, $N_{\max}$, remains close to the number of cells. Lastly, the topological behavior of both $\mathcal{T}_\varsigma(\xi)$ and $\mathcal{T}_\varsigma(n_0)$ is also regular: with minor rectification algorithms that do not change significantly the complex's structure or alter the appearance rate of simplexes, correct topological shapes can be attained as fast and as reliably as with the entire set of the place cell coactivities, without compromising the place cell coverage of the environment or fragmenting the map (Babichev et al., 2016a). Thus, the generic biological requirements listed above are met and we may conclude that the selected "critical mass" of coactive place cell combinations can produce viable cell assembly complexes $\mathcal{T}_{CA}(\xi)$ and $\mathcal{T}_{CA}(n_0)$ (Babichev et al., 2016a).

3.6. Synaptic Parameters

The physiologically implementable cell assembly complexes $\mathcal{T}_{CA}$ set the stage for further developments of the topological model. For example, the simplexes of $\mathcal{T}_{CA}$ can be rigged with parameters describing transferring, detecting and interpreting neuronal (co)activity in the corresponding cell assemblies, allowing us to account for the effects of the hippocampal network's synaptic architecture and providing a basic description of the synaptic computations in the cell assemblies.

In a phenomenological approach, synaptic connections can be characterized simply by the probabilities of transmitting spikes from a place cell to a readout neurons' membranes and by the probabilities that the latter will spike upon collecting their inputs. If the cell assemblies are modeled as cliques of the coactivity graph, then the key role is played by the probability of transmitting the coactivity from the pairs of coactive place cells to the corresponding readout neurons' and response probabilities. In principle, these probabilities can be evaluated using detailed neuronal and synaptic models; however, in a simpler phenomenological approach, they may be regarded as random variables drawn from stationary, unimodal distributions with the modes p_* (transmission) and q_* (response) and the variances Δ_p and Δ_q. The stationarity here implies that we disregard synaptic plasticity processes (Brunel et al., 2004; Barbour et al., 2007; Buzsáki and Mizuseki, 2014).

Under such assumptions, it is possible to study how the large scale, systemic characteristics of the spatial memory map depend on the synaptic strengths, at what point spatial learning may fail, and so forth. It can be shown, e.g., that if the characteristic coactivity transmission probability is high ($0.9 \leq p_* \leq 1$) then its small variations do not produce strong effects on the spatial map. On the other hand, as p_* decreases further, the changes accumulate and, as p_* approaches a certain critical value p_{crit}, learning times diverge at a power rate,

$$T_{\min} \propto (p_* - p_{crit})^{-\kappa},$$

with κ ranging typically between 0.1 and 0.5 (**Figure 4B**). The effects produced by the diminishing probability of the post-synaptic neurons' responses, q_*, are qualitatively similar but weaker than the effects of lowering the spike transmission probability p_* (Dabaghian, 2019).

These results suggest explanations for numerous observations of correlative links between weakening memory capacity and deterioration of synapses, broadly discussed in neuroscience literature (Shapiro, 2001; Selkoe, 2002; Toth et al., 2012). According to model, weakening synapses reduce the size the coactivity complex and degrade its topological structure. For example, simulations demonstrate the number of connections in the coactivity graph $\mathcal{G}$ near p_{crit} drops as $N_2 \propto (p_* - p_{crit})^{\delta}$, $\delta \sim 1$, whereas the number of longer-lasting $1D$ spurious loops in the corresponding coactivity complex grows exponentially, $\log(b_1) \propto (p_{crit} - p_*)$ (**Figure 4C**), suggesting a phase transition from a regular to an irregular state (Donato et al., 2016). In physiological terms, this implies that synaptic depletion reduces the number of detectable coactivities while generating defects in the cognitive map, which results in a rapid increase of the learning time.

Moreover, weakening synapses reduce the learning region down to its compete disappearance at $p_* = p_{crit}$, which suggests that spatial learning may fail not only because the parameters of neuronal firing are pushed beyond a certain fixed "working range," but also because that range itself may shrink or cease to exist. In particular, the fact that the learning region disappears if the transmission probability drops below the critical value implies that deterioration of memory capacity produced by the synaptic failure cannot be compensated by increasing the place field's firing rates or by recruiting a larger population of active neurons—for more details see Dabaghian (2019).

3.7. Dynamical Cell Assemblies

Physiologically, cell assemblies are *dynamic* structures: they may form among the cells that demonstrate repeated coactivity and disband as a result of deterioration of synaptic connections, caused by reduction or cessation of spiking, then reappear during a subsequent surge of coactivity, disband again and so forth (Harris et al., 2003; Buzsaki, 2010). In the model, the formation and disbanding of the cell assemblies is represented by the appearances and disappearances of the corresponding simplexes, so that the net dynamics of the cell assembly network and the evolution of the resulting cognitive map is represented by a "flickering" cell assembly complex, denoted as $\mathcal{F}(t)$. Unlike its "perennial" counterpart $\mathcal{T}(t)$, which can only grow and stabilize with time (**Figures 5A,B**), the flickering complex $\mathcal{F}(t)$ may inflate, shrink, fragment into pieces that may fuse back together, produce transient holes, fractures, gaps, and other dynamic "topological defects" (**Figures 5C,D**).

One of the key questions that can be addressed by the model is the following: experimentally, cell assemblies' lifetimes range between minutes (Goldman-Rakic, 1995; Billeh et al., 2014; Hiratani and Fukai, 2014) and hundreds of milliseconds (Whittington et al., 2000a,b; Bi and Poo, 2001; Bennett et al., 2018), whereas cognitive representations of environments can last for days and months (Clayton et al., 2003; Brown et al., 2007; Meck et al., 2013). How can a rapidly rewiring network sustain stable representations of the world? In model's terms, can the large-scale topological properties of $\mathcal{F}(t)$ be stable, despite rapid recycling of its simplexes? Computationally, this question can be addressed using Zigzag Persistent Homology theory (Edelsbrunner et al., 2002; Carlsson et al., 2009; Carlsson and Silva, 2010).

(i) Decaying flickering coactivity complexes. Flickering of the coactivity complexes and their topological dynamics can be simulated in many ways (see Battiston et al., 2020 for a broad review). A simple model can be based on the dynamics of links of the coactivity graph $\mathcal{G}$ as follows.

- Vertexes ς_i of $\mathcal{G}$ appear at the moment of the corresponding place cells' first activation and thereupon remain stable, as place cells do in learned environments (Thompson and Best, 1990).
- A link ς_{ij} between vertexes ς_i and ς_j appears with probability $p_{ij}^{+} = 1$ at the moment when cells c_i and c_j become coactive and disappears with probability $p_{ij}^{-}(t) =\sim e^{-t/\tau_{ij}}$, where time t is counted from the moment of ς_{ij}'s last activation and τ_{ij} defines its *proper* decay time. Below we consider a simple case in which all connections decay at the same rate, $\tau_{ij} = \tau$,

$$p_{ij}^{-}(t) =\sim e^{-t/\tau}, \tag{2}$$

  so that the decay dynamics of the flickering coactivity graph depends on a single parameter τ.
- The behavior of the higher-order cliques and hence of the flickering complex $\mathcal{F}_\tau$ are also defined by the link decay period τ. Note that pairs of place cells may coactivate before decaying, i.e., links in $\mathcal{G}_\tau$ can rejuvenate; hence cliques of orders $m \geq 1$ may acquire *effective lifetimes* $\tau_e^{(m)} > \tau$.

As mentioned in §1 of this section, details of the coactivity complex' dynamics depend on the sequence in which the rat traverses place fields in a map $M_{\mathcal{E}}$. For a given map $M_{\mathcal{E}}$, a trajectory $\gamma(t)$ and fixed physiological parameters (firing rates, place field sizes, etc.), the Betti numbers $b_k(\mathcal{F}_\tau(t))$ depend primarily on the links' decay time τ (Babichev and Dabaghian, 2017a,b; Babichev et al., 2018, 2019). One would expect that if τ is too small (e.g., if the coactivity simplexes tend to disappear between two consecutive co-activations of the corresponding cells), then the flickering complex should rapidly deteriorate without attaining an adequate topological shape. If τ is too large, then the effect of the decaying connections should be insignificant, i.e., the flickering complex $\mathcal{F}_\tau(t)$ should follow the dynamics of its "perennial" counterpart $\mathcal{T}(t) \equiv \mathcal{F}_\infty(t)$, constructed for the same spiking parameters. In particular, if the place cells' coactivity complex $\mathcal{T}(t)$ assumes the correct topological shape in a biologically viable time $T_{\min}(\mathcal{T})$, then a similar behavior should be expected from its slowly decomposing counterpart $\mathcal{F}_\tau(t)$. For intermediate values of τ, the topological dynamics of $\mathcal{F}_\tau(t)$ may exhibit a rich variety of behaviors.

Simulations show that a characteristic interval between successive activations of links in the environment shown on

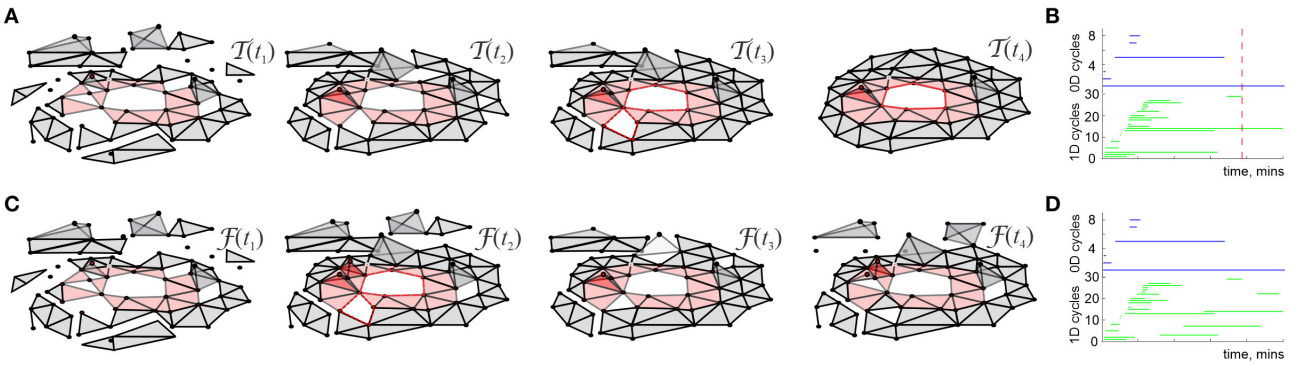

FIGURE 5 | Perennial and flickering coactivity complexes. **(A)** In absence of decay, the coactivity complex $\mathcal{T}$ steadily grows. At first, it contains many pieces riddled transient holes, but then, as the place cells' spiking information accumulates, there emerges just one stable piece and only one hole survives (shown by red chain of simplexes)—the one that corresponds to non-contractible simplicial path encircling a hole in the environment $\mathcal{E}$ (**Figures 3B**, **4A**). **(B)** The timelines of 0*D* (top) and 1*D* (bottom) topological loops in $\mathcal{T}$, computed using Persistent Homology theory methods, show that the topological shape of the coactivity complex stabilizes. One persistent loop in each dimension remaining after the minimal learning time T_{min} (vertical dashed line) indicate the stable topological barcode of $\mathcal{T}$. **(C)** If decay is allowed, then the coactivity simplexes may not only appear but also disappear, yielding a "flickering" coactivity complex $\mathcal{F}$. Unlike the perennial complexes $\mathcal{T}$, flickering complexes $\mathcal{F}$ may never stabilize, i.e., transient topological defects, described by Zigzag Persistent Homology theory **(D)** may persist indefinitely.

Figures 5A,C is about $\Delta t \approx 30$ s. If the proper decay times are not too large ($2.5\Delta t \lesssim \tau \lesssim 4.5\Delta t$), then the time intervals between consecutive births and deaths of a link ς distribute bimodally: the relatively short lifetimes distribute exponentially, with about twice longer *effective* lifetimes $\tau_e^{(2)} \approx 2\tau$ (higher-order simplexes decay more rapidly, e.g., $\tau_e^{(3)} \approx \tau$, etc.). In addition, there appears a pool of long-living connections that persist throughout the entire navigation period (**Figure 6A**). In other words, the flickering coactivity complex $\mathcal{F}_\tau(t)$ acquires a stable "core" formed by a population of "surviving simplexes," enveloped by a population of "rapidly fluttering" simplexes.

The resulting mix of skeletal (stable) and fluttering simplexes rapidly grows at the onset of the navigation and begins to saturate by the time a typical link makes an appearance, which, incidentally, is comparable to the "perennial" learning time $T_{\min}(\mathcal{T})$ (a few minutes). The characteristic size of $\mathcal{F}_\tau(t)$ grows to about a half of the size of $\mathcal{F}_\infty(t)$, with about 15% fluctuations (**Figure 6B**). Thus, the population of simplexes in $\mathcal{F}_\tau(t)$ is indeed transient: although the size of $\mathcal{F}_\tau(t)$ fluctuates slowly from one moment of time to the next, the set of simplexes that are present at a given moment of time t but missing at a later moment t', grows as a function of temporal separation $|t - t'|$, becoming close to the sizes of either $\mathcal{F}_\tau(t)$ or $\mathcal{F}_\tau(t')$ in approximately one learning period $T_{\min}(\mathcal{T})$ (Babichev et al., 2018, 2019).

The topological shape of $\mathcal{F}_\tau(t)$ changes much slower: after a brief initial stabilization period, the topological barcode $\mathfrak{b}(\mathcal{F}_\tau)$ remains similar to the barcode of the navigated environment $\mathcal{E}$, exhibiting occasional topological fluctuations at the $T_{\min}$-timescale (**Figure 6C**). Thus, the coactivity complex $\mathcal{F}_\tau$ can preserve not only its approximate size but also its topological structure, despite the ongoing recycling of its simplexes.

As τ grows, the effective lifetimes $\tau_e^{(2)}$ and $\tau_e^{(3)}$, as well as the number of simplexes actualized at a given moment increase approximately linearly, yielding a growing "stable core" (**Figure 6**). As a result, a *complete suppression* of topological fluctuations in the coactivity complex is achieved at a finite values of $\tau = \tau_*$ (**Figure 7**), which gives a theoretical estimate for the rate of physiological transience that permits stable representations of the environment $\mathcal{E}$ (Babichev et al., 2018). This observation illustrates the phenomenon of *emergent topological stability* in flickering complexes, which may provide insight into how transient networks sustain lasting representations of stable physical reality.

(ii) Finite latency flickering coactivity complexes. An alternative model of flickering clique complexes can be built by restricting the period over which the coactivity graph is formed to a shorter "spike integration" time window ϖ (Theunissen and Miller, 1995; Hoffman et al., 2016; Perea, 2019). In such approach, the coactivity simplexes that emerge within the starting ϖ-period, ϖ_1, will constitute a coactivity complex $\mathcal{F}(\varpi_1)$; the simplexes appearing within the next window, ϖ_2 will form the complex $\mathcal{F}(\varpi_2)$ and so forth. A given clique-simplex ς (as defined by the set of its vertexes) may therefore appear through a chain of consecutive windows, $\varpi_1, \varpi_2, \ldots, \varpi_{k-1}$, then disappear at the k^{th} step ϖ_k (i.e., $\varsigma \in \mathcal{F}(\varpi_{k-1})$, but $\varsigma \notin \mathcal{F}(\varpi_k)$), then reappear in a later window $\varpi_{l\geq k}$, then disappear again, and so forth. The duration of ς's existence between its k-th consecutive appearance and disappearance, $\delta t_{\varsigma,k}$, can then be as short as the shift between the consecutive windows $\Delta\varpi$ or as long as the animal's navigation session.

It is natural to view the individual, "instantaneous" complexes $\mathcal{F}(\varpi_i)$ as instantiations of a single "finite latency" flickering coactivity complex, $\mathcal{F}(\varpi_i) = \mathcal{F}_\varpi(t_i)$. As it turns out, such complexes exhibit a number similarities with the decaying complexes $\mathcal{F}_\tau(t)$, e.g., for $\varpi \geq T_{\min}(\mathcal{T})$ the pool of maximal simplexes is renewed at about ϖ timescale, but the net number of simplexes contained in $\mathcal{F}_\varpi(t)$ changes within about $5 - 10\%$ of its mean value (**Figure 8A**). Biologically, this implies that a cell assembly network that described by $\mathcal{F}_\varpi(t)$ fully rewires in about a ϖ period, without changing its overall size. Specifically, for ϖ exceeding the perennial learning time $T_{min}(\mathcal{T})$ and small

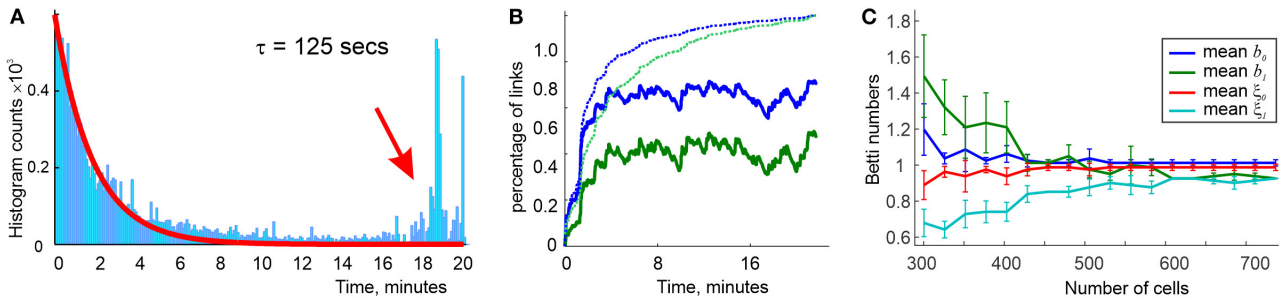

FIGURE 6 | Topological dynamics of the decaying coactivity complex. **(A)** The histogram of the time intervals between the connections' consecutive appearances and disappearances: the lifetimes the rapidly "fluttering" simplexes are distributed exponentially (red line fit). The population of the "survivor" links produces a bulging tail of the distribution (red arrow). **(B)** The population of 1*D* (blue trace) and 2*D* (green trace) simplexes in the decaying "flickering" complex $\mathcal{F}_\tau(t)$, compared to the population of 1*D* and 2*D* simplexes in the perennial complex $\mathcal{T}(t)$ (dashed lines). The size of $\mathcal{F}_\tau(t)$ remains dynamic, whereas $\mathcal{T}(t)$ saturates in about 10 min. **(C)** At $\tau = 125$ s decay period, the mean Betti numbers $b_0(\mathcal{F}_\tau)$ (blue) and $b_1(\mathcal{F}_\tau)$ (green) converge to their "physical" values $b_{0,1} = 1$ as the active place cell population increases from $N = 300$ to $N = 750$ units; the $b_{0,1}$-fluctuations decrease (shrinking error bars) and the rates $\xi_{0,1}$ of producing the correct Betti values grow to nearly 100%. Statistics evaluated over a $T = 25$ min navigation period.

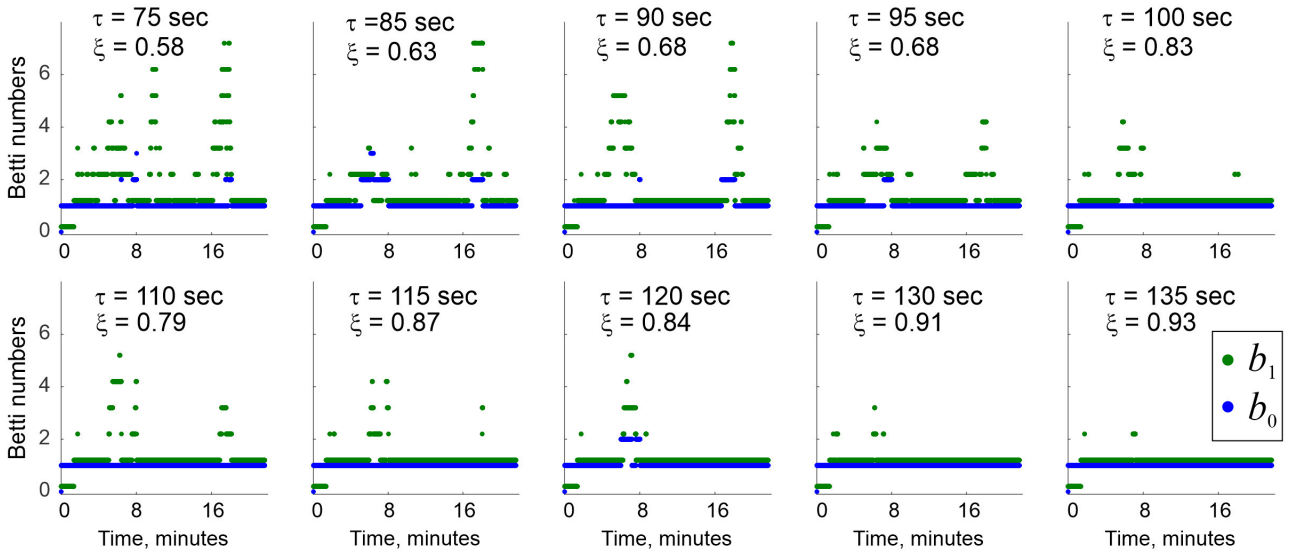

FIGURE 7 | Topological stabilization. As the decay constant τ grows from ~ 75 to ~ 135 s, the topological shape of $\mathcal{F}_\tau(t)$ stabilizes. A complete suppression of topological fluctuations is achieved for $\tau \approx 2 - 2.5$ min, with the other system parameters (rat's speed, place cell firing rates, place field sizes, etc.) within a physiological range. Blue and green dots show Betti numbers b_0 and b_1 at select moments of time. The ξ values show the percentage of times when physically correct topological signature was captured.

time steps $\Delta\varpi \gtrsim 0.01\varpi$, the intervals $\delta t_{\varsigma,k}$, as well as their means, $t_\varsigma = \langle \delta t_{\varsigma,k} \rangle_k$, are exponentially distributed, which allows characterizing the simulated cell assemblies by a half-life, τ_ϖ that typically varies within $\tau_\varsigma \approx 3 - 20$ s. As ϖ widens, the mean lifetimes t_ς of the maximal simplexes grow, and vice versa, as the memory window shrinks, simplex-flickerings intensify.

On the other hand, the large-scale shape of $\mathcal{F}_\varpi(t)$ is much more stable than its individual simplexes, as in the case of the "decay model" (2). The topological fluctuation reduce with growing ϖ, and, for sufficiently long latency periods $\varpi \geq \varpi_* \approx 1.5 T_{\min}$ they tend to *disappear completely* (**Figure 8B**)—even though the simplexes' lifetimes remain short ($\tau^*_\varpi \approx 15$ s for the environment illustrated on **Figure 5A**). For sufficiently long latencies, $\varpi \gtrsim 1.2\, T_{\min}(\mathcal{T})$, the time required to produce physical barcode $\mathfrak{b}(\mathcal{F}_\varpi) = \mathfrak{b}(\mathcal{E})$ within typical window ϖ_k is similar to the perennial learning time, $\bar{T}_{min} = \langle T^{(k)}_{min} \rangle_k \approx T_{\min}(s, f, N)$, with a variance of about 20 − 40% of the mean, which shows that topological dynamics of the simulated cognitive maps is largely time-invariant. In plain words, this result shows that accumulation of the topological information can start at any point (e.g., at the onset of the navigation or after an exploratory delay) and produce the desired stable map after about the same period of learning. In effect, this observation justifies using perennial coactivity complexes for estimating T_{min} in Dabaghian et al. (2012), Arai et al. (2014), Basso et al. (2016), Hoffman et al. (2016), and Dabaghian (2019).

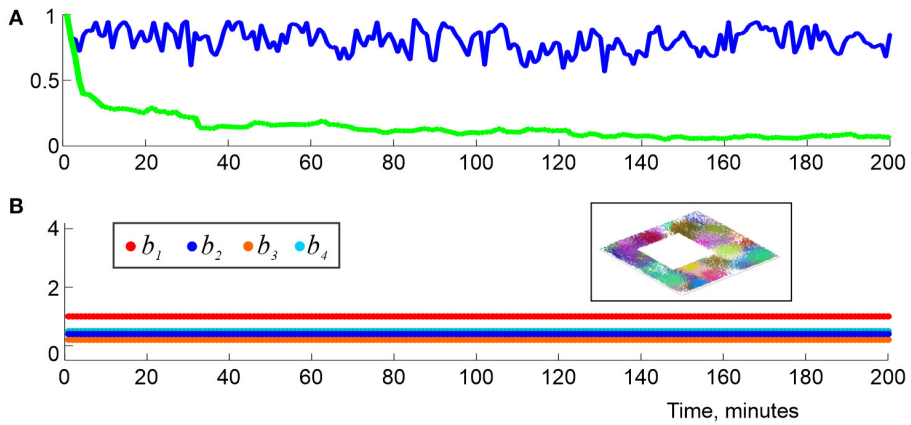

FIGURE 8 | Topological dynamics in the finite latency flickering complexes. **(A)** At each moment, t_n, the blue line shows the proportion of maximal simplexes of $\mathcal{F}_{\varpi}(t_n)$ that were present at the previous time t_{n-1}, i.e., in $\mathcal{F}_{\varpi}(t_{n-1})$. The green line shows the proportion of maximal simplexes of $\mathcal{F}_{\varpi}(t_n)$ that were present at the onset of the navigation, i.e., in $\mathcal{F}_{\varpi}(t_1)$. The latter population changes by about 95% in about 2 min. **(B)** The low-dimensional Betti numbers, b_1, b_2, b_3, and b_4 (colors shown in the left box) as a function of time, computed using $\varpi = 1.5T_{\min}$ demonstrate full topological stabilization of $\mathcal{F}_{\varpi}(t)$, whose shape fully matches the topological shape of the underlying environment (right box) at all times. Here (s, f, N) is $(23, 28, 350)$.

Moreover, for these latencies $[\varpi \gtrsim \varpi_{crit} \approx 1.4T_{\min}(\mathcal{T})]$, the instantaneous learning times $T^{(k)}_{min}$ become ϖ-independent, i.e., the finite latency model can provide a *parameter-free* characterization of the time required by a network of place cell assemblies to represent the topology of the environment and establishes the timescale for the topological fluctuations in the simulated cognitive map.

Note that finite latency model (*ii*) cannot be naïvely reduced to the decay model (*i*) by fixing the links' lifetimes, i.e., by using the decay probability

$$p_{-}(t) = \begin{cases} 1 & \text{if } t = \tau \\ 0 & \text{if } t \neq \tau. \end{cases} \tag{3}$$

The topological structure of the "quenched-decay" coactivity complex $\mathcal{F}^{*}_{\tau}(t)$ controlled by the distribution (3) exhibits more unstable dynamics than either $\mathcal{F}_{\varpi}(t)$ or $\mathcal{F}_{\tau}(t)$, even for the τ-values that reliably produce physical Betti numbers for the exponentially distributed lifetimes. As decay slows down (i.e., as τ grows), the population of survivor links produced by (3) also grows and the topological structure of $\mathcal{F}^{*}_{\tau}(t)$ eventually stabilizes; nevertheless, robust Betti numbers appear at much higher values of τ than with the exponentially decaying links, and the match between them and the physical Betti numbers is much less frequent. Thus, the statistical spread of the connections' lifetimes produced by the tail of the exponential distribution (2) plays an important role in attaining the net complex' stability, i.e., that a certain "synaptic disorder" is required for effective learning (Chowdhury et al., 2018).

Overall, the model suggests that although many details of topological dynamics of flickering complexes may depend on the simplexes' lifetimes and other parameters, several qualitative features, notably the emergent topological stability of $\mathcal{F}(t)$ are *universal*, i.e., largely independent from the simplex-recycling mechanisms. In fact, even if the functional connections between place cells are established and pruned randomly, at a rate that matches the statistics (2), the resulting random connectivity graph $\mathcal{G}_r(t)$ produces a random clique complex $\mathcal{F}_r(t)$ whose Betti numbers converge to the Betti numbers of the environment at the same timescale as the Betti numbers of $\mathcal{F}_{\tau}(t)$ or $\mathcal{F}_{\varpi}(t)$, exhibiting similar pattern of the topological fluctuations. Importantly, in the latter case, the details of these processes are controlled by the physiological parameters, e.g., by the number of active cells and their firing rates (see **Figure 6C** and Babichev et al., 2018, 2019).

3.8. Memory Spaces

In the above discussion, the coactivity complexes were used to describe topological structure of the hippocampal spatial memory frameworks—cognitive maps (Moser et al., 2008; Schmidt and Redish, 2013). However, it is well-known that hippocampus encodes not only spatial but also generic, non-spatial memories (Wood et al., 2000; Ginther et al., 2011; Wixted et al., 2018; Wu et al., 2020), embedding them into broader contexts, placing them in sequence of preceding and succeeding events (Agster et al., 2002; Fortin et al., 2002). In Eichenbaum et al. (1999) it was suggested that the resulting integrated memory structure may be viewed as a *memory space* $\mathcal{M}$ that subjects can "mentally explore" or "mentally navigate" (Theves et al., 2020). In other words, it was suggested that individual memory episodes and the spatiotemporal relationships between them may be viewed as "locations" or "regions" that may overlap, contain one another or be otherwise related in a spatial manner (Babichev and Dabaghian , 2018). In particular, the standard spatial inferences that enable spatial cognition and behavior are viewed as particular examples of the memory space navigations (Johnson and Redish, 2007; Hopfield, 2010; Issa and Zhang, 2012; Dabaghian, 2016).

From a physiological perspective, the fact that a memory space associated with a given environment $\mathcal{E}$ is encoded by the same place cell population that produces a cognitive map of $\mathcal{E}$,

suggests that the corresponding coactivity complex $\mathcal{T}_{CA}$ may be used to represent both structures. To gain an insight into this representation, notice that any simplicial complex, in particular $\mathcal{T}_{CA}$, defines a finite topological space $\mathcal{A}(\mathcal{T}_{CA})$, endowed with Alexandrov topology: the locations in $\mathcal{A}(\mathcal{T}_{CA})$ correspond to the coactivity simplexes and the topological neighborhoods of a given location represented by a simplex ς are formed by the locations whose simplexes include ς (Alexandroff, 1937; Babichev and Dabaghian , 2018). Since the simplexes of $\mathcal{T}_{CA}$ represent combinations of coactive place cells, which, in turn, are believed to represent memory elements, one may view the resulting "topological space of coactivities" $\mathcal{A}(\mathcal{T}_{CA})$ as a representation of the topological memory space encoded by the corresponding cell assembly network, $\mathcal{M} = \mathcal{A}(\mathcal{T}_{CA})$. There are three immediate implications of this construction.

i. The dynamics of the large-scale topological structure of memory space can be inferred directly from the algebro-topological studies of the corresponding coactivity complexes, since the (singular) homologies of $\mathcal{M}(\mathcal{T}_{CA})$ are identical to the (simplicial) homologies of the coactivity complex $\mathcal{T}_{CA}$ (McCord, 1966; Stong, 1966; Babichev et al., 2016a). This implies, e.g., that a memory space that contains a topological map of a given environment emerges over the same learning period $T_{\min}$ and within the same scope of spiking parameters $\mathcal{L}$ as the cognitive map, that it is similarly affected by the brain waves, by the deteriorating synapses, etc., and by the remappings (Babichev and Dabaghian , 2018).

ii. It can be shown that neuronal activity representing a trajectory γ traced by the animal in physical space maps continuously into path $\wp$ navigated in the Alexandrov topology of the memory space $\mathcal{M}(\mathcal{T}_{CA})$. This provides a theoretical base for the intuition of "mental exploration," allowing to interpret the succession of the place cell activities as a representation of a continuous succession of memory episodes (Samsonovich and McNaughton, 1997; Issa and Zhang, 2012; Buzsáki et al., 2014; Dabaghian, 2016).

iii. In neuroscience literature it is recognized that "*space is constructed in the brain rather than perceived, and the hippocampus is central to this construction*," and yet its meaning remains unclear: "*how can spaceless data enter the hippocampal system and spatial cognitive maps come out*" (O'Keefe and Nadel, 1978; Nadel and Hardt, 2004). The topological model may shed light on these problems, because it allows interpreting spatiality *intrinsically*, as a certain *relational* structure defined on spiking activity (Vickers, 1989; Roeper, 1997; Cohn and Hazarika, 2001), thus providing an ontological foundation for the emergent spatiality of the cognitive map, mentioned in the Introduction.

4. DISCUSSION

Extensive studies are dedicated to establishing correlations between parameters of neuronal activity and the characteristics of cognitive phenomena that emerge from this activity (Postle, 2006). The approach discussed above aims at filling the "semantic gap" between these two scales of information processing within a unified framework, based on the conjecture about topological nature of the hippocampal memory organization (Dabaghian et al., 2014; Babichev et al., 2016b; Babichev and Dabaghian , 2018). A formal connection with the realm of simplicial topology is made based on an observation that neuronal computations may be described as operations over spike combinations—which ones are produced over a given period, which ones are detected or transformed into specific outputs, etc. Viewing each particular collection of spikes as an abstract simplex allows representing large volumes of spiking data as abstract simplicial complexes whose topological properties describe the net qualitative information emerging at the neuronal ensemble level. With this approach, the simplicial complex' dynamics may be used as a metaphor for the learning processes, which permits not only phenomenological descriptions at different spatiotemporal scales but also possesses explanatory power, i.e., allows embedding empirical data into qualitative and quantitative schemas for reasoning about cognitive phenomena.

The framework also allows describing the flow of information in transient networks, which significantly expands the scope of the modeled phenomena. The net structure of this information is represented by flickering coactivity complexes that exhibit topological dynamics at three complementary timescales. The fastest timescale corresponds to rapid recycling of the local connections, which represents the flow of the ongoing, temporary information—the *short-term memory* (Hebb, 1949; Cowan, 2008). The net topological dynamics, described by the time-dependent invariants, e.g., Betti numbers, unfolds at a timescale that is by about an order of magnitude slower than the simplex-level fluctuations. Physiologically, this "operational" timescale corresponds to the *intermediate-term memory* (Eichenbaum et al., 1994; Kesner and Hunsaker, 2010). Lastly, the topological variations occur over a robust base that marks persistent, qualitative characteristics that marks the *long-term memory*. Such stratification indicates functional importance of the complementary learning systems for processing information at different levels of spatiotemporal granularity (O'Reilly and McClelland, 1994; McClelland et al., 1995; Fusi et al., 2005).

The model reveals complex interactions between these dynamics; for example, for sufficiently slow transience rates, the fluctuations of the topological shapes encoded by the network freeze out, i.e., the simulated cognitive map can acquire topological stability. Physiologically, this implies that if the cell assemblies rewire sufficiently slowly, then the net map encoded by the corresponding network may retain its structure despite the recycling connections in its neuronal substrate. In other words, the model suggests that synaptic and structural plasticity, which are ultimately responsible for the network's ability to incorporate new information (McHugh and Tonegawa, 2009; Leuner et al., 2010; Schaefers et al., 2010), do not necessarily compromise the qualitative information represented by the system. Rather, renewing connections allow correcting errors, e.g., removing spurious topological defects that may have appeared by an accident. As a result, a network capable of recycling information demonstrates better learning capacity, suggesting that both learning and forgetting components are necessary for physiological learning (Dupret et al., 2010; Kuhl

et al., 2010; Murre et al., 2013). The model also suggests that memory deterioration caused by an overly rapid decay of the network's connections may be compensated by increasing neuronal activity, e.g., by boosting the neuronal firing rates (Babichev et al., 2018) or by increasing the "off-line," endogenous activity of the hippocampal network that can occur in wake or in sleep states (Ji and Wilson, 2007; Karlsson and Frank, 2009; Dragoi and Tonegawa, 2011, 2013). In certain contexts, such *replays* can be viewed as manifestations of the animal's "mental explorations" of its cognitive map (Foster and Wilson, 2006; Johnson and Redish, 2007; Hopfield, 2010; Issa and Zhang, 2012; Dabaghian, 2016), which are believed to help learning and memory consolidation (Girardeau et al., 2010; Roux et al., 2017). Indeed, the model shows that frequent place cell replays significantly reduce the structural fluctuations in the cognitive map, thus helping to separate the fast and the slow timescales and to extract stable, qualitative representation of the external world (Babichev et al., 2019).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

AUTHOR CONTRIBUTIONS

YD conceived and wrote this manuscript.

FUNDING

The work was supported by the NSF grants 1422438 and 1901338.

ACKNOWLEDGMENTS

The author thanks V. Brant for helping with the text and his collaborators, notably A. Babichev.

REFERENCES

Agarwal, G., Stevenson, I., Berényi, A., Mizuseki, K., Buzsáki, G., and Sommer, F. (2014). Spatially distributed local fields in the hippocampus encode rat position. *Science* 344, 626–630. doi: 10.1126/science.1250444

Agster, K., Fortin, N., and Eichenbaum, H. (2002). The hippocampus and disambiguation of overlapping sequences. *J. Neurosci.* 22, 5760–5768. doi: 10.1523/JNEUROSCI.22-13-05760.2002

Albert, R., and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Rev. Mod. Phys.* 74, 47–97. doi: 10.1103/RevModPhys.74.47

Alexandroff, P. (1928). Untersuchungen über Gestalt und Lage abgeschlossener Mengen beliebiger Dimension. *Ann. Math.* 30, 101–187. doi: 10.2307/1968272

Alexandroff, P. (1937). Diskrete Räume. *Rec. Math.* 2, 501–518.

Alexandrov, P. (1965). *Elementary Concepts of Topology*. New York, NY: F. Ungar Pub. Co.

Alvernhe, A., Sargolini, F., and Poucet, B. (2012). Rats build and update topological representations through exploration. *Anim. Cogn.* 15, 359–368. doi: 10.1007/s10071-011-0460-z

Ang, C., Carlson, G., and Coulter, D. (2005). Hippocampal CA1 circuitry dynamically gates direct cortical inputs preferentially at theta frequencies. *J. Neurosci.* 25, 9567–9580. doi: 10.1523/JNEUROSCI.2992-05.2005

Arai, M., Brandt, V., and Dabaghian, Y. (2014). The effects of theta precession on spatial learning and simplicial complex dynamics in a topological model of the hippocampal spatial map. *PLoS Comput. Biol.* 10:e1003651. doi: 10.1371/journal.pcbi.1003651

Babichev, A., Cheng, S., and Dabaghian, Y. (2016b). Topological schemas of cognitive maps and spatial learning. *Front. Comput. Neurosci.* 10:18. doi: 10.3389/fncom.2016.00018

Babichev, A., and Dabaghian, Y. (2017a). Persistent memories in transient networks. *Springer Proc. Phys.* 191, 179–188 doi: 10.1007/978-3-319-47810-4_14

Babichev, A., and Dabaghian, Y. (2017b). Transient cell assembly networks encode stable spatial memories. *Sci. Rep.* 7:3959. doi: 10.1038/s41598-017-03423-3

Babichev, A., and Dabaghian, Y. (2018). Topological schemas of memory spaces. *Front. Comput. Neurosci.* 12:27. doi: 10.3389/fncom.2018.00027

Babichev, A., Ji, D., Memoli, F., and Dabaghian, Y. (2016a). A topological model of the hippocampal cell assembly network. *Front. Comput. Neurosci.* 10:50. doi: 10.3389/fncom.2016.00050

Babichev, A., Morozov, D., and Dabaghian, Y. (2018). Robust spatial memory maps encoded by networks with transient connections. *PLoS Comput. Biol.* 14:e1006433. doi: 10.1371/journal.pcbi.1006433

Babichev, A., Morozov, D., and Dabaghian, Y. (2019). Replays of spatial memories suppress topological fluctuations in cognitive map. *Netw. Neurosci.* 3, 707–724. doi: 10.1162/netn_a_00076

Barabási, A.-L, and Albert, R. (1999). Emergence of scaling in random networks. *Science* 286, 509–512. doi: 10.1126/science.286.5439.509

Barbieri, R., Frank, L., Nguyen, D., Quirk, M., Solo, V., Wilson, M., et al. (2004). Dynamic analyses of information encoding in neural ensembles. *Neural Comput.* 16, 277–307. doi: 10.1162/089976604322742038

Barbour, B., Brunel, N., Hakim, V., and Nadal, J.-P. (2007). What can we learn from synaptic weight distributions? *Trends Neurosci.* 30, 622–629. doi: 10.1016/j.tins.2007.09.005

Basso, E., Arai, M., and Dabaghian, Y. (2016). Gamma synchronization influences map formation time in a topological model of spatial learning. *PLoS Comput. Biol.* 12:e1005114. doi: 10.1371/journal.pcbi.1005114

Battiston, F., Cencetti, G., Iacopini, I., Latora, V., Lucash, M., Patania, A., et al. (2020). Networks beyond pairwise interactions: structure and dynamics. *Phys. Rep.* 874, 1–92. doi: 10.1016/j.physrep.2020.05.004

Bellmund, J., de Cothi, W., Ruiter, T., Nau, M., Barry, C., and Doeller, C. (2020). Deforming the metric of cognitive maps distorts memory. *Nat. Hum. Behav.* 4, 177–188. doi: 10.1038/s41562-019-0767-3

Bennett, S., Kirby, A., and Finnerty, G. (2018). Rewiring the connectome: evidence and effects. *Neurosci. Biobehav. Rev.* 88, 51–62. doi: 10.1016/j.neubiorev.2018.03.001

Bi, G., and Poo, M. (2001). Synaptic modification by correlated activity: Hebb's postulate revisited. *Annu. Rev. Neurosci.* 24, 139–166. doi: 10.1146/annurev.neuro.24.1.139

Billeh, Y., Schaub, M., Anastassiou, C., Barahona, M., and Koch, C. (2014). Revealing cell assemblies at multiple levels of granularity. *J. Neurosci. Methods* 236, 92–106. doi: 10.1016/j.jneumeth.2014.08.011

Bonifazi, P., Goldin, M., Picardo, M., Jorquera, I., Cattani, A., Bianconi, G., et al. (2009). GABAergic hub neurons orchestrate synchrony in developing hippocampal networks. *Science* 326, 1419–1424. doi: 10.1126/science.1175509

Brette, R. (2012). Computing with neural synchrony. *PLoS Comput. Biol.* 8:e1002561. doi: 10.1371/journal.pcbi.1002561

Brown, E., Frank, L., Tang, D., Quirk, M., and Wilson, M. (1998). A statistical paradigm for neural spike train decoding applied to position prediction from ensemble firing patterns of rat hippocampal place cells. *J. Neurosci.* 18, 7411–7425. doi: 10.1523/JNEUROSCI.18-18-07411.1998

Brown, M., Farley, R., and Lorek, E. (2007). Remembrance of places you passed: social spatial working memory in rats. *J. Exp. Psychol. Anim. Behav. Process.* 33, 213–224. doi: 10.1037/0097-7403.33.3.213

Brunel, N., Hakim, V., Isope, P., Nadal, J.-P., and Barbour, B. (2004). Optimal information storage and the distribution of synaptic weights: perceptron versus purkinje cell. *Neuron* 43, 745–757. doi: 10.1016/S0896-6273(04)00528-8

Burgess, N., and O'Keefe, J. (1996). Cognitive graphs, resistive grids, and the hippocampal representation of space. *J. Gen. Physiol.* 107, 659–662. doi: 10.1085/jgp.107.6.659

Buzsáki, G. (2002). Theta oscillations in the hippocampus. *Neuron* 33, 325–340. doi: 10.1016/S0896-6273(02)00586-X

Buzsáki, G. (2005). Theta rhythm of navigation: link between path integration and landmark navigation, episodic and semantic memory. *Hippocampus* 15, 827–840. doi: 10.1002/hipo.20113

Buzsaki, G. (2010). Neural syntax: cell assemblies, synapsembles, and readers. *Neuron* 68, 362–385. doi: 10.1016/j.neuron.2010.09.023

Buzsáki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929. doi: 10.1126/science.1099745

Buzsáki, G., and Mizuseki, K. (2014). The log-dynamic brain: how skewed distributions affect network operations. *Nat. Rev. Neurosci.* 15, 264–278. doi: 10.1038/nrn3687

Buzsáki, G., Peyrache, A., Kubie, J. (2014). Emergence of cognition from action. *Cold Spring Harb. Symp. Quant. Biol.* 79, 41–50. doi: 10.1101/sqb.2014.79.024679

Buzsáki, G., and Wang, X. (2012). Mechanisms of gamma oscillations. *Annu. Rev. Neurosci.* 35, 203–225. doi: 10.1146/annurev-neuro-062111-150444

Cacucci, F., Yi, M., Wills, T., Chapman, P., and O'Keefe, J. (2008). Place cell firing correlates with memory deficits and amyloid plaque burden in Tg2576 Alzheimer mouse model. *Proc. Natl. Acad. Sci. U.S.A.* 105, 7863–7868. doi: 10.1073/pnas.0802908105

Carlsson, G. (2009). Topology and data. *Bull. Amer. Math. Soc.* 46, 255–308. doi: 10.1090/S0273-0979-09-01249-X

Carlsson, G., and Silva, Vd. (2010). Zigzag persistence. *Found. Comput. Math*. 10, 367–405. doi: 10.1007/s10208-010-9066-0

Carlsson, G., Silva, Vd., and Morozov, D. (2009). "Zigzag persistent homology and real-valued functions," in *Proceedings of the 25th Annual Symposium on Computational Geometry* (Aarhus: ACM), 247–256. doi: 10.1145/1542362.1542408

Čech, E. (1932). Théorie générale de l'homologie dans un espace quelconque. *Fundam. Math.* 19, 149–183. doi: 10.4064/fm-19-1-149-183

Chen, Z., Kloosterman, F., Brown, E., and Wilson, M. (2012). Uncovering spatial topology represented by rat hippocampal population neuronal codes. *J. Comput. Neurosci.* 33, 227–255. doi: 10.1007/s10827-012-0384-x

Chowdhury, S., Dai, B., and Memoli, F. (2018). The importance of forgetting: limiting memory improves recovery of topological characteristics from neural data. *PLoS ONE* 13:e0202561. doi: 10.1371/journal.pone.0202561

Clayton, N., Bussey, T., and Dickinson, A. (2003). Can animals recall the past and plan for the future? *Nat. Rev. Neurosci.* 4, 685–691. doi: 10.1038/nrn1180

Cohen, R., Rezai-Zadeh, K., Weitz, T., Rentsendorj, A., Gate, D., Spivak, I., et al. (2013). A transgenic Alzheimer rat with plaques, tau pathology, behavioral impairment, oligomeric Aβ, and Frank neuronal loss. *J. Neurosci.* 33, 6245–6256. doi: 10.1523/JNEUROSCI.3672-12.2013

Cohn, A. G., and Hazarika, S. M. (2001). Qualitative spatial representation and reasoning: an overview. *Fundam. Inf.* 46, 1–29. doi: 10.5555/1219982.1219984

Colgin, L., Denninger, T., Fyhn, M., Hafting, T., Bonnevie, T., Jensen, O., et al. (2009). Frequency of gamma oscillations routes flow of information in the hippocampus. *Nature* 462, 353–357. doi: 10.1038/nature08573

Colgin, L., and Moser, E. (2010). Gamma oscillations in the hippocampus. *Physiology* 25, 319–329. doi: 10.1152/physiol.00021.2010

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Prog. Brain Res.* 169, 323–338. doi: 10.1016/S0079-6123(07)00020-9

Curto, C., and Itskov, V. (2008). Cell groups reveal structure of stimulus space. *PLoS Comput. Biol.* 4:e1000205. doi: 10.1371/journal.pcbi.1000205

Dabaghian, Y. (2016). Maintaining consistency of spatial information in the hippocampal network: a combinatorial geometry model. *Neural Comput.* 28, 1051–1071. doi: 10.1162/NECO_a_00840

Dabaghian, Y. (2019). Through synapses to spatial memory maps: a topological model. *Sci. Rep.* 9:572. doi: 10.1038/s41598-018-36807-0

Dabaghian, Y., Brandt, V., and Frank, L. (2014). Reconceiving the hippocampal map as a topological template. *eLife* 3:e03476. doi: 10.7554/eLife.03476.009

Dabaghian, Y., Memoli, F., Frank, L., and Carlsson, G. (2012). A topological paradigm for hippocampal spatial map formation using persistent homology. *PLoS Comput. Biol.* 8:e1002581. doi: 10.1371/journal.pcbi.1002581

De Silva, V., and Ghrist, R. (2007). Coverage in sensor networks via persistent homology. Algebr. Geometr. Topol. 7, 339–358. doi: 10.2140/agt.2007.7.339

Dilgen, J., Tompa, T., Saggu, S., Naselaris, T., and Lavin, A. (2013). Optogenetically evoked gamma oscillations are disturbed by cocaine administration. *Front. Cell Neurosci.* 7:213. doi: 10.3389/fncel.2013.00213

Donato, I., Gori, M., Pettini, M., Petri, G., De Nigris, S., Franzosi, R., et al. (2016). Persistent homology analysis of phase transitions. *Phys. Rev. E* 93:052138. doi: 10.1103/PhysRevE.93.052138

Dragoi, G., and Tonegawa, S. (2011). Preplay of future place cell sequences by hippocampal cellular assemblies. *Nature*469, 397–401. doi: 10.1038/nature09633

Dragoi, G., and Tonegawa, S. (2013). Distinct preplay of multiple novel spatial experiences in the rat. *Proc. Natl. Acad. Sci. U.S.A.* 110, 9100–9105. doi: 10.1073/pnas.1306031110

Dupret, D., Pleydell-Bouverie, B., and Csicsvari, J. (2010). Rate remapping: when the code goes beyond space. *Neuron* 68, 1015–1016. doi: 10.1016/j.neuron.2010.12.011

Eckert, M., and Abraham, W. (2010). Physiological effects of enriched environment exposure and LTP induction in the hippocampus *in vivo* do not transfer faithfully to *in vitro* slices. *Learn. Mem.* 17, 480–484. doi: 10.1101/lm.1822610

Edelsbrunner, H., and Harer, J. (2010). Computational topology: an introduction. *Am. Math. Soc.* (Cambridge, NY: Cambridge University Press), 241. doi: 10.1090/mbk/069

Edelsbrunner, H., Letscher, D., and Zomorodian, A. (2002). Topological persistence and simplification. *Discrete Computat. Geom.* 28, 511—533. doi: 10.1007/s00454-002-2885-2

Eichenbaum, H., Dudchenko, P., Wood, E., Shapiro, M., and Tanila, H. (1999). The hippocampus, memory, and place cells: is it spatial memory or a memory space? *Neuron* 23, 209–226. doi: 10.1016/S0896-6273(00)80773-4

Eichenbaum, H., Otto, T., and Cohen, N. (1994). Two functional components of the hippocampal memory system. *Behav. Brain Sci.* 17, 449–472. doi: 10.1017/S0140525X00035391

Fenton, A., Kao, H., Neymotin, S., Olypher, A., Vayntrub, Y., Lytton, W., et al. (2008). Unmasking the CA1 ensemble place code by exposures to small and large environments: more place cells and multiple, irregularly arranged, and expanded place fields in the larger space. *J. Neurosci.* 28, 11250–11262. doi: 10.1523/JNEUROSCI.2862-08.2008

Fenton, A., and Muller, R. (1998). Place cell discharge is extremely variable during individual passes of the rat through the firing field. *Proc. Natl. Acad. Sci. U.S.A.* 95, 3182–3187. doi: 10.1073/pnas.95.6.3182

Fortin, N., Agster, K., and Eichenbaum, H. (2002). Critical role of the hippocampus in memory for sequences of events. *Nat. Neurosci.* 5, 458–462. doi: 10.1038/nn834

Foster, D., and Wilson, M. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature* 440, 680–683. doi: 10.1038/nature04587

Fusi, S., Drew, P., and Abbott, L. (2005). Cascade models of synaptically stored memories. *Neuron* 45, 599–611. doi: 10.1016/j.neuron.2005.02.001

Geisler, C., Diba, K., Pastalkova, E., Mizuseki, K., Royer, S., and Buzsáki, G. (2010). Temporal delays among place cells determine the frequency of population theta oscillations in the hippocampus. *Proc. Natl. Acad. Sci. U.S.A.* 107, 7957–7962. doi: 10.1073/pnas.0912478107

Ghrist, R. (2008). Barcodes: the persistent topology of data. *Bull. Am. Math. Soc.* 45, 61–75. doi: 10.1090/S0273-0979-07-01191-3

Ginther, M., Walsh, D., and Ramus, S. (2011). Hippocampal neurons encode different episodes in an overlapping sequence of odors task. *J. Neurosci.* 31, 2706–2711. doi: 10.1523/JNEUROSCI.3413-10.2011

Girardeau, G., Benchenane, K., Wiener, S., Buzsaki, G., and Zugaro, M. (2010). Selective suppression of hippocampal ripples impairs spatial memory. *Nat. Neurosci.* 12, 1222–1223. doi: 10.1038/nn.2384

Goldman-Rakic, P. (1995). Cellular basis of working memory. *Neuron* 14, 477–485. doi: 10.1016/0896-6273(95)90304-6

Gothard, K., Skaggs, W., and McNaughton, B. (1996). Dynamics of mismatch correction in the hippocampal ensemble code for space: interaction between path integration and environmental cues. *J. Neurosci.* 16, 8027–8040. doi: 10.1523/JNEUROSCI.16-24-08027.1996

Guger, C., Gener, T., Pennartz, C., Brotons-Mas, J., Edlinger, G., Badia, S. B. I., et al. (2011). Real-time position reconstruction with hippocampal place cells. *Front. Neurosci.* 5:85. doi: 10.3389/fnins.2011.00085

Harris, K. (2005). Neural signatures of cell assembly organization. *Nat. Rev. Neurosci.* 6, 399–407. doi: 10.1038/nrn1669

Harris, K., Henze, D., Hirase, H., Leinekugel, X., Dragoi, G., Czurkó, A., et al. (2002). Spike train dynamics predicts theta-related phase precession in hippocampal pyramidal cells. *Nature* 417, 738–741. doi: 10.1038/nature00808

Harris, K. D., Csicsvari, J., Hirase, H., Dragoi, G., and Buzsaki, G. (2003). Organization of cell assemblies in the hippocampus. *Nature* 424, 552–556. doi: 10.1038/nature01834

Hasselmo, M., Bodelon, C., and Wyble, B. (2002). A proposed function for hippocampal theta rhythm: separate phases of encoding and retrieval enhance reversal of prior learning. *Neural Comput.* 14, 793–817. doi: 10.1162/089976602317318965

Hatcher, A. (2002). *Algebraic Topology*. Cambridge: Cambridge University Press.

Hebb, D. (1949). *The Organization of Behavior; A Neuropsychological Theory*. New York, NY: Wiley.

Hiratani, N., and Fukai, T. (2014). Interplay between short- and long-term plasticity in cell-assembly formation. *PLoS ONE* 9:e101535. doi: 10.1371/journal.pone.0101535

Hoffman, K., Babichev, A., and Dabaghian, Y. (2016). A model of topological mapping of space in bat hippocampus. *Hippocampus* 26, 1345–1353. doi: 10.1002/hipo.22610

Hopfield, J. (2010). Neurodynamics of mental exploration. *Proc. Natl. Acad. Sci. U.S.A.* 107, 1648–1653. doi: 10.1073/pnas.0913991107

Huhn, Z., Orbán, G., Érdi, P., and Lengyel, M. (2005). Theta oscillation-coupled dendritic spiking integrates inputs on a long time scale. *Hippocampus* 15, 950–962. doi: 10.1002/hipo.20112

Huxter, J., Senior, T., Allen, K., and Csicsvari, J. (2008). Theta phase-specific codes for two-dimensional position, trajectory and heading in the hippocampus. *Nat. Neurosci.* 11, 587–594. doi: 10.1038/nn.2106

Issa, J., and Zhang, K. (2012). Universal conditions for exact path integration in neural systems. *Proc. Natl. Acad. Sci. U.S.A.* 109, 6716–6720. doi: 10.1073/pnas.1119880109

Jackson, J., and Redish, A. (2007). Network dynamics of hippocampal cell-assemblies resemble multiple spatial maps within single tasks. *Hippocampus* 17, 1209–1229. doi: 10.1002/hipo.20359

Jensen, O., and Lisman, J. (1996). Hippocampal CA3 region predicts memory sequences: accounting for the phase precession of place cells. *Learn. Mem.* 3, 279–287. doi: 10.1101/lm.3.2-3.279

Jensen, O., Lisman, J. (2000). Position reconstruction from an ensemble of hippocampal place cells: contribution of theta phase coding. *J. Neurophysiol.* 83, 2602–2609. doi: 10.1152/jn.2000.83.5.2602

Jezek, K., Henriksen, E., Treves, A., Moser, E., and Moser, M.-B. (2011). Theta-paced flickering between place-cell maps in the hippocampus. *Nature* 478, 246–249. doi: 10.1038/nature10439

Ji, D., and Wilson, M. (2007). Coordinated memory replay in the visual cortex and hippocampus during sleep. *Nat. Neurosci.* 10, 100–107. doi: 10.1038/nn1825

Jia, X., and Kohn, A. (2011). Gamma rhythms in the brain. *PLoS Biol.* 9:e1001045. doi: 10.1371/journal.pbio.1001045

Johnson, A., and Redish, A. (2007). Neural ensembles in CA3 transiently encode paths forward of the animal at a decision point. *J. Neurosci.* 27, 12176–12189. doi: 10.1523/JNEUROSCI.3761-07.2007

Jonsson, J. (2008). *Simplicial Complexes of Graphs*. New York, NY: Springer.

Kang, L., Xu, B., and Morozov, D. (2020). State space discovery in spatial representation circuits with persistent cohomology. *bioRxiv* 2020.2010.2006.328773. doi: 10.1101/2020.10.06.328773

Karlsson, M., and Frank, L. (2009). Awake replay of remote experiences in the hippocampus. *Nat. Neurosci.* 12, 913–918. doi: 10.1038/nn.2344

Katz, Y., Kath, W., Spruston, N., and Hasselmo, M. (2007). Coincidence detection of place and temporal context in a network model of spiking hippocampal neurons. *PLoS Comput. Biol.* 3:e234. doi: 10.1371/journal.pcbi.0030234

Kesner, R., and Hunsaker, M. (2010). The temporal attributes of episodic memory. *Behav. Brain Res.* 215, 299–309. doi: 10.1016/j.bbr.2009.12.029

König, P., Engel, A., and Singer, W. (1996). Integrator or coincidence detector? The role of the cortical neuron revisited. *Trends Neurosci.* 19, 130–137. doi: 10.1016/S0166-2236(96)80019-1

Kuhl, B., Shah, A., DuBrow, S., and Wagner, A. (2010). Resistance to forgetting associated with hippocampus-mediated reactivation during new learning. *Nat. Neurosci.* 13, 501–506. doi: 10.1038/nn.2498

Lee, I., Yoganarasimha, D., Rao, G., and Knierim, J. (2004). Comparison of population coherence of place cells in hippocampal subfields CA1 and CA3. *Nature* 430, 456–459. doi: 10.1038/nature02739

Leutgeb, J., Leutgeb, S., Treves, A., Meyer, R., Barnes, C., McNaughton, B., et al. (2005). Progressive transformation of hippocampal neuronal representations in "morphed" environments. *Neuron* 48, 345–358. doi: 10.1016/j.neuron.2005.09.007

Li, X., Ouyang, G., Usami, A., Ikegaya, Y., and Sik, A. (2010). Scale-free topology of the CA3 hippocampal network: a novel method to analyze functional neuronal assemblies. *Biophys. J.* 98, 1733–1741. doi: 10.1016/j.bpj.2010.01.013

Lisman, J. (2005). The theta/gamma discrete phase code occuring during the hippocampal phase precession may be a more general brain coding scheme. *Hippocampus* 15, 913–922. doi: 10.1002/hipo.20121

Lisman, J., and Idiart, M. (1995). Storage of 7 ± 2 short-term memories in oscillatory subcycles. *Science* 267, 1512–1515. doi: 10.1126/science.7878473

London, M., and Häusser, M. (2005). Dendritic computation. *Ann. Rev. Neurosci.* 28, 503–532. doi: 10.1146/annurev.neuro.28.061604.135703

Lu, C., Hamilton, J., Powell, A., Toescu, E., and Vreugdenhil, M. (2011). Effect of ageing on CA3 interneuron sAHP and γ oscillations is activity-dependent. *Neurobiol. Aging* 32, 956–965. doi: 10.1016/j.neurobiolaging.2009.05.006

Lum, P., Singh, G., Lehman, A., Ishkanov, T., Vejdemo-Johansson, M., Alagappan, M., et al. (2013). Extracting insights from the shape of complex data using topology *Sci. Rep.* 3:1236. doi: 10.1038/srep01236

Lundqvist, M., Herman, P., and Lansner, A. (2011). Theta and gamma power increases and alpha/beta power decreases with memory load in an attractor network model. *J. Cogn. Neurosci.* 23, 3008–3020. doi: 10.1162/jocn_a_00029

Magee, J. (2000). Dendritic integration of excitatory synaptic input. *Nat. Rev. Neurosci.* 1, 181–190. doi: 10.1038/35044552

Matthews, D., Simson, P., and Best, P. (1996). Ethanol alters spatial processing of hippocampal place cells: a mechanism for impaired navigation when intoxicated. *Alcohol Clin. Exp. Res.* 20, 404–407. doi: 10.1111/j.1530-0277.1996.tb01660.x

Maurer, A., Cowen, S., Burke, S., Barnes, C., and McNaughton, B. (2006). Organization of hippocampal cell assemblies based on theta phase precession. *Hippocampus* 16, 785–794. doi: 10.1002/hipo.20202

McClelland, J., McNaughton, B., and O'Reilly, R. (1995). Why there are complementary learning systems in the hippocampus and neocortex: insights from the successes and failures of connectionist models of learning and memory. *Psychol. Rev.* 102, 419–457 doi: 10.1037/0033-295X.102.3.419

McCord, M. C. (1966). Singular homology groups and homotopy groups of finite topological spaces. *Duke Math. J.* 33, 465–474. doi: 10.1215/S0012-7094-66-03352-7

McCracken, C., and Grace, A. (2013). Persistent cocaine-induced reversal learning deficits are associated with altered limbic cortico-striatal local field potential synchronization. *J. Neurosci.* 33, 17469–17482. doi: 10.1523/JNEUROSCI.1440-13.2013

McHugh, T., and Tonegawa, S. (2009). CA3 NMDA receptors are required for the rapid formation of a salient contextual representation. *Hippocampus* 19, 1153–1158. doi: 10.1002/hipo.20684

McHugh, T., Tonegawa, S., Leuner, B., and Gould, E. (2010). Structural plasticity and hippocampal function. *Annu. Rev. Psychol.* 61, 111–140. doi: 10.1146/annurev.psych.093008.100359

Meck, W., Church, R., and Olton, D. (2013). Hippocampus, time, and memory. *Behav. Neurosci.* 127, 655–668. doi: 10.1037/a0034188

Mizuseki, K., Sirota, A., Pastalkova, E., and Buzsáki, G. (2009). Theta oscillations provide temporal windows for local circuit computation in the entorhinal-hippocampal loop. *Neuron*, 64, 267–280. doi: 10.1016/j.neuron.2009.08.037

Moretti, D., Fracassi, C., Pievani, M., Geroldi, C., Binetti, G., Zanetti, O., et al. (2009). Increase of θ/γ ratio is associated with memory impairment. *Clin. Neurophysiol.* 120, 295–303. doi: 10.1016/j.clinph.2008.11.012

Moser, E. I., Kropff, E., and Moser M-B. (2008). Place cells, grid cells, and the brain's spatial representation system. *Annu. Rev. Neurosci.* 31, 69–89. doi: 10.1146/annurev.neuro.31.061307.090723

Muller, R., Stead, M., and Pach, J. (1996). The hippocampus as a cognitive graph. *J. Gen. Physiol.* 107, 663–694. doi: 10.1085/jgp.107.6.663

Murre, J., Chessa, A., and Meeter, M. (2013). A mathematical model of forgetting and amnesia. *Front. Psychol.* 4:76. doi: 10.3389/fpsyg.2013.00076

Nadel, L., and Hardt, O. (2004). The spatial brain. *Neuropsychology* 18, 473–476. doi: 10.1037/0894-4105.18.3.473

Neves, G., Cooke, S., and Bliss, T. (2008). Synaptic plasticity, memory and the hippocampus: a neural network approach to causality. *Nat. Rev. Neurosci.* 9, 65–75. doi: 10.1038/nrn2303

Nikoli, D., Fries, P., and Singer, W. (2013). Gamma oscillations: precise temporal coordination without a metronome. *Trends Cogn. Sci.* 17, 54–55. doi: 10.1016/j.tics.2012.12.003

Nithianantharajah, J., and Hannan, A. (2006). Enriched environments, experience-dependent plasticity and disorders of the nervous system. *Nat. Rev. Neurosci.* 7, 697–709. doi: 10.1038/nrn1970

O'Keefe, J., and Nadel, L. (1978). *The Hippocampus as a Cognitive Map*. New York, NY: Clarendon Press; Oxford University Press.

O'Neill, J., Senior, T., Allen, K., Huxter, J., and Csicsvari, J. (2008). Reactivation of experience-dependent cell assembly patterns in the hippocampus. *Nat. Neurosci.* 11, 209–215. doi: 10.1038/nn2037

O'Reilly, R., and McClelland, J. (1994). Hippocampal conjunctive encoding, storage, and recall: avoiding a trade-off. *Hippocampus* 4, 661–682. doi: 10.1002/hipo.450040605

Perea, J. (2019). Topological time series analysis. *Notices Am. Math. Soc.* 66, 686–693. doi: 10.1090/noti1869

Place, R., and Nitz, D. (2020). Cognitive maps: distortions of the hippocampal space map define neighborhoods. *Curr. Biol.* 30, R340–R342. doi: 10.1016/j.cub.2020.02.085

Postle, B. (2006). Working memory as an emergent property of the mind and brain. *Neuroscience* 139, 23–38. doi: 10.1016/j.neuroscience.2005.06.005

Pouget, A., Dayan, P., and Zemel, R. (2000). Information processing with population codes. *Nat. Rev. Neurosci.* 1, 125–132. doi: 10.1038/35039062

Ratté, S., Lankarany, M., Rho, Y.-A., Patterson, A., and Prescott, S. (2015). Subthreshold membrane currents confer distinct tuning properties that enable neurons to encode the integral or derivative of their input. *Front. Cell Neurosci.* 8:452. doi: 10.3389/fncel.2014.00452

Robbe, D., and Buzsáki, G. (2009). Alteration of theta timescale dynamics of hippocampal place cells by a cannabinoid is associated with memory impairment. *J. Neurosci.* 29, 12597–12605. doi: 10.1523/JNEUROSCI.2407-09.2009

Roeper, P. (1997). Region-based topology. *J. Philos. Logic* 26, 251–309. doi: 10.1023/A:1017904631349

Roux, L., Hu, B., Eichler, R., Stark, E., and Buzsaki, G. (2017). Sharp wave ripples during learning stabilize the hippocampal spatial map. *Nat. Neurosci.* 20, 845–853. doi: 10.1038/nn.4543

Samsonovich, A., and McNaughton, B. L. (1997). Path integration and cognitive mapping in a continuous attractor neural network model. *J Neurosci.* 17, 5900–5920. doi: 10.1523/JNEUROSCI.17-15-05900.1997

Schaefers, A., Grafen, K., Teuchert-Noodt, G., and Winter, Y. (2010). Synaptic remodeling in the dentate gyrus, CA3, CA1, subiculum, and entorhinal cortex of mice: effects of deprived rearing and voluntary running. *Neural Plast.* 2010:11. doi: 10.1155/2010/870573

Schmidt, B., and Redish, A. (2013). Neuroscience: navigation with a cognitive map. *Nature* 497, 42–43. doi: 10.1038/nature12095

Selkoe, D. (2002). Alzheimer's disease is a synaptic failure. *Science* 298, 789–791. doi: 10.1126/science.1074069

Shapiro, M. (2001). Plasticity, hippocampal place cells, and cognitive maps. *Arch. Neurol.* 58, 874–881. doi: 10.1001/archneur.58.6.874

Shepherd, G. (2004). *The Synaptic Organization of the Brain. 5th Edn.* Oxford; New York, NY: Oxford University Press.

Skaggs, W., McNaughton, B., Wilson, M., and Barnes, C. (1996). Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* 6, 149–172. doi: 10.1002/(SICI)1098-1063(1996)6:2<149::AID-HIPO6>3.0.CO;2-K

Spruston, N. (2008). Pyramidal neurons: dendritic structure and synaptic integration. *Nat. Rev. Neurosci.* 9, 206–221. doi: 10.1038/nrn2286

Stong, R. E. (1966). Finite topological spaces. *Trans. Am. Math. Soc.* 123, 325–340. doi: 10.1090/S0002-9947-1966-0195042-2

Tancer, M. (2013). "Intersection patterns of convex sets via simplicial complexes: a survey," in *Thirty Essays on Geometric Graph Theory*, ed J. Pach (New York, NY: Springer), 521–540. doi: 10.1007/978-1-4614-0110-0_28

Theunissen, F., and Miller, J. (1995). Temporal encoding in nervous systems: a rigorous definition. *J. Comput. Neurosci.* 2, 149–162. doi: 10.1007/BF00961885

Theves, S., Fernández, G., and Doeller, C. (2020). The hippocampus maps concept space, not feature space. *J. Neurosci.* 40, 7318–7325. doi: 10.1523/JNEUROSCI.0494-20.2020

Thompson, L., and Best, P. (1990). Long-term stability of the place-field activity of single units recorded from the dorsal hippocampus of freely behaving rats. *Brain Res.* 509, 299–308. doi: 10.1016/0006-8993(90)90555-P

Toth, M., Melentijevic, I., Shah, L., Bhatia, A., Lu, K., Talwar, A., et al. (2012). Neurite sprouting and synapse deterioration in the aging *Caenorhabditis elegans* nervous system. *J. Neurosci.* 32, 8778–8790. doi: 10.1523/JNEUROSCI.1494-11.2012

Touretzky, D., Weisman, W., Fuhs, M., Skaggs, W., Fenton, A., and Muller, R. (2005). Deforming the hippocampal map. *Hippocampus* 15, 41–55. doi: 10.1002/hipo.20029

Trimper, J., Stefanescu, R., and Manns, J. (2014). Recognition memory and $\theta - \gamma$ interactions in the hippocampus. *Hippocampus* 24, 341–353. doi: 10.1002/hipo.22228

Ulanovsky, N., and Moss, C. (2007). Hippocampal cellular and network activity in freely moving echolocating bats. *Nat. Neurosci.* 10, 224–233. doi: 10.1038/nn1829

van Vugt, M., Schulze-Bonhage, A., Litt, B., Brandt, A., and Kahana, M. (2010). Hippocampal gamma oscillations increase with memory load. *J. Neurosci.* 30, 2694–2699. doi: 10.1523/JNEUROSCI.0567-09.2010

Vickers, S. (1989). *Topology via Logic*. Cambridge: Cambridge University Press.

Vreugdenhil, M., and Toescu, E. (2005). Age-dependent reduction of γ oscillations in the mouse hippocampus *in vitro*. *Neuroscience* 132, 1151–1157. doi: 10.1016/j.neuroscience.2005.01.025

White, A., and Best, P. (2000). Effects of ethanol on hippocampal place-cell and interneuron activity. *Brain Res.* 876, 154–165. doi: 10.1016/S0006-8993(00)02629-9

Whittington, M., Faulkner, H., Doheny, H., and Traub, R. (2000a). Neuronal fast oscillations as a target site for psychoactive drugs. *Pharmacol Ther.* 86, 171–190. doi: 10.1016/S0163-7258(00)00038-3

Whittington, M., Traub, R., Kopell, N., Ermentrout, B., and Buhl, E. (2000b). Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *Int. J. Psychophysiol.* 38, 315–336. doi: 10.1016/S0167-8760(00)00173-2

Wills, T., Lever, C., Cacucci, F., Burgess, N., and O'Keefe, J. (2005). Attractor dynamics in the hippocampal representation of the local environment. *Science* 308, 873–876. doi: 10.1126/science.1108905

Wilson, M., and McNaughton, B. (1993). Dynamics of the hippocampal ensemble code for space. *Science* 261, 1055–1058. doi: 10.1126/science.8351520

Wixted, J., Goldinger, S., Squire, L., Kuhn, J., Papesh, M., Smith, K., et al. (2018). Coding of episodic memory in the human hippocampus. *Proc. Natl. Acad. Sci. U.S.A.* 115, 1093–1098. doi: 10.1073/pnas.1716443115

Wood, E., Dudchenko, P., Robitsek, R., and Eichenbaum, H. (2000). Hippocampal neurons encode information about different types of memory episodes occurring in the same location. *Neuron* 27, 623–633. doi: 10.1016/S0896-6273(00)00071-4

Wu, C., Schulz, E., Garvert, M., Meder, B., and Schuck, N. (2020). Similarities and differences in spatial and non-spatial cognitive maps. *PLoS Comput. Biol.* 16:e1008149. doi: 10.1371/journal.pcbi.1008149

Wu, X., and Foster, D. (2014). Hippocampal replay captures the unique topological structure of a novel environment. *J. Neurosci.* 34, 6459–6469. doi: 10.1523/JNEUROSCI.3414-13.2014

Yartsev, M., and Ulanovsky, N. (2013). Representation of three-dimensional space in the hippocampus of flying bats. *Science* 340, 367–372. doi: 10.1126/science.1235338

Zomorodian, A. (2005). *Topology for Computing*. Cambridge: Cambridge University Press.

Zomorodian, A., and Carlsson, G. (2005). Computing persistent homology. *Discrete Comput. Geom.* 33, 249–274. doi: 10.1007/s00454-004-1146-y

Conflict of Interest: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 14 April 2021
doi: 10.3389/fncom.2021.590019

A Computational Framework for Controlling the Self-Restorative Brain Based on the Free Energy and Degeneracy Principles

Hae-Jeong Park [1,2,3,4,5*†] *and Jiyoung Kang* [1,2*†]

[1] Center for Systems and Translational Brain Science, Institute of Human Complexity and Systems Science, Yonsei University, Seoul, South Korea, [2] Department of Nuclear Medicine, Yonsei University College of Medicine, Seoul, South Korea, [3] Department of Psychiatry, Yonsei University College of Medicine, Seoul, South Korea, [4] Brain Korea 21 Project, Graduate School of Medical Science, Yonsei University College of Medicine, Seoul, South Korea, [5] Department of Cognitive Science, Yonsei University, Seoul, South Korea

The brain is a non-linear dynamical system with a self-restoration process, which protects itself from external damage but is often a bottleneck for clinical treatment. To treat the brain to induce the desired functionality, formulation of a self-restoration process is necessary for optimal brain control. This study proposes a computational model for the brain's self-restoration process following the free-energy and degeneracy principles. Based on this model, a computational framework for brain control is established. We posited that the pre-treatment brain circuit has long been configured in response to the environmental (the other neural populations') demands on the circuit. Since the demands persist even after treatment, the treated circuit's response to the demand may gradually approximate the pre-treatment functionality. In this framework, an energy landscape of regional activities, estimated from resting-state endogenous activities by a pairwise maximum entropy model, is used to represent the pre-treatment functionality. The approximation of the pre-treatment functionality occurs via reconfiguration of interactions among neural populations within the treated circuit. To establish the current framework's construct validity, we conducted various simulations. The simulations suggested that brain control should include the self-restoration process, without which the treatment was not optimal. We also presented simulations for optimizing repetitive treatments and optimal timing of the treatment. These results suggest a plausibility of the current framework in controlling the non-linear dynamical brain with a self-restoration process.

Keywords: free energy principle, resting state, brain dynamics, energy landscape, self-restoration, maximum entropy model, degeneracy

OPEN ACCESS

Edited by:
Tatyana Sharpee,
Salk Institute for Biological Studies, United States

Reviewed by:
Fei He,
Coventry University, United Kingdom
Sungho Hong,
Okinawa Institute of Science and Technology Graduate University, Japan
Adeel Razi,
Monash University, Australia

***Correspondence:**
Hae-Jeong Park
parkhj@yonsei.ac.kr
Jiyoung Kang
jykang@yuhs.ac

† These authors have contributed equally to this work

Received: *31 July 2020*
Accepted: *18 March 2021*
Published: *14 April 2021*

Citation:
Park H-J and Kang J (2021) A Computational Framework for Controlling the Self-Restorative Brain Based on the Free Energy and Degeneracy Principles. Front. Comput. Neurosci. 15:590019. doi: 10.3389/fncom.2021.590019

INTRODUCTION

The goal of clinical treatment for the brain is to modify the brain circuit to yield a desirable brain function. Since the brain is a complex non-linear dynamic system, clinical treatment can be considered a control problem. For the clinical treatment to the human brain, various methods have been developed, such as thermal ablation with the high intensity focused ultrasound (Park et al., 2017), deep brain stimulation (DBS) (Park et al., 2015), vagus nerve stimulation (Yu et al., 2018),

and transcranial magnetic stimulation (TMS) (Park et al., 2013; Kar, 2019) as well as conventional medications and traditional surgical operations (Schreglmann et al., 2018). Despite the remarkable advancement of these therapeutic techniques, the optimal control of the brain by treatments has many practical challenges due to the complexity of the brain and ethical issues. Since various experiments are not allowed in the human brain, establishing optimal control procedures for the brain is very slow and limited. Therefore, optimal brain control remains mostly theoretical and based on computational models.

The brain control studies with computational models have been conducted mainly in two approaches: characterization of the network controllability and prediction of the network behavior based on the state dynamic equation of the brain. To characterize the brain network, the controllability of a brain system has been evaluated in the graph-theoretic perspective (Liu et al., 2011; Gu et al., 2017; Tang et al., 2017; Cornblath et al., 2019; Lee et al., 2019; Stiso et al., 2019; Karrer et al., 2020). This approach optimizes input signals to increase or decrease activity at some brain network nodes to induce the desired brain activity at all the brain nodes. The other approach is to predict a brain system's behavior by altering some nodes or edges of the brain. In this approach, the optimal control is determined by evaluating the outcome after removing nodes or edges or changing the system's parameters in the virtual brain model (Falcon et al., 2016; Jirsa et al., 2017; Proix et al., 2017; An et al., 2019; Olmi et al., 2019). Those two types of computational approaches on brain control have primarily focused on the immediate changes in the brain network's activity or function. Those studies, however, did not consider the fact that the brain has self-restorative plasticity, making the system resilient to external treatments or perturbations.

Self-restoration capacity in the brain has been found after damage or stress via neural, molecular, and hormonal mechanisms (Russo et al., 2012; King, 2016; Murrough and Russo, 2019). At the macroscopic level, the self-restoration process toward the initial functionality has widely been reported in clinical neuroscience, for e.g., functional recovery after stroke (Murphy and Corbett, 2009; Malone and Felling, 2020), recovery of the language capacity (Saur et al., 2006), recovery of the vision after surgery (Mikellidou et al., 2019). This self-restorative property of the brain is advantageous in protecting the brain after various external attacks (Glassman, 1987). In terms of clinical treatment, however, this self-restoration process is a bottleneck as it tends to recover the initial abnormal functionality, acting against the aim of any treatment. Examples of this anti-treatment action can be found with neurological or antipsychotic medication (Abbott, 2010) showing drug resistance in most brain disease such as schizophrenia (Potkin et al., 2020), depression (Bennabi et al., 2019), Parkinson's diseases (Vorovenci et al., 2016), and epilepsy (Lee et al., 2017). Goellner et al. (2013) showed that the late seizure recurrence after temporal lobe epilepsy surgery was as much as 48.9%. Despite the anatomical alteration by resection, the treated brains, initially showing free or reduction of abnormal function (seizure behavior) after surgical dissection, returned to the initial state of abnormal functionality in a certain period after treatment.

In this respect, the brain's self-restoration process may well be included as an essential part of the computational model of brain control. Despite the criticality of the brain's self-restoration process, research that has a self-restoration process in the control problem is hardly found. This may be partly attributable to the difficulty in defining the driving force of the self-restoration process and its mathematical formulation. In contrast to the microscopic level, where the mechanism of the self-recovery process has actively been researched in terms of neurogenesis (e.g., Mattson, 2008), a systematic understanding of the self-restoration process at the macroscopic level is still lacking. Are there any principles that we may refer to formulate the brain's self-restoration process at the macroscopic level?

In the current study, as an extension of our previous study (Kang et al., 2021), we propose a computational framework for controlling the self-restorative brain by formulating the driving force of self-restoration based on the free-energy principle (Friston et al., 2006; Friston, 2010) and the degeneracy nature of a non-linear complex system (Glassman, 1987). According to the free-energy principle, the brain acts based on a model established to minimize long-term average surprise from the external environment (Friston et al., 2006; Friston, 2010). The brain network and its behavior can be considered a result of long-term adjustment to meet environmental demands (Park and Friston, 2013). For a neural population at any level of the information hierarchy, neural populations that send signals to and receive signals from the neural population are environment to the neural population. For a neural circuit of the brain, the circuit's environment can involve lower-level and higher-level neural populations, affecting the neural circuit by sensation from the lower-level neural populations and regulation from the higher-level neural populations (Friston, 2008). The long-term demands of the environment to a neural circuit or system can be represented by the statistics of bottom-up (from the lower-level neural populations) and top-down (from the higher-level neural populations) signals. Although any clinical treatment may alter a neural circuit, the circuit's environmental demands persist and do not change rapidly even after alteration in the circuitry. Since the functionality before treatment has been developed as an optimal solution to environmental demands, we posit that the altered neural circuit gradually approximates the pretreated neural circuit's functionality while adjusting itself to meet the environmental demands after treatment.

In this framework, the pre-treatment state the brain circuit tends to recover is not the same circuit, but the functionality that the circuit has established to satisfy the external demands. In most cases, the functionality for ongoing environmental demands has to be approximated via reallocation of the reduced circuit resources after treatment (e.g., after removing a node of the circuit). In this respect, the recovery of functionality via an altered circuit can be referred to as a well-known property of the complex biological system called "degeneracy" (Glassman, 1987; Edelman and Gally, 2001). Degeneracy of a system indicates a function (or behavior) can be implemented with different network configurations (Friston and Price, 2003). Degeneracy of the self-reorganizing biological system is essential to manage and protect its functionality from damage (Marder and Goaillard,

2006; Marder et al., 2017). This study is based on the degeneracy principle as it focuses on the restoration of the pre-treatment functionality, not the pre-treatment circuit.

The final component of the current framework for brain control is to define the pre-treatment functionality of the brain. As a representation of the brain's functionality, we used the non-linear dynamics of regional activity while the brain is at rest (called resting-state). A strong relationship exists between the resting-state connectivity (or distributed patterns of endogenous activity) with task-related brain activation or connectivity (Biswal et al., 1995; Smith et al., 2009; Cole et al., 2014, 2016; Krienen et al., 2014; Park et al., 2014; Yeo et al., 2015; Tavor et al., 2016; Jung et al., 2018). Fox and Raichle (2007) argued that resting-state connectivity might serve as a potential scaffold that supports diverse configurations subserving the functional elements of a given task (Fox and Raichle, 2007). In this respect, a stabilized brain network's resting-state dynamics were considered to summarize the environmental demands established by long-term interactions with the environment.

The resting-state brain network may behave as a non-linear dynamical system with its microstate (defined by distributed regional activity pattern) transitioning over the energy landscape of multistable microstates (attractors) (Freyer et al., 2011, 2012; Rabinovich and Varona, 2011; Deco and Jirsa, 2012; Kelso, 2012; Cabral et al., 2014; Tognoli and Kelso, 2014; Deco et al., 2015; Breakspear, 2017). The non-linear dynamics of the brain circuit can be modeled in terms of non-linear interactions among nodes in the network using a pairwise maximum entropy model (MEM) (Watanabe et al., 2013, 2014a,b,c; Kang et al., 2017, 2019; Ezaki et al., 2018; Gu et al., 2018). From the pairwise MEM, we can infer the probability distribution of each microstate (an activation pattern), the microstates' energy landscape. In the energy landscape, a microstate's energy is the minus (scaled) log of its probability of occurrence. In this model, the microstate dynamics (represented in microstates' energy landscape) are emergent from the underlying complex network (or circuitry). They are considered to represent the gross functionality of a brain.

In summary, our proposal for the recovery process can be explained by the free energy principle to satisfy environmental demands by reconfiguring the remained resources after treatment according to the degeneracy principle of the complex brain. Utilizing this self-restoration model, we could develop a strategy to identify the optimal treatment target region (nodes or edges in the network) and the amount of treatment strength within a source system to be treated (e.g., a disease system) to induce microstate dynamics of the desired goal system (a healthy system). We call this procedure optimal brain control throughout the paper. In the conventional control theory problems, control signals are inputs to the system to achieve the desired system's state without changing the system parameters. Meanwhile, the optimal brain control in this study refers to adjusting the source system's model parameters to approximate the desired functionality of the goal system. In this study based on the pairwise MEM, the model parameters include the sensitivity of a brain region and interaction among brain regions, which indicate neurobiological connectivity or synaptic efficacy that modulate the input and output relationship, i.e., the functionality of the brain.

We used the term "optimal brain control" in consideration of clinical treatment settings that call for optimal selection of treatment target gray matter regions (nodes) or white matter regions (edges) and treatment strength, within limited access to the brain circuit at a time. For example, the circuit that generates epileptic seizures is the source system, and the goal system is a healthy functioning state without a seizure. The treatment target can be multiple nodes in the medication. For example, lorazepam enhances the effect of the inhibitory neurotransmitter gamma-aminobutyric acid (GABA) receptors distributed in multiple brain areas. The target can be a single node by temporal lobectomy, which removes a part of the anterior temporal lobe. Callosotomy, which dissects interhemispheric fibers, is an example of targeting edges in the network.

The current paper is composed in the following order. It begins with a mathematical description of pairwise MEM and its energy landscape analysis. We then formulated a self-restoration process and optimal control in a non-linear dynamical system. Based on this formulation, we present diverse simulations to illustrate the self-restoration process and show the effect of modeling the self-restoration process in brain control. We also present simulations for optimizing repetitive treatment strategy and its timing in consideration of the clinical practice, where any treatment is highly restricted. Using these simulations, we expected to show the construct validity of the current framework in the brain's control.

BACKGROUND

Dynamic Properties of the Brain Using Pairwise Maximum Entropy Model

To define the dynamic properties of a system, we used the energy landscape of microstate established on pairwise MEM. Here, we briefly explain the pairwise MEM model. The details for deriving the MEM of the resting-state functional magnetic resonance imaging (rsfMRI) can be found elsewhere (Watanabe et al., 2013, 2014a; Kang et al., 2017, 2019).

In the pairwise MEM model of a brain with N regions, a brain state V_k, is defined as an N-dimensional binary vector;

$$V_k = (\sigma_1, ..., \sigma_N), \tag{1}$$

where $\sigma_i = 1$ for an activated state and $\sigma_i = 0$ for an inactivated state at region i. Thus, totally 2^N states exist. An energy $E(V_k)$ of a state V_k is defined as

$$E(V_k) = -\sum_{i=1}^{N} H_i \sigma_i(V_k) - \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} J_{ij} \sigma_i(V_k) \sigma_j(V_k), \tag{2}$$

where H_i and J_{ij} are model parameters that represent weights for independent activation of region i and pairwise interaction (coactivation) between regions i and j, respectively. For simplicity, we used $A = \{H_i, J_{ij}\}|_{i=1,\cdots,N, j=1,\cdots,N}$ to express all the model parameters. These model parameters were estimated using

the maximum likelihood estimation approach. For a detailed mathematical review, see Yeh et al. (2010).

The probability of a state V_k is given by the Boltzmann distribution $p(V_k)$,

$$p\left(V_k\right) = \frac{\exp\left(-E(V_k)\right)}{\sum_{j=1}^{2^N} \exp\left(-E\left(V_j\right)\right)}. \tag{3}$$

To analyze the energy landscape of state dynamics, we defined local minima (attractors, *LM*) and occupation time ratio of each local minimum ($OCR(LM_i)$) as below.

Local Minimum

A local minimum (LM_i) of an energy landscape of a system with parameters A is a state with lower energy than its neighbor states. Neighbor states are defined as states that differ from each other by only one element (one region) of the activation state.

Occupation Time Ratio

$OCR(LM_i)$: Occupation time ratio of LM_i is the sum of probabilities of all states in the basin region of LM_i. The basin region of LM_i is the set of states that belong to the LM_i. To determine whether a state belongs to the LM_i, each element of the state is gradually changed along the energy gradient until it reaches one of the local minima.

Functional Distance Between Energy Landscapes

To measure the functional distance between the target and source systems, A^t and A^s, in the energy landscape, we defined a distance function between states in terms of dynamic properties (energy landscape) of the two systems, governed by the system's network parameters.

To focus on the functional distance between major attractors and their properties in the optimal treatment, we use a partial Kullback–Leibler (KL)-divergence defined as follows.

$$\mathcal{D}\left(A^t, A^s\right) = \sum_{k\in\mathbb{R}} p\left(V_k|A^t\right) \ln \frac{p\left(V_k|A^t\right)}{p\left(V_k|A^s\right)}, \tag{4}$$

where $\mathbb{R}$ represents a set of states that belong to basin regions of major attractors.

We also defined the similarity between two systems in terms of system parameters by the root-mean-square deviation (RMSD) of the two systems' parameter vectors.

$$RMSD(A^t, A^s) = \left|\left|A^t - A^s\right|\right|. \tag{5}$$

Recovery Process

We modeled the recovery process based on three assumptions: (1) recovery occurs by adjusting the network connectivity (interactions) of the neighbors of the treated node or edge; (2) adjustment of connectivity is performed within a range of its flexibility, and (3) recovery occurs to meet the external demands, which were represented in the state dynamics of the pretreated stabilized system.

The treatment at region m (a node or an edge, for simplicity, we call it "region") is denoted by changing the element A^p_m in the pretreated network parameters A^p with

$$A^{tr}_m \Leftarrow A^p_m + \alpha \tag{6}$$

where α is the amount of treatment. The system parameters right after treatment can be expressed as $A^t = \left\{A^{tr}_m, A^p_{\setminus m}\right\}$, where m and $\setminus m$ represent the treated and untreated regions, respectively (**Figure 2D**). m can be multiple nodes or edges. In this theoretical study, we assumed that we know how to achieve the desired level α and achieve the desired parameter A^{tr}_m. The treated state A^t of a system is the starting point of the recovery ${A_0}^r$.

The system proceeds with its recovery to minimize the functional distance between state dynamics before treatment and recovery (**Figure 2B**). When we decompose the network elements (nodes or edges) into recovery regions (strongly connected neighbors of the treated node or edge), $\mathbb{R}_m$, and unchanged (weakly or unconnected) regions, $\setminus\mathbb{R}_m$, for a treated region (node or edge) m with a treatment strength A^{tr}_m, the network state of the treated system just after treatment, $A^t = \{A^{tr}_m, A^p_{\setminus m}\}$, can be written as $A^t = \{A^t_{\mathbb{R}_m}, A^t_{\setminus\mathbb{R}_m}\}$. The recovery then begins from A^t and the recovery regions $\mathbb{R}_m$ cooperate to find an optimal parameter set $A^*_{\mathbb{R}_m}$ within a constrained bound $\mathbb{C}$ to return to the pre-treatment state A^p. This recovery process can be written as below:

$$A^*_{\mathbb{R}_m} = \arg \min_{A^{t'}_{\mathbb{R}_m}, \left|A^{t'}_{\mathbb{R}_m}\right| \le \mathbb{C}} \mathcal{D}_r\left(A^p, A^{t'} \middle| A^{t'}_{\mathbb{R}_m}\right) \tag{7}$$

$$A^{t'} = \{A^{t'}_{\mathbb{R}_m}, A^t_{\setminus\mathbb{R}_m}\}, A^r = \left\{A^*_{\mathbb{R}_m}, A^t_{\setminus\mathbb{R}_m}\right\},$$

where $\mathcal{D}_r$ indicates the distance function between the stabilized state before treatment A^p and a plausible treatment solution $A^{t'}$ by adjusting parameters $A^{t'}_{\mathbb{R}_m}$ in the recovery region $\mathbb{R}_m$ while keeping the other region $\setminus\mathbb{R}_m$ unchanged after treatment A^t. The final recovered state A^r is composed of the optimal parameter set within the recovered regions $A^*_{\mathbb{R}_m}$ and the unchanged regions of the treated system $A^t_{\setminus\mathbb{R}_m}$. Considering the limited capacity of the biological change, we restricted maximum changes at the recovered regions $A^*_{\mathbb{R}_m}$ to be <20 % of those of the previous step.

In this study, we define the pre-treatment network parameter A^p as a stable state after a long period of adaptation to the environment. The optimally recovered state A^r can be a new pre-treatment network state A^p for a subsequent treatment, after stabilization, e.g., ${A_\infty}^r = A^p$, where ∞ indicates a sufficient time for stabilization. Since the treated system may not revert completely to the pre-treatment network state by utilizing the constrained resources of the recovery regions, A^p is a function of trial number or time, moving toward the target system over a very long time scale. It should also be noted that a subsequent treatment can be applied to a system before the system is stabilized. We refer to treatment before stabilization as the treatment at the transient network stage. We considered that the transient state does not satisfy the environment's demands. In this case, the pre-treatment network parameter A^p was not

updated with the state right before treatment, but instead referred to the recent stabilized state. We utilized functional distance $\mathcal{D}_r$ to generate similar dynamics instead of generating similar network parameters for the recovering and initial systems.

The reorganization is performed by modifying the network parameters of neighbors $\mathbb{R}_m$ but within a plausible range of each parameter (connectivity)'s flexibility. In Eq. 7, we denote this with $\left|A^{t'}_{\mathbb{R}_m}\right| \leq \mathbb{C}$, indicating plausible parameters within a constrained bound $\mathbb{C}$. The amount of change at each node or edge in the recovery regions can be defined proportionally to the treated system's baseline (pretreated state). For the treated region m, we can assign less flexibility than for neighbors or assign inflexibility (i.e., no change) after treatment during the recovery process.

To measure effects of a treatment on brain dynamics, we define a recovery capacity as the difference between functional distance (partial KL divergence) of the treated state A^t and pretreated state A^p, $D\left(A^p, A^t\right)$, and the functional distance between the recovered state A^r and pretreated state A^p, $D\left(^p, A^r\right)$ as follows,

$$\text{Recovery capacity} = \triangle D = D\left(A^p, A^t\right) - D\left(A^p, A^r\right).$$

Optimal Control

A self-restorative system A^s is decomposed into the region that requires treatment m and the unaffected (untreated) region $\backslash m$ and is represented as $A^s = \{A^s_m, A^s_{\backslash m}\}$. When a target system A^g is given as the goal to achieve for a source system A^s, the optimal treatment is to find a region m and its treatment level A^{t*}_m to minimize the distance function $\mathcal{D}$ between the goal system A^g and recovered system $A^{s'r}$ that develops following the self-restoration process of the source system A^s in response to the treatment. To differentiate this from the distance function $\mathcal{D}$ between A^g and $A^{s'r}$, we use $\mathcal{D}^+$ to indicate the functional distance $\mathcal{D}$ between A^g and A^s. The optimal control is defined as below,

$$A^{t*}_m = \arg\min\ \mathcal{D}^+\left(A^g, A^{s'}\middle|A^{s'}_{m'}\right) \tag{8}$$

$$= \arg\min\ \mathcal{D}\left(A^g, A^{s'r}\middle|A^{s'}_{m'}\right), \tag{9}$$

$$A^{s'} = \left\{A^{s'}_{m'}, A^s_{\backslash m'}\right\},$$

$$A^{s'r} = \left\{A^{s'*}_{\mathbb{R}_{m'}}, A^{s'}_{\backslash \mathbb{R}_{m'}}\right\},$$

$$A^t = \{A^{t*}_m, A^s_{\backslash\, m}\},$$

where the recovered system $A^{s'r}$ is achieved following a self-restoration process after changes in the neighbors $\mathbb{R}_{m'}$ of the treated region m', according to Equation 7. The optimal control is conducted by searching for the best solution to achieve the goal system's dynamics by adjusting the parameter $A^{s'}_{m'}$ in the treated region m' while maintaining the other parameters $A^s_{\backslash m'}$ in untreated regions m' unchanged. The final treated system A^t is composed of the optimal treatment region with its strength A^{t*}_m and the unaffected parameters of the treated system $A^s_{\backslash\, m}$.

Note that the distance function $\mathcal{D}$ is defined in functional space (between energy landscapes), not in parameter space. In other words, $\mathcal{D}$ indicates a distance between the source dynamics that emerge from the source system with a parameter A^s, and the target dynamics that emerge from the goal system with a parameter A^g. From the perspective of degeneracy, the minimal distance function $\mathcal{D}$ in the dynamics space does not necessarily indicate the closeness in the network parameter space. Even though the two parameter sets, A^g and A^s, are distant in the parameter space, they can be close in the dynamics space.

Strategy for Iterative Optimal Treatment

Optimal treatment is a recursive procedure between treatment planning by the operator and the restoration process in the treated system (**Figures 1**, **2C**). The target region (node and edge) to be treated and the strength of treatment was chosen using a grid search algorithm in this study. In practice, the treatment to the system was performed by altering the MEM parameter A_i (an activity of a region H_i or a pairwise interaction J_{ij}) by an amount of α. We assumed that only neighboring nodes and edges of the treated node participate in the recovery process to return the brain dynamics to the pre-treatment state (**Figures 1B,C**). The treatment strength induces changes in the energy landscape in a non-linear manner (**Figure 2A**). When a node is altered (i.e., H_i is changed), edges that are strongly connected with the node (**Figure 1B**) undergo self-restoration steps, gradually changing the energy landscape (**Figure 2B**). When an edge is treated (i.e., J_{ij} is selected for treatment), two nodes that are connected with the treated edge and the strongly connected edges of the two nodes undergo self-restoration (**Figure 1C**). A threshold ($|J_{ij}| \geq 0.1$) was used to determine strongly connected edges. If we applied treatments multiple times, the energy landscape evolved as the iteration of treatment and restoration (**Figure 2C**).

The restoration process is an optimization procedure with reference to the stable pre-treatment state (**Figure 2B**) as shown in Equation 7. When the restoration process is saturated (no significant improvement in minimizing the functional distance by changing parameters), the saturated network becomes a new pre-treatment state for a new restoration step, i.e., ${A_\infty}^r = A^p$ (**Figure 2C**).

To implement the recovery process in Equation 7, we adopted the gradient ascent method, which is generally used to estimate the pairwise MEM model parameters from the experimental data by maximizing the log-likelihood (Watanabe et al., 2013, 2014a; Kang et al., 2017, 2019). To maximize the log-likelihood, model parameters, H_i and J_{ij}, are updated iteratively according to differences between data-driven and model-driven expectations of activations and coactivations, as shown below.

$$H_i(t+1) \leftarrow H_i(t) + \alpha_g\left(\log\langle\sigma_i\rangle - \log\langle\sigma_i\rangle_A\right), \tag{10}$$

$$J_{ij}(t+1) \leftarrow J_{ij}(t) + \alpha_g\left(\log\langle\sigma_i\sigma_j\rangle - \log\langle\sigma_i\sigma_j\rangle_A\right), \tag{11}$$

where α_g is a learning rate, $\langle\sigma_i\rangle$ and $\langle\sigma_i\sigma_j\rangle$ are expectations of activations and coactivations of the brain regions evaluated using the empirical data. From the pairwise MEM parameter A, probability $p(V_k|A)$ for each state V_k can be derived using Equation 3, based on which the expected activations and coactivations of the brain regions are derived using the

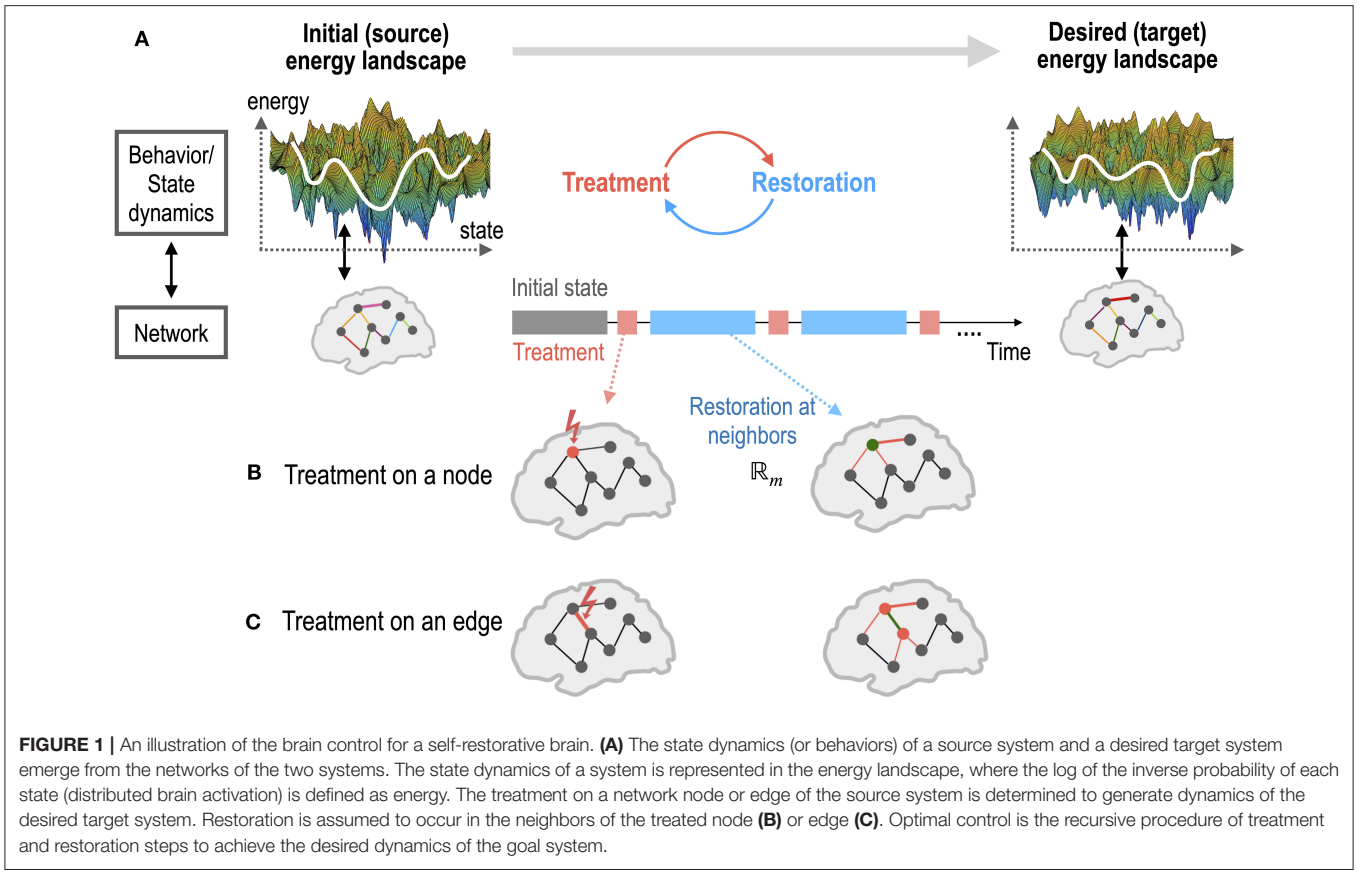

FIGURE 1 | An illustration of the brain control for a self-restorative brain. **(A)** The state dynamics (or behaviors) of a source system and a desired target system emerge from the networks of the two systems. The state dynamics of a system is represented in the energy landscape, where the log of the inverse probability of each state (distributed brain activation) is defined as energy. The treatment on a network node or edge of the source system is determined to generate dynamics of the desired target system. Restoration is assumed to occur in the neighbors of the treated node **(B)** or edge **(C)**. Optimal control is the recursive procedure of treatment and restoration steps to achieve the desired dynamics of the goal system.

following equation.

$$\langle \sigma_i \rangle_A = \sum_{k=1}^{2^N} \sigma_i(V_k) p(V_k|A), \tag{12}$$

$$\langle \sigma_i \sigma_j \rangle_A = \sum_{k=1}^{2^N} \sigma_i \sigma_j(V_k) p(V_k|A). \tag{13}$$

If the dynamics of a pretreated system A^p follows the pairwise MEM (we assumed in this study), the expected activations $\langle \sigma_i \rangle$ and coactivations $\langle \sigma_i \sigma_j \rangle$ of sufficiently large samples from the stabilized system A^p, equal to the model-driven expectations of the activation $\langle \sigma_i \rangle_{A^p}$ and coactivation $\langle \sigma_i \sigma_j \rangle_{A^p}$ of the brain regions. Then, a recovery process can be explained as follows:

$$H_i(t+1) \leftarrow H_i(t) + \alpha_g \left(\log \langle \sigma_i \rangle_{A^p} - \log \langle \sigma_i \rangle_{A^r} \right), \tag{14}$$

$$J_{ij}(t+1) \leftarrow J_{ij}(t) + \alpha_g \left(\log \langle \sigma_i \sigma_j \rangle_{A^p} - \log \langle \sigma_i \sigma_j \rangle_{A^r} \right), \tag{15}$$

$$A^r = \{H_i, J_{ij}\}\,|_{i=1,\cdots,N, j=1,\cdots,N}$$

The recovery is proceeded by adjusting the network parameters of neighbors ($A^t_{\mathbb{R}_m}$) of the treated node or edge m.

In general, a treatment is applied to a system when the recovery process is saturated for a sufficient time after each treatment. However, we also presented a simulation of treatment in the transient state before full saturation. We denote the transition state as the proportion of time relative to the time for full recovery. In this case, we used a grid search method to determine the optimal treatment time without waiting for full recovery, as well as the target region and its scale. In this situation, we denote the functional distance $\mathcal{D}^+$ as a function of treatment timing T.

$$A^{t*}_m = \arg\min \mathcal{D}^+ \left(A^g, A^{s'} \middle| A^{s'}_{m'}, T \right) \tag{16}$$

MATERIALS AND METHODS

A Test System: the Subcortical Limbic Brain

As a test system for optimal control, we reused the MEM for the subcortical human brain (Kang et al., 2017). Briefly, the system consists of 15 subcortical regions of interests (ROIs): the hippocampus (HIPP), amygdala (AMYG), caudate (CAU), putamen (PUT), pallidum (PAL), thalamus (THL), nucleus accumbens (NACC) of the left (L), and right (R) hemispheres, and the brainstem (BSTEM). The MEM parameters were estimated from the resting-state fMRI data of 470 participants in the HCP database (Van Essen et al., 2012). The estimated parameters are presented in **Figures 3A,B**. The subcortical-limbic system has highly symmetric interactions across hemispheres and appears to be modular. We sorted local minima (LMs) with their occupation time ratios and selected

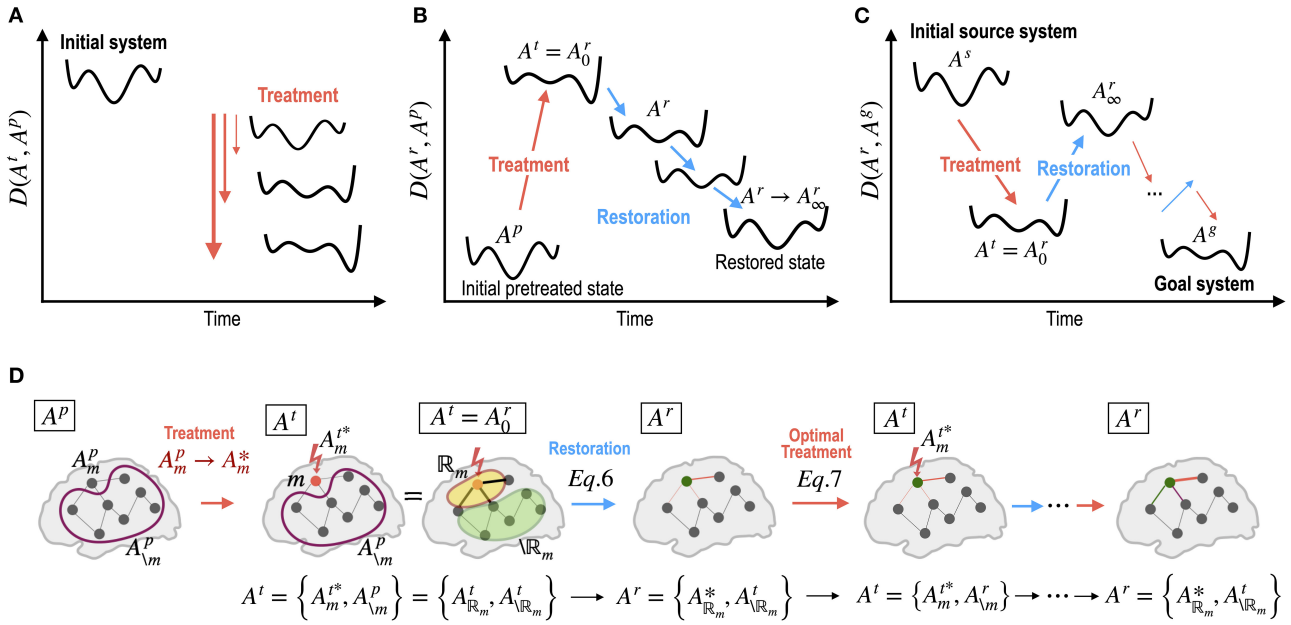

FIGURE 2 | Dynamics of energy landscapes (state dynamics) by treatment in the network parameters and by self-restoration processes. **(A)** Different treatments on the network alter the energy landscape of the system non-linearly. **(B)** The brain has a tendency to revert to its initial state before treatment via non-linear state transitions. **(C)** Optimal treatment to induce desired behaviors (dynamics) can be achieved by iterative treatment and self-restoration via an optimal adjustment in the network parameters. **(D)** Notations for iterative procedures in the optimal control of the brain are explained. $\mathbf{A^p}$, $\mathbf{A^t}$, and $\mathbf{A^r}$ indicate system (network) parameters for the pretreated state, treated state, and self-recovery state; $\mathbf{A^{t^*}_m}$: optimal strength of a treated node or edge (m); $\mathbf{A^p_{\backslash m}}$: the pre-treatment parameters of untreated node or edge (\m); $\mathbf{A^*_{\mathbb{R}_m}}$: optimal parameters within the recovered regions $\mathbb{R}_\mathbf{m}$; $\mathbf{A^t_{\backslash\mathbb{R}_m}}$: parameters of the unchanged regions in the recovery process. See the Method section for details.

the five top local minima that have the highest occupation time ratios. Among 18 LMs, the five major LMs, i.e., LM1(8000), LM2(24769), LM3(32768), LM4(1), and LM5(25286), occupied 83.5% of all possible states (**Figure 3E**). The brain activation patterns for the two major local minima are displayed in **Figure 3F**. In this study, we set this system as a goal of the control for the virtual abnormal systems. By controlling the regional activity parameter H_i and pairwise interaction parameter J_{ij} (node and edge of the MEM parameters) of a virtual abnormal system, the abnormal system is expected to be guided to have dynamics of this goal system.

Overview of Simulations

We conducted six simulation experiments to show the construct validity of the proposed framework. In simulation 1, we present an example to show the self-restoration process in the energy landscape of the brain network after treatment (or damage) in a region. In simulation 2, we present the effects of treatment on nodes and edges according to the number of neighbors to show the advantage of more neighboring edges in the restoration process. In simulation 3, we show the need for a self-restoration process in the control model by comparing treatment effects with and without considering the system's self-restoration. In simulation 4, we show the effects of repetitive treatment on each node of a source system to induce the desired dynamics. In simulation 5, we further control the timing of subsequent treatments before full recovery when treating a system. In simulation 6, we optimize the dissection of interhemispheric connectivity to simulate a corpus callosotomy for epilepsy surgery. All of these simulations were conducted to show the effects of the self-restoration process and how to treat the system to achieve the desired behaviors.

Simulation 1: The System's Self-Restoration Process After Treatment

To illustrate the self-restoration process, we presented a perturbation simulation of the right thalamus (R THL) by adding 0.5 to its H_i parameter. After this perturbation, the neighboring edges connected to the node were gradually reconfigured to generate state dynamics similar to those of the initial pre-treatment state (**Figure 4A**). For a treatment that induced a deviation of state dynamics from the pre-treatment state, the self-restoration procedure gradually moved the system toward the pre-treatment state, which induced a shorter functional distance $\mathcal{D}$ (partial KL-divergence) between the recovering and pre-treatment states (**Figure 4B**), and receded its parameters from those of the pre-treatment state (**Figures 4B,C**). Despite the increasing distance in the parameter space (RMSD, **Figure 4C**), the distance in the state dynamics from the pre-treatment state is reduced ($\mathcal{D}$, **Figure 4B**). This is an example of degeneracy, which refers to the phenomenon where similar behaviors can be formulated using different networks.

Treatment of the right thalamus changed the energy landscape significantly from that of the pre-treatment state (**Figure 4D**); the OCR of LM1 increased from 28.8 to 40.0, and the OCR of

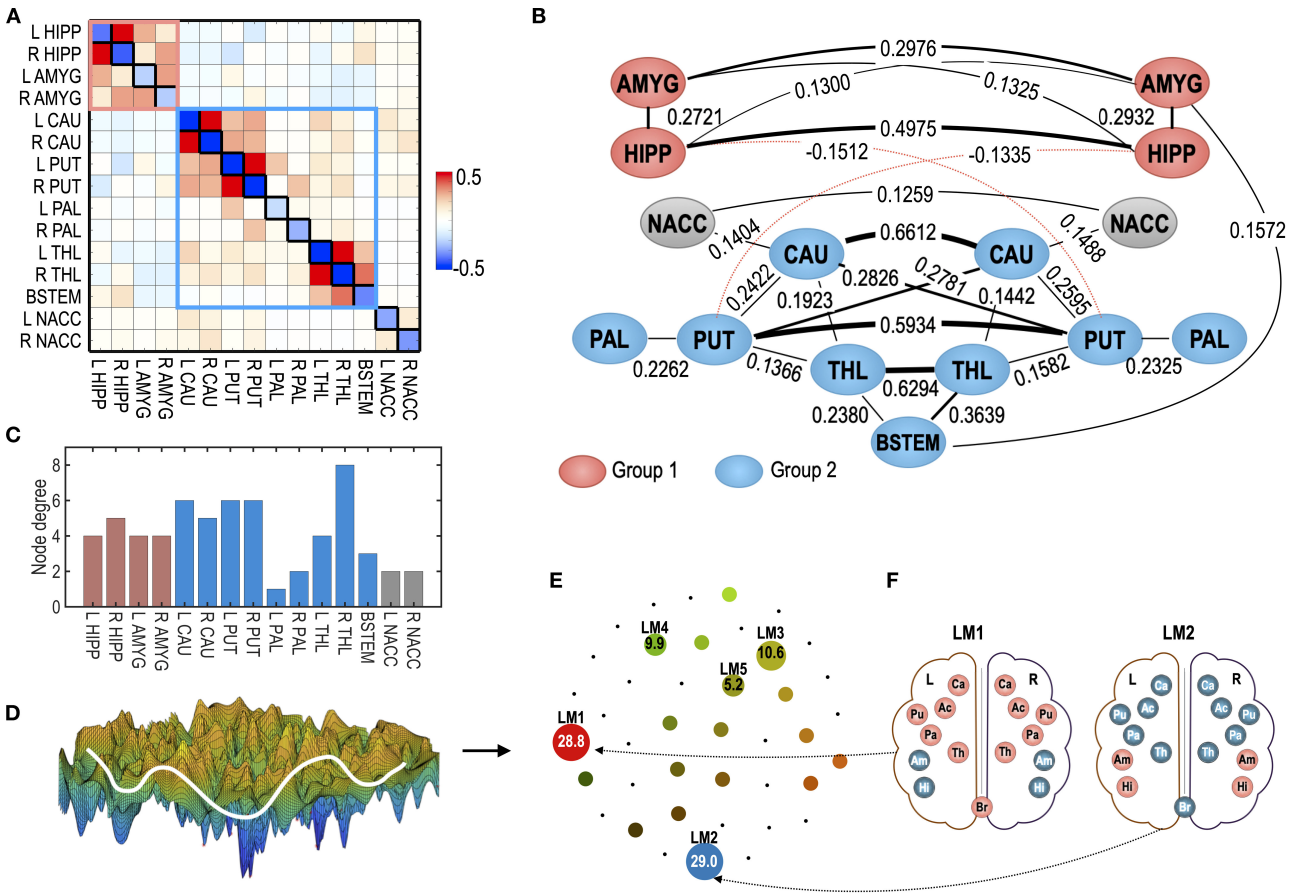

FIGURE 3 | The human subcortical-limbic system used in the study as a test system. **(A)** MEM parameters of the baseline sensitivity parameter (H_i) and pairwise interaction parameters (J_{ij}) are displayed as diagonal and off-diagonal elements in the matrix. **(B)** The system is displayed in the interaction diagram. The pairwise interaction is displayed for $|J_{ij}| > 0.12$. The thickness of the lines represents the strength of the pairwise interactions. The red dotted line represents negative pairwise interactions ($J_{ij} < 0$). Red and blue nodes represent ROIs that belong to modules 1 and 2. Modified from Kang et al. (2017). **(C)** Node degree of the subcortical-limbic system under a threshold $|J_{ij}| = 0.1$ is shown. **(D)** A schematic illustration of the energy landscape of the subcortical-limbic system is displayed. **(E)** The energy landscape of the human subcortical system is represented in terms of local minima (LM) displayed in circles. The sizes of circles for LMs reflect the occupation time ratios (OCR) of LMs. The OCRs of LMs are displayed in percentiles. **(F)** Activation patterns corresponding to two major local minima (LM), i.e., LM1 and LM2, are displayed. The subcortical-limbic system includes the hippocampus (HIPP, Hi), amygdala (AMYG, Am), caudate (CAU, Ca), putamen (PUT, Pu), pallidum (PAL, Pa), thalamus (THL, Th), nucleus accumbens (NACC, Ac), and brainstem (BSTEM, Br). Red and blue colors represent activated and inactivated regions, respectively.

LM2 decreased from 29.02 to 18.9. During self-restoration, this asymmetric energy landscape (between LM1 and LM2) gradually recovered (**Figure 4D**). In the final stage of the self-restoration, the OCRs of LM1 and LM2 were 37.5 and 31.9, similar to those of the initial system. Full recovery was not achieved in this system as it utilized only the limited resources of neighboring edges. **Figure 4E** shows the changes in the major two local minima (LM1 and LM4) along with the treatment and transient states in the recovery process. This suggests that the system recovers similar energy landscapes after recovery.

Simulation 2: Region-Specific Self-Restoration Capacity

To test the node- or edge-specific recovery capacity, we evaluated the self-restoration process after treating each node (**Figure 5**) and each edge (**Figure 6**) one by one. The degree of freedom was defined by the number of neighboring edges that participated in the self-restoration process. In the node's treatment, the neighboring edges were strongly (threshold ($|J_{ij}| \geq 0.1$) connected with the treated node (**Figure 1B**). In the treatment of an edge, neighboring nodes connected to the treated edge and strongly connected edges connected to these neighboring nodes were considered to participate in the self-restoration process (**Figure 1C**).

Figure 5 presents the node-specific restoration process in the real subcortical-limbic system. Nodes with diverse node degrees (**Figure 5B**) have different recovery capacities (**Figure 5A**) and recovery curves (**Figure 5C**). The finally recovered network parameters and energy landscapes differed from each other (**Figures 5C,D**). The recovery capacity depended highly on the number of neighbors (or node degrees); nodes with more

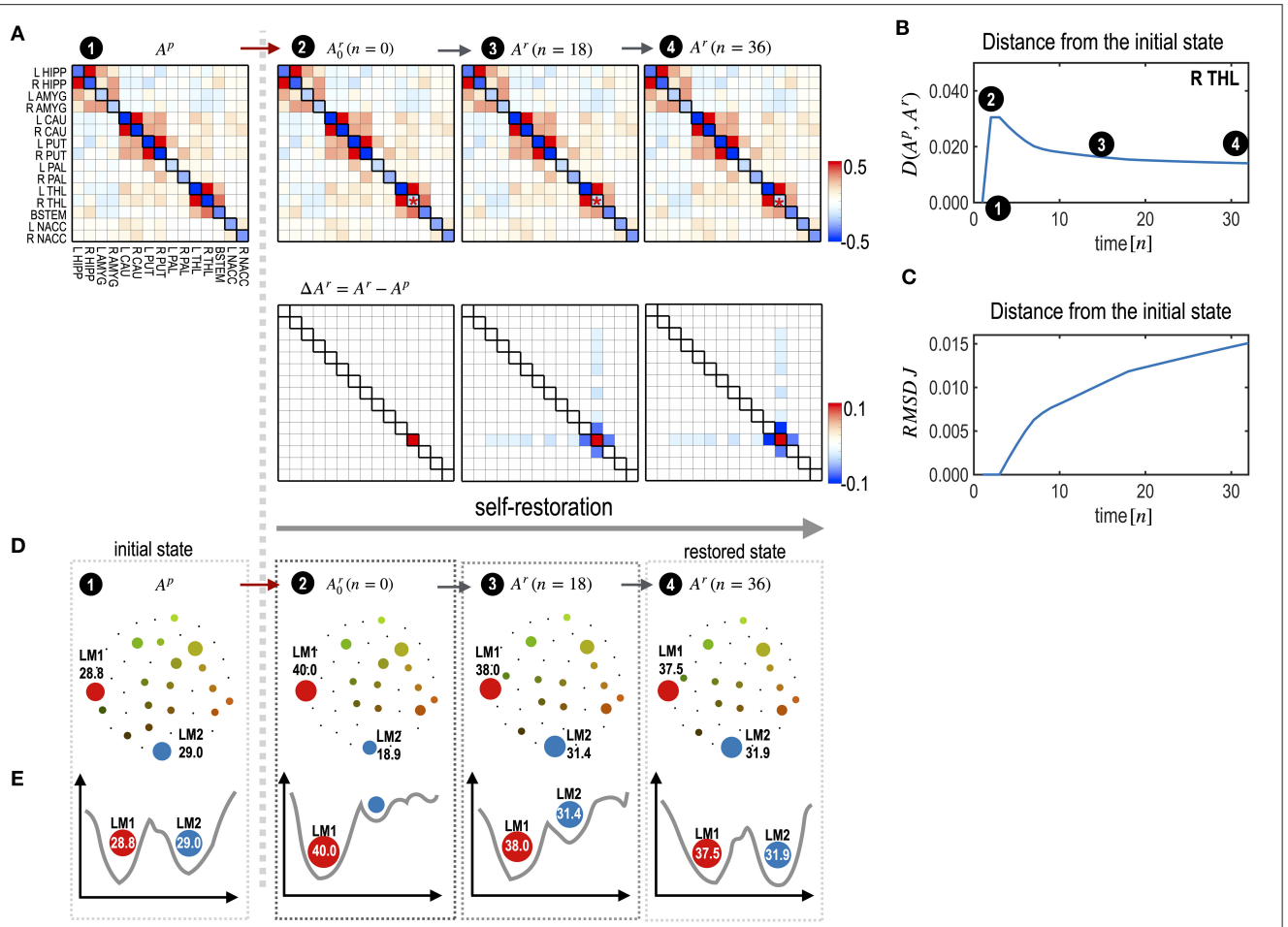

FIGURE 4 | Explanation of the self-restoration process. **(A)** MEM parameters (and differences from the initial system) at the initial, treated (the time point $n = 0$), half-restored ($n = 18$), and fully restored ($n = 36$) states are displayed. **(B,C)** A distance curve (partial KL-divergence) **(B)** and an RMSD curve of J (interaction parameters) **(C)** from the pre-treatment state are displayed. Despite the increasing distance in parameter space, the distance of state dynamics from the pre-treatment state is reduced in the functional space. **(D)** Schematic diagrams for energy landscapes of the pre-treatment state (A^p) and self-restored state (A^r) are displayed along the time course. **(E)** Alterations in the energy landscapes following treatment and during the restoration process are explained in terms of the two major local minima LM1 and LM2. After restoration, the occupation time ratios of LM1 and LM2 become closer to those of the pretreated state.

neighbors had a higher recovery capacity. In this case, the node degree acted as a degree of freedom of the system. This was also found in the treatment of edges shown in **Figure 6**. In this case, the number of neighboring edges of the system again explained the recovery capacity. The edges with a higher number of neighboring edges (size of neighbors $\mathbb{R}_m$ in **Figure 6B**) showed better restoration (**Figure 6C**). For example, greater self-restoration occurred after treatment in the edge between R CAU and L THL compared to the edge between L PAL and L NACC (**Figure 6D**). The latter utilizes adjustments of more edges than the former one in the self-restoration (**Figure 6E**).

Simulation 3. Effects of the Self-Restoration Process in Controlling the Brain System

We simulated treatments with and without considering the self-restorative properties of the system. In this study, we generated a virtual system by adding a Gaussian random noise $\sim N(0, 0.1)$ to the parameters of the human subcortical-limbic system presented in **Figure 3**. We considered the virtual system as a source system and the human subcortical system as a goal system (**Figure 7A**).

For each node, the best strength of treatment ($|\alpha|$) was identified using a grid search method among a set of α, 0.05, 0.10, 0.15, …, and 0.5. We selected the best treatment strength that minimizes the functional distance between the final restoration state and goal state based on Equations 8, 9. **Figure 7B** shows the expected treatment effects without considering the restoration process, while **Figure 7C** shows the actual treatment results for the self-restorative system. Discrepancies between the expected and treatment effects occurred when the self-restoration process was not considered. In contrast, when the recovery process was considered, we obtained increased treatment effects (**Figure 7D**), with functional distances from the goal system shorter than that of treatments without considering the system's recovery process (**Figure 7C**). Optimal nodes differed according to how the nodes

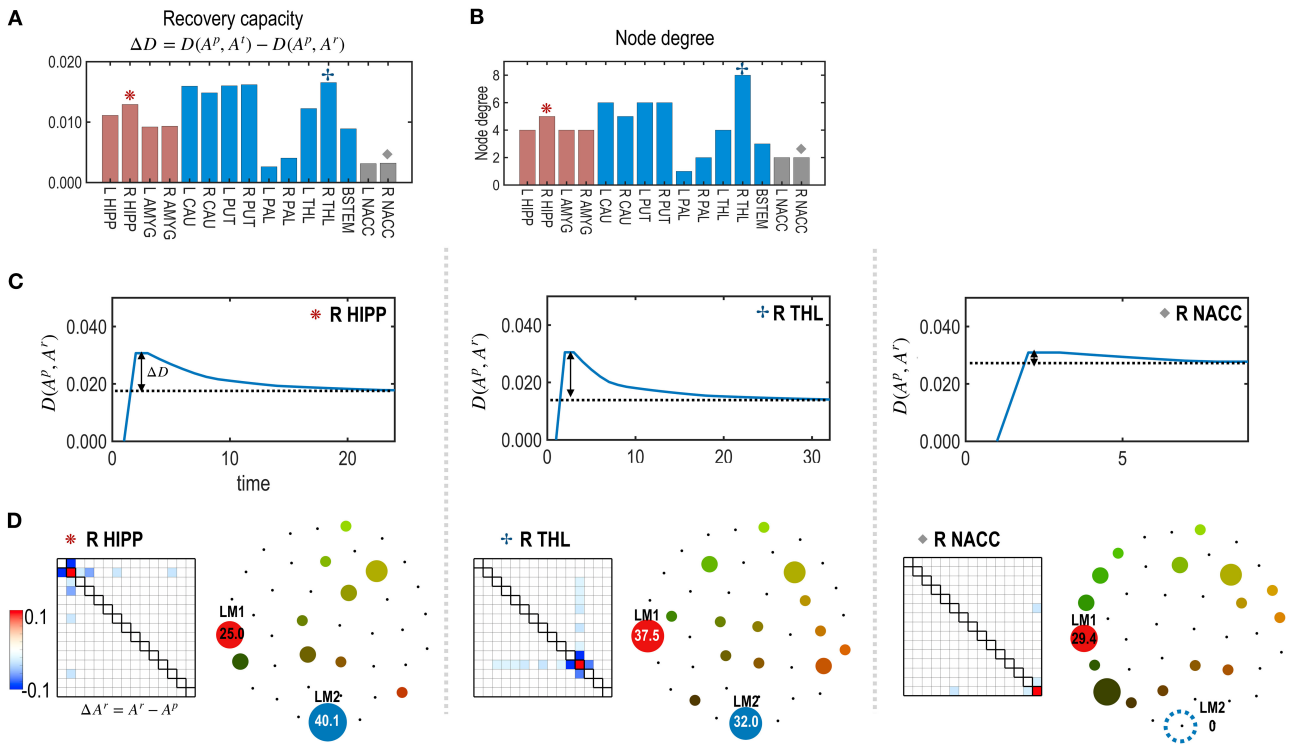

FIGURE 5 | The recovery capacity at the nodes of the subcortical-limbic system. **(A)** Recovery capacity **D** is shown for each node. **(B)** The node degree plays as degrees of freedom, i.e., numbers of adjustable parameters for the recovery. **(C)** For the three representative nodes (R HIPP, R THR, and R NACC), recovery curves are displayed in terms of the functional distance (partial KL divergence) between the restoring state and the pre-treatment state. **(D)** Differences between parameters of the restored and pre-treatment states, and schematic energy landscapes at the final restoration state are displayed. The restoration at the R THL was most successful in terms of recovering the energy landscape of the pretreated state.

were chosen with or without a self-restorative model. When we determine the optimal node and its treatment strength to treat without considering the recovery process, an optimal treatment target was chosen in the left putamen (L PUT) with a treatment strength of 0.45 (**Figure 7B**) but the optimal treatment did not effectively change the system to the desired goal after restoration (**Figures 7C,E**). When we consider the restoration's effects, the best treatment was selected on BSTEM with−0.5 (**Figure 7D**). For this treatment, functional distance from the desired state decreased right after treatment, followed by an increase during the restoration process (**Figure 7G**). The final treatment effects by considering the restorative process are better in this treatment than the treatment without considering restoration (**Figure 7E**).

Simulation 4. Repetitive Treatment at a Single Node

In the clinic, most treatments are repetitive, particularly concerning medications. We simulated repetitive treatments without changing the treated node. In the repetitive treatment simulation, the subsequent treatment was applied after the effects of the previous treatment had become saturated (i.e., reached an equilibrium state). **Figure 8** shows the results of the repetitive treatment at each node. Compared to the single treatment shown in **Figure 7D**, repetitive treatments with a sequence of different strengths (**Figures 8C,F**) made the system closer to the desired goal (**Figure 8B**). After repetitive treatment, the treated system's final energy landscape gets closer to that of the target goal system (**Figure 8H**).

In **Figure 8E**, the state change due to the second treatment suggests the non-linear nature of the treatment vs. the behavioral response (dynamics). The optimal treatment was not always chosen to minimize the distance to the goal system in the early stages of the time curves, as shown in the first treatment effect (**Figure 8E**). Instead, the optimal treatment made the system deviate from the desired goal system but eventually get closer to the desired system than other treatments that are initially effective but finally ineffective (**Figure 8E**).

Simulation 5. Repetitive Treatment at a Single Node With Flexible Timing

Most previous studies did not consider the timing of the treatment under the dynamically responding brain. When a treatment is applied, the brain gradually recovers and transitions to an equilibrium state. Considering the restoration process, one may apply the subsequent treatment at transient states without waiting for the equilibrium state. We simulated

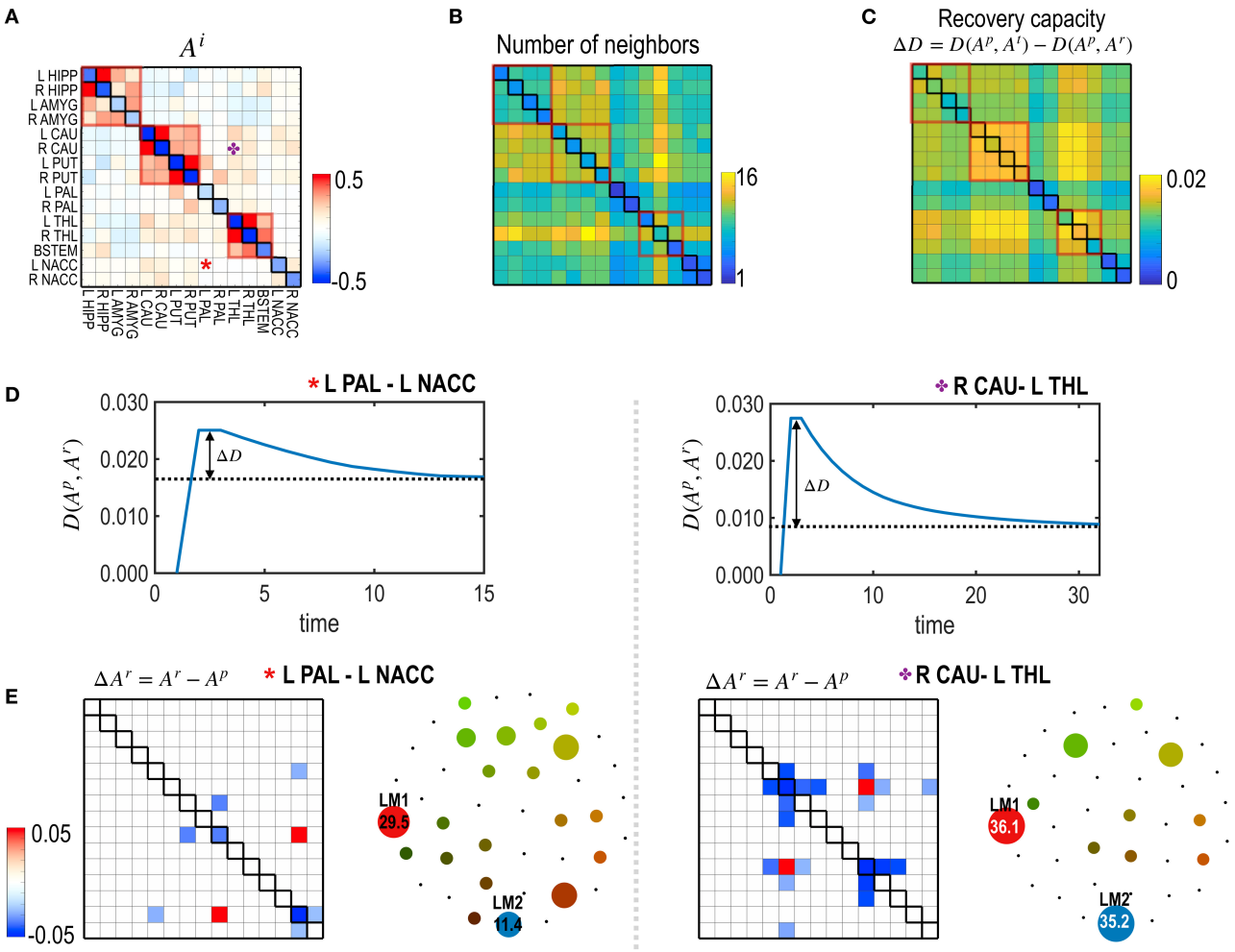

FIGURE 6 | The recovery capacity at the edges of the subcortical-limbic system. **(A,B)** The model parameters **(A)** and the number of neighbors $\mathbb{R}_m$ as degrees of freedom of the subcortical-limbic system **(B)** are presented. **(C)** Recovery capacities ΔD are shown for all the edges. Larger self-restoration occurred in the edges with high degrees of freedom (e.g., orange asterisk; edge between R CAU-L THL) than edges with lower degrees of freedom (e.g., red asterisk; edge between L PAL-L NACC). Two representative examples of treatments on the two edges are presented to show their differences in the recovery processes. **(D)** The distance curves (the partial KL-divergence) from the pretreated state are presented for the two representative edges; L PAL–L NACC (red asterisk) and R CAU–L THL (orange asterisk). **(E)** Differences in the MEM parameters of the recovered states from the pre-treatment state and the schematic energy landscapes of the final states for the recovered two edges are illustrated.

optimal repetitive treatment by optimizing the timing of subsequent treatments. In this simulation, we used the same simulation setting as simulation 3, except for the flexible timing of the treatment. Compared to simulation 4, we explored the best strength of the treatment and the best timing of subsequent treatments for each node. As shown in **Figure 9**, the treatment with the best timing increased the treatment effect compared to treatment after full recovery for each treatment (**Figure 8**).

Simulation 6. Optimal Removal of Edges

We simulated dissections of interhemispheric connections to imitate a corpus callosotomy for epilepsy surgery. We generated an abnormal brain that has stronger connections between the left and right hemispheres (**Figure 10A**). This was performed by increasing positive inter-hemispheric connectivity and decreasing the negative interhemispheric connectivity of the subcortical-limbic system by adding a Gaussian noise $\sim N(0, 0.1)$ according to the polarity of the initial connectivity. We tested the optimal treatment strategy for different numbers and targets of interhemispheric edges (one, two, and three) to be dissected (**Figures 10B,C**). For each number of edges (49 single interhemispheric edges, i.e., left 7 × right 7, 2,352 combinations for two edges, and 110,544 combinations for three edges), we evaluated the best edges to remove. For the source brain, dissection of LAMYG-RCAU (one edge); LAMYG-RCAU and LAMYG- RPAL (two edges); LAMYG-RCAU, LAMYG- RPAL, and LHIPP- RCAU (three edges) were the best combinations to decrease the functional distance from the goal system (**Figure 10D**).

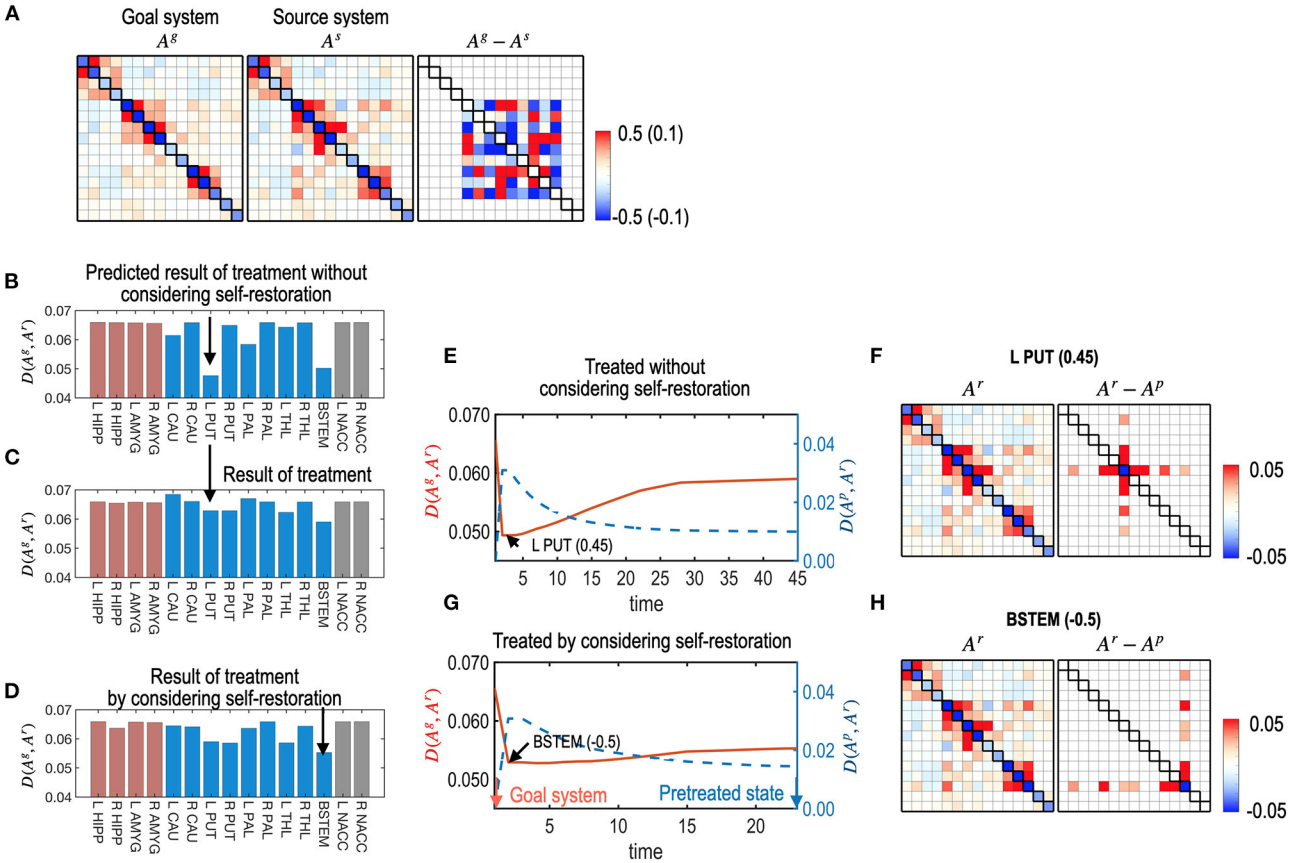

FIGURE 7 | Treatments with and without considerations of the self-restoration. **(A)** Model parameters of a virtual (source) system and the subcortical-limbic system as a goal system, and the difference between the goal and the source systems are displayed. The virtual system is generated by adding a noise ($\mathbf{A^g} - \mathbf{A^s}$) to the subcortical-limbic system. **(B)** The distance between the goal state and the predicted final state, $\mathbf{D}(\mathbf{A^g}, \mathbf{A^r})$, after treatment at each node without considering the self-restoration, was minimal for the treatment on the L PUT. **(C)** The final results of treatments are displayed in **(B)**, which show longer distances from the goal state than predicted due to the restoration process. For example, the treatment on node L PUT with strength 0.45 results in distance 0.074, which was expected 0.055 in the prediction without considering the self-restoration process. **(D)** The final distance for each node treated by considering self-restoration is displayed. In this case, BSTEM was selected as an optimal treatment node with −0.5 strength. **(E–H)** The treatment and restoration curves in terms of the distance from the goal system (red) and the pretreated state (blue) for the treatment on L PUT with strength of 0.45 **(E)** and for the treatment on BSTEM with strength of −0.5 **(G)** are displayed. MEM parameters (and differences from the pretreated state) at the final state are displayed for the treatment on L PUT **(F)** and for the treatment on BSTEM **(H)**.

DISCUSSION

Although brain control has garnered increasing interest, brain control research has mainly been conducted based on theoretical and computational models as the practical control of the brain has many challenges due to the complexity of the brain and ethical issues. Several computational models to control the brain network have been proposed to characterize the graph-theoretic properties of the system (Tang et al., 2017; Lee et al., 2019; Stiso et al., 2019; Karrer et al., 2020) or a purpose of predicting outcomes after treatment (Falcon et al., 2016; Jirsa et al., 2017; Proix et al., 2017; An et al., 2019; Olmi et al., 2019). These previous studies have assumed the brain as a dynamic system, immediately responding to the incoming treatment. However, the system's self-restoration process after the cessation of the treatment has not been fully considered, without which the brain control may not be optimal. Compared to the brain circuit's various dynamic state equations, the formulation of the self-restoration process has been rarely researched.

To account for the effect of the restoration process on brain control, we propose a formulation of the brain's recovery process that drives the system to perform the function before treatment. The driving force of this self-restoration process is based on the free-energy principle (Friston et al., 2006; Friston, 2010) over a non-linear complex system, with degeneracy in terms of generating the same behaviors from diverse network configurations. According to the free-energy principle (Friston et al., 2006; Friston, 2010), the network is configured to respond or predict the environment's statistical demands, making the system energy-efficient. As long as the environment's statistics do not change, the treated or partially lesioned system may well-adjust its remaining subnetwork (neighbors of the treated node) to satisfy those demands. Since the altered (treated) node cannot participate in the organized work of the

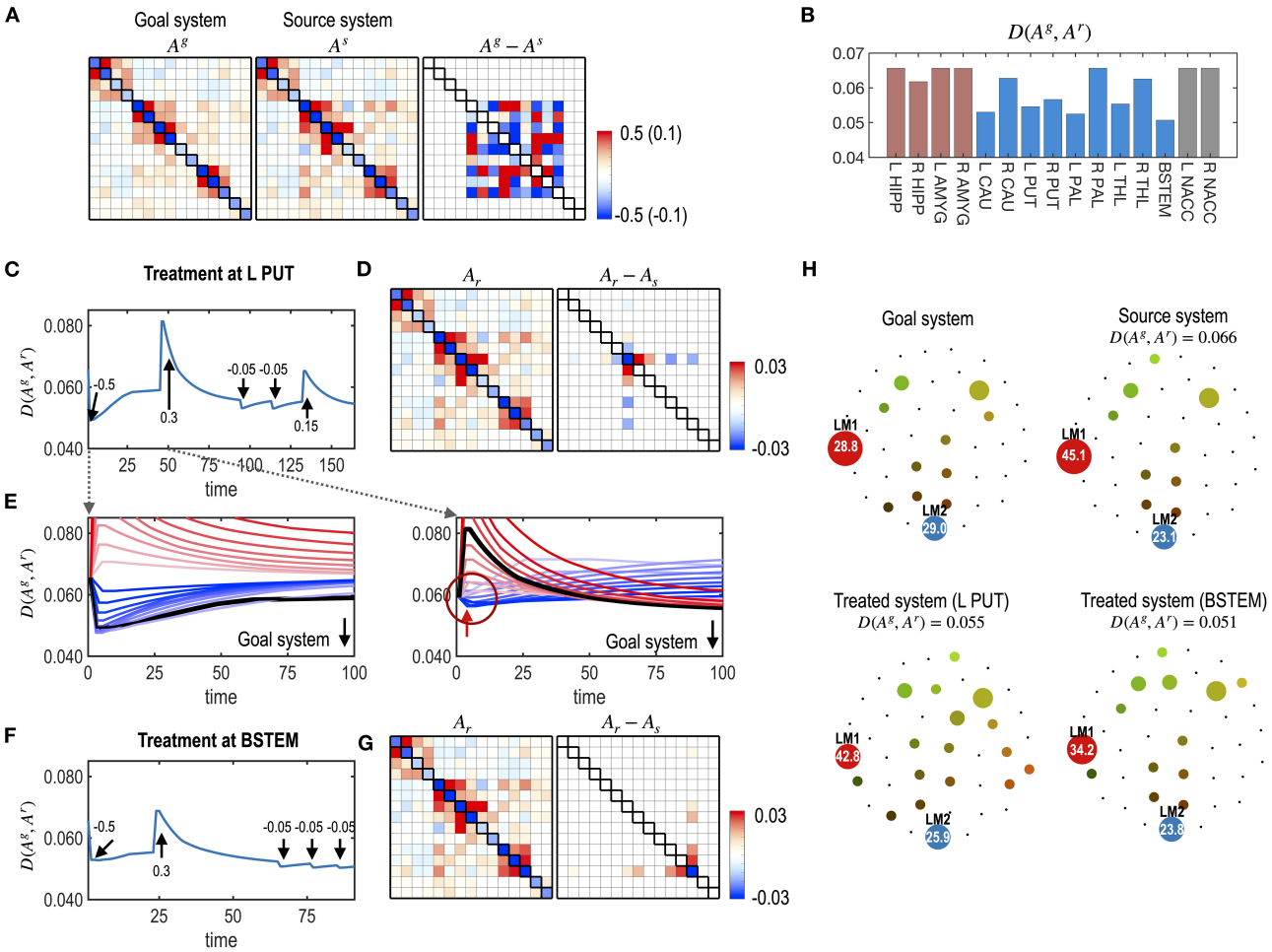

FIGURE 8 | Multiple treatments for the nodes of a system. **(A)** MEM parameters of the virtual source and goal systems, and their differences are displayed. **(B)** The distances (partial KL divergence) between the final states and the desired state are displayed for each treatment. The optimal treatment was chosen at the BSTEM. Results of the treatment at L PUT **(C–E)** and BSTEM **(F,G)** are shown. **(C)** A treatment curve (distance between the transient state and the desired state) with a treatment strength at each treatment (arrows) is displayed. **(D)** MEM model of the final state and its difference from the pre-treated state are shown. **(E)** Restoration curves with potential strengths of the treatment are displayed for the first and second treatments. In the first treatment, the treatment strength that induces the black line was chosen as the optimal strength since it finally becomes closest to the desired state after saturation. After a treatment, it gets close to the desired state and then recedes slightly from it. In the second treatment, some treatment strengths make the system closer to the desired state at the early stage of the treatment (blue lines in the circle). However, those curves eventually diverge from the goal system. The optimal treatment strength induced a curve colored in black, which initially deviates from the goal system but eventually comes closer to the desired state than any other treatment strengths. For clarity, we scaled the restoration time for each treatment. **(F,G)** A treatment curve at the BSTEM **(F)** and its final MEM parameter and its differences from the pre-treated state **(G)** are shown. **(H)** Energy landscapes of the goal, the virtual source, and final systems after treatment at L PUT and BSTEM are displayed.

subnetwork at the same performance level as a pre-treatment state, the system tries to compensate for the role of the altered node by reorganizing interactions with its neighbors. This restoration process in the brain can be called a type of optimization process in that the system tries to adjust itself and gradually approximates the desired functionality of the pre-treatment state by interaction with the environment under biological constraints.

The other central concept of the current study is the redundant nature of the non-linear brain (Glassman, 1987; Edelman and Gally, 2001). A complex system has degeneracy, i.e., the same or similar functions (behavior) can be achieved using different configurations of networks (or connectivity). Since it is complicated to restore all connectivity after damage, optimal control utilizes non-linearity between networks and behaviors by reconfiguring networks among neighbors within limited ranges to approximate the goal system dynamics. In this non-linear relationship, the closeness in the system parameters (e.g., connectivity) does not necessarily indicate closeness in behaviors. Instead of matching network connectivity, the current framework fits behaviors (i.e., microstate dynamics) by modulating a smaller number of network parameters. This is possible as the brain is a complex non-linear system from which non-linear microstate dynamics emerge.

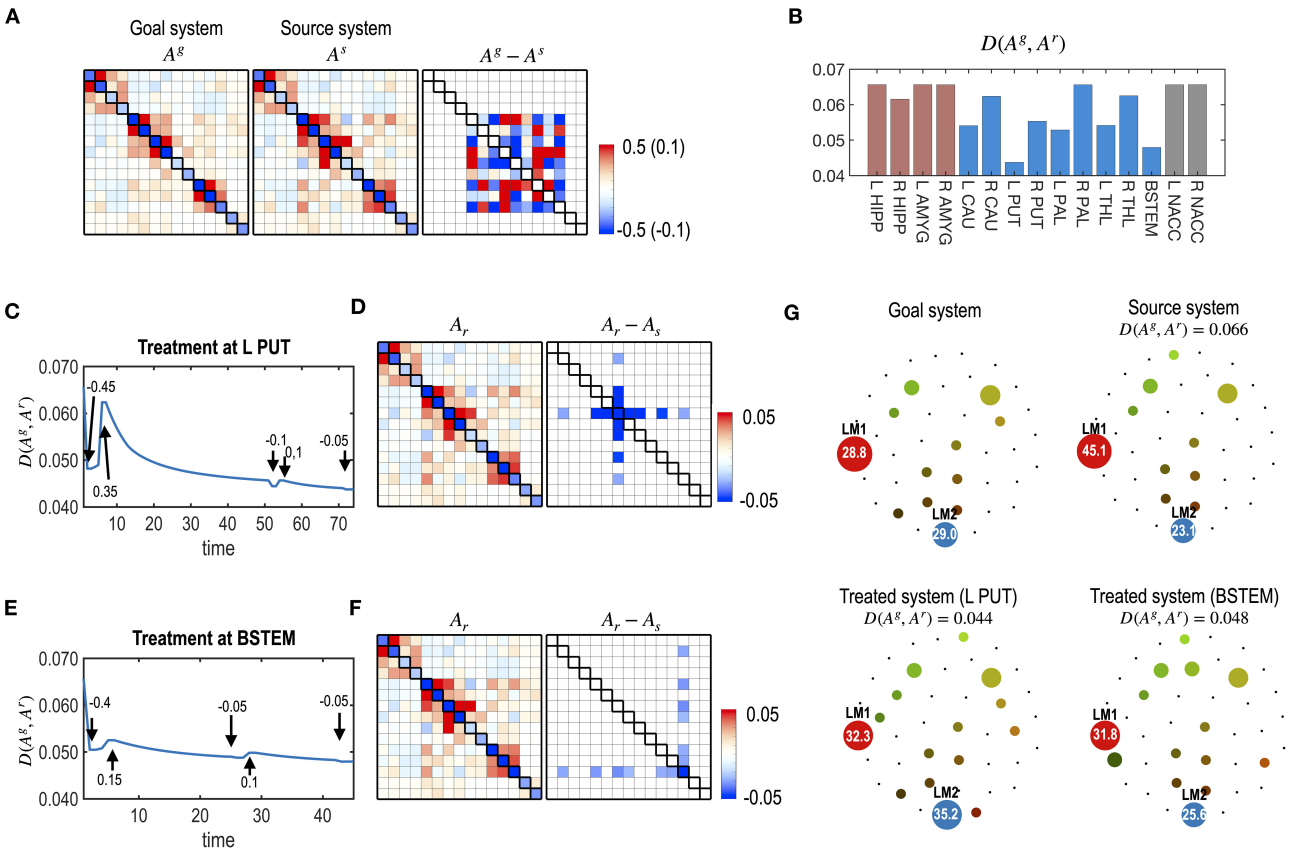

FIGURE 9 | Repetitive treatments with flexible treatment timing. **(A)** The same model in **Figure 8** is used for the current simulation. **(B)** The distance between the goal and final states, **D** ($\mathbf{A}^g$,$\mathbf{A}^r$), from multiple treatments with flexible timing of treatment is evaluated for each node. The best node for the multiple treatments was the L PUT. **(C–F)** Results of the treatments at L PUT **(C,D)** and at BSTEM **(E,F)** are displayed. **(C,E)** Treatment curves (the distance between the transient states and the goal system) are displayed for L PUT and BSTEM. The arrows indicate the timing of each treatment, and the values are the strengths of the treatment. **(D,F)** MEM parameters of the final system and its difference from the initial system are shown. **(G)** Energy landscapes of the goal system, the virtual source system, and final systems after treatment at the L PUT and BSTEM are displayed.

The non-linear microstate dynamics of a brain are represented in the microstate's energy landscape, the microstate of which is often defined by a distributed activity pattern over the temporal scale of a second. Energy landscape analysis has been applied to explore dynamics in large-scale functional brain networks, such as the default mode and pre-frontal networks, on resting-state fMRI (Watanabe et al., 2013, 2014a; Kang et al., 2017, 2019) and in sleep (Watanabe et al., 2014b). In our previous study (Kang et al., 2017), the energy landscape analysis revealed that the subcortical brain at rest exhibits the maximal number of stable states and small sets of stable states account for most of the occupation time. Furthermore, a graph theory analysis of the energy landscape revealed a hub-like state transition organization embedded in the resting-state human brain (Kang et al., 2019). The energy landscape of brain states is governed by a set of network parameters in the pairwise MEM, upon which treatment is imposed.

The brain control extends the energy landscape concept at the temporal scale of a second (microstates) to the energy landscape over a more extended period. Over a longer period of years, a brain can be considered in a network state of the macroscopic energy landscape. For example, the brain develops from one network state (a set of network parameters) to another (another set of network parameters). A network state in the macroscopic energy landscape is defined by a network parameter, which differs from the definition of a microstate in the microscopic energy landscape by a distributed activity pattern. In this respect, the brain control problem is to choose an optimal way to guide a brain network to a desired network along the macroscopic energy landscape.

We assumed that the restoration process is conducted by the cooperative activity among neural populations within the brain network, which tries to generate similar functionality established before treatment. In this process, the brain network's modularity, an essential property to protect against damage to a complex brain (Park and Friston, 2013; Sporns and Betzel, 2016), would be crucial for the recovery. By rearranging resources within a module (e.g., altering connectivity within a biological range), the modularity actuates the system's reorganization to construct a similar behavior. There is plentiful evidence of modular-based reorganization in brain diseases (Balenzuela

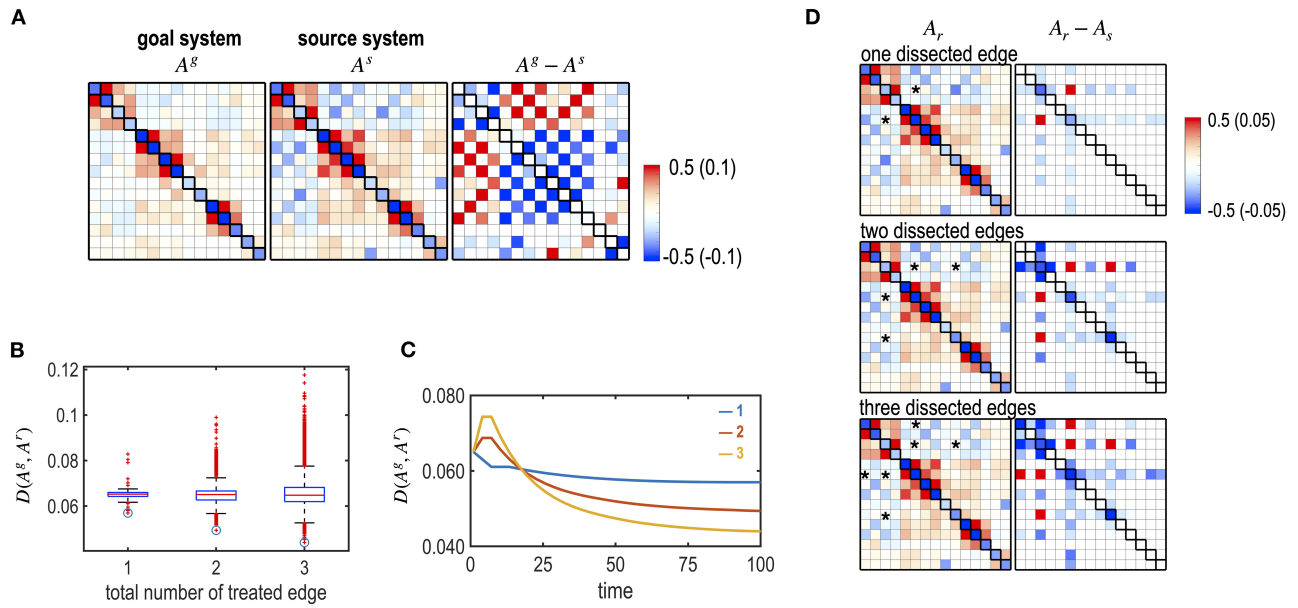

FIGURE 10 | Optimal dissection of edges, simulating a corpus callosotomy for epilepsy surgery. **(A)** MEM parameters of the virtual (abnormal source) system and the goal system (subcortical-limbic system), and their differences are displayed. The virtual source system was generated by increasing the strengths of the connectivity between the left and right hemispheres. **(B)** The box graph shows the final distance between the goal and the treated system, $\mathbf{D}(\mathbf{A^g}, \mathbf{A^r})$ for all the combinations of dissected edges by dissecting one, two, and three inter-hemispheric edges. Optimally treated results for each number of edges are denoted with circles. **(C)** Treatment curves (the distance between the transient state and the goal system) are displayed for the optimally chosen edges among the one (blue line), two (red line), and three (yellow line) edge combinations of dissection. **(D)** MEM parameters of the final systems after optimal treatments with different numbers of dissected edges are displayed with their differences from the initial source system.

et al., 2010; Chen et al., 2013; Siegel et al., 2018). In this study, nodes with more (stronger) connections (functional neighbors) play a more efficient role in restoration than nodes with fewer connections. Since the current test system has a relatively small network size (nodes = 15), we did not restrict the neighbors within a module but functionally close regions (**Figure 6**). Even though not strictly the same as the modularity concept in systems science, the functional neighbors work as a functional module in terms of cooperation within the module. Consideration of neighbors restricted within a functional module and within an anatomical limit of a larger network would be more realistic in modeling the recovery process.

Clinical treatment is generally exerted on the brain multiple times. After treatment, e.g., antipsychotic medication, clinicians wait to stabilize the brain to avoid transient states. However, one may consider applying subsequent treatment before stabilization. Until fully stabilized, the system has multiple transient states for network parameters. Some transient states may be more efficient in achieving a desired goal than the stabilized state. However, the transient state may be unstable, and finding an optimal strategy may be unpredictable. In the current study, we showed a possibility to optimize the right timing without waiting for complete stabilization when we have a model for self-restorative process.

The current framework as computational modeling takes advantage of prediction capacity by simulation. It is theoretically possible that some treatment parameters may lead the treated system unstable, generating abnormal functionality. The self-restoration process may also cause the malformation of the function. By evaluating treatment outcomes for all possible ranges of parameters, we may check unstable points before deciding the treatment. The model-based prediction could also be used in evaluating the treatment effect due to noise in the restoration process as a type of Monte-Carlo simulation (See the **Supplementary Material**). As the noise effects differ across brain regions, one may choose a reliable target that is less sensitive to noise in the restoration process. The treatment could benefit from evaluating the treatment outcomes with noise in any parameters or any updating rules besides the restoration process. This Monte-Carlo simulation may complement the limitation of the current deterministic approach. We used a simple deterministic model and its solver for the control problem to explain the basic framework of brain control and show the current framework's construct validity. More sophisticated models based on more advanced control theory methods, such as a stochastic model proposed by Todorov (2009), could be further researched.

In this study, we showed the construct validity of the proposed framework using various simulations to consider the clinical environment. A simulation suggests that the optimal brain control should include the system's self-restoration process, without which a (so called optimal) treatment is not optimal. Using simulation, we also proposed how to control the

self-restoration process by choosing the optimal target region and treatment strength. We then presented simulations for optimizing repetitive treatments and the optimal timing of treatment. We found that some treatment choices led to a degraded performance at an early stage but eventually showed a better treatment effect (**Figure 8**). This is a typical example of the non-linear property of the self-restoration system that should be considered in optimal control. All of these simulations suggest the plausibility and rationale of the proposed brain control framework.

The current study is theoretical, and we acknowledge all possible limitations of the theoretical framework. The current brain control framework will be more practical when we know more about the system's reorganization mechanisms. Empirical experiments and validation are most demanding. Determining the means of achieving the desired treatment level at the right target for each treatment is one of the fundamental challenges. The details of the restoration process require extensive research and experiments. There exist many challenges before brain control can be applied to actual experiments. However, the current conceptual framework with the self-restorative process in the treatment is highly needed in clinical practices, which calls for personalized treatments based on individualized self-restoration systems and basic neuroscience to understand how the brain works.

In summary, we propose an optimal brain control framework by introducing self-restoration processes in the brain after treatment. Simulation results showing the responses and movement of a source system toward the desired system in diverse testing sets suggest the framework's plausibility in optimal brain control within a restricted treatment environment. Although further research with experimental data should be conducted, we believe the proposed computational framework would help attain optimal brain control of the dynamic self-restorative brain after treatment.

DATA AVAILABILITY STATEMENT

The MATLAB codes used in the current study are available by request or online from a public drive, http://bit.ly/monet-optimal-control-MEM.

AUTHOR CONTRIBUTIONS

H-JP developed the idea and JK performed experiments. Both wrote the manuscript.

FUNDING

This research was supported by Brain Research Program and the Brain Pool Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (NRF-2017M3C7A1049051 and NRF-2017H1D3A1A01053094).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncom.2021.590019/full#supplementary-material

REFERENCES

Abbott, A. (2010). Schizophrenia: the drug deadlock. *Nature* 468, 158–159. doi: 10.1038/468158a

An, S., Bartolomei, F., Guye, M., and Jirsa, V. (2019). Optimization of surgical intervention outside the epileptogenic zone in the virtual epileptic patient (VEP). *PLoS Comput. Biol.* 15:e1007051. doi: 10.1371/journal.pcbi.1007051

Balenzuela, P., Chernomoretz, A., Fraiman, D., Cifre, I., Sitges, C., Montoya, P., et al. (2010). Modular organization of brain resting state networks in chronic back pain patients. *Front. Neuroinform* 4:116. doi: 10.3389/fninf.2010.00116

Bennabi, D., Charpeaud, T., Yrondi, A., Genty, J. B., Destouches, S., Lancrenon, S., et al. (2019). Clinical guidelines for the management of treatment-resistant depression: french recommendations from experts, the french association for biological psychiatry and neuropsychopharmacology and the fondation fondamental. *BMC Psychiatry* 19:262. doi: 10.1186/s12888-019-2237-x

Biswal, B., Yetkin, F. Z., Haughton, V. M., and Hyde, J. S. (1995). Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* 34, 537–541. doi: 10.1002/mrm.1910340409

Breakspear, M. (2017). Dynamic models of large-scale brain activity. *Nat. Neurosci.* 20, 340–352. doi: 10.1038/nn.4497

Cabral, J., Kringelbach, M. L., and Deco, G. (2014). Exploring the network dynamics underlying brain activity during rest. *Prog. Neurobiol.* 114, 102–131. doi: 10.1016/j.pneurobio.2013.12.005

Chen, G., Zhang, H. Y., Xie, C., Chen, G., Zhang, Z. J., Teng, G. J., et al. (2013). Modular reorganization of brain resting state networks and its independent validation in Alzheimer's disease patients. *Front. Hum. Neurosci.* 7:456. doi: 10.3389/fnhum.2013.00456

Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., and Petersen, S. E. (2014). Intrinsic and task-evoked network architectures of the human brain. *Neuron* 83, 238–251. doi: 10.1016/j.neuron.2014.05.014

Cole, M. W., Ito, T., Bassett, D. S., and Schultz, D. H. (2016). Activity flow over resting-state networks shapes cognitive task activations. *Nat. Neurosci.* 19, 1718–1726. doi: 10.1038/nn.4406

Cornblath, E. J., Tang, E., Baum, G. L., Moore, T. M., Adebimpe, A., Roalf, D. R., et al. (2019). Sex differences in network controllability as a predictor of executive function in youth. *Neuroimage* 188, 122–134. doi: 10.1016/j.neuroimage.2018.11.048

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., 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

Edelman, G. M., and Gally, J. A. (2001). Degeneracy and complexity in biological systems. *Proc. Natl. Acad. Sci. U.S.A.* 98, 13763–13768. doi: 10.1073/pnas.231499798

Ezaki, T., Sakaki, M., Watanabe, T., and Masuda, N. (2018). Age-related changes in the ease of dynamical transitions in human brain activity. *Hum. Brain Mapp.* 39, 2673–2688. doi: 10.1002/hbm.24033

Falcon, M. I., Riley, J. D., Jirsa, V., Mcintosh, A. R., Elinor Chen, E., and Solodkin, A. (2016). Functional mechanisms of recovery after chronic stroke: modeling with the virtual brain. *eneuro* 3, 1–14. doi: 10.1523/ENEURO.0158-15.2016

Fox, M. D., and Raichle, M. E. (2007). Spontaneous fluctuations in brain activity observed with functional magnetic resonance

imaging. *Nat. Rev. Neurosci.* 8, 700–711. doi: 10.1038/nrn2201

Freyer, F., Roberts, J. A., Becker, R., Robinson, P. A., Ritter, P., and Breakspear, M. (2011). Biophysical mechanisms of multistability in resting-state cortical rhythms. *J. Neurosci.* 31, 6353–6361. doi: 10.1523/JNEUROSCI.6693-10.2011

Freyer, F., Roberts, J. A., Ritter, P., and Breakspear, M. (2012). A canonical model of multistability and scale-invariance in biological systems. *PLoS Comput. Biol.* 8:e1002634. doi: 10.1371/journal.pcbi.1002634

Friston, K. (2008). Hierarchical models in the brain. *PLoS Comput. Biol.* 4:e1000211. doi: 10.1371/journal.pcbi.1000211

Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi: 10.1038/nrn2787

Friston, K., Kilner, J., and Harrison, L. (2006). A free energy principle for the brain. *J. Physiol. Paris* 100, 70–87. doi: 10.1016/j.jphysparis.2006.10.001

Friston, K. J., and Price, C. J. (2003). Degeneracy and redundancy in cognitive anatomy. *Trends Cogn. Sci.* 7, 151–152. doi: 10.1016/S1364-6613(03)00054-8

Glassman, R. B. (1987). An hypothesis about redundancy and reliability in the brains of higher species: analogies with genes, internal organs, and engineering systems. *Neurosci. Biobehav. Rev.* 11, 275–285. doi: 10.1016/S0149-7634(87)80014-3

Goellner, E., Bianchin, M. M., Burneo, J. G., Parrent, A. G., and Steven, D. A. (2013). Timing of early and late seizure recurrence after temporal lobe epilepsy surgery. *Epilepsia* 54, 1933–1941. doi: 10.1111/epi.12389

Gu, S., Betzel, R. F., Mattar, M. G., Cieslak, M., Delio, P. R., Grafton, S. T., et al. (2017). Optimal trajectories of brain state transitions. *Neuroimage* 148, 305–317. doi: 10.1016/j.neuroimage.2017.01.003

Gu, S., Cieslak, M., Baird, B., Muldoon, S. F., Grafton, S. T., Pasqualetti, F., et al. (2018). The energy landscape of neurophysiological activity implicit in brain network structure. *Sci. Rep.* 8:2507. doi: 10.1038/s41598-018-20123-8

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. doi: 10.1016/j.neuroimage.2016.04.049

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

Kang, J., Eo, J., Lee, D. M., and Park, H. J. (2021). A computational framework for optimal control of a self-adjustive neural system with activity-dependent and homeostatic plasticity. *Neuroimage* 230:117805. doi: 10.1016/j.neuroimage.2021.117805

Kang, J., Pae, C., and Park, H. J. (2017). Energy landscape analysis of the subcortical brain network unravels system properties beneath resting state dynamics. *Neuroimage* 149, 153–164. doi: 10.1016/j.neuroimage.2017.01.075

Kang, J., Pae, C., and Park, H. J. (2019). Graph-theoretical analysis for energy landscape reveals the organization of state transitions in the resting-state human cerebral cortex. *PLoS ONE* 14:e0222161. doi: 10.1371/journal.pone.0222161

Kar, S. K. (2019). Predictors of response to repetitive transcranial magnetic stimulation in depression: a review of recent updates. *Clin. Psychopharmacol. Neurosci.* 17, 25–33. doi: 10.9758/cpn.2019.17.1.25

Karrer, T. M., Kim, J. Z., Stiso, J., Kahn, A. E., Pasqualetti, F., Habel, U., et al. (2020). A practical guide to methodological considerations in the controllability of structural brain networks. *J. Neural Eng.* 17:026031. doi: 10.1088/1741-2552/ab6e8b

Kelso, J. A. (2012). Multistability and metastability: understanding dynamic coordination in the brain. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 906–918. doi: 10.1098/rstb.2011.0351

King, A. (2016). Neurobiology: rise of resilience. *Nature* 531, S18–S19. doi: 10.1038/531S18a

Krienen, F. M., Yeo, B. T. T., and Buckner, R. L. (2014). Reconfigurable task-dependent functional coupling modes cluster around a core functional architecture. *Philos. Trans. R. Soc. B Biol. Sci.* 369. doi: 10.1098/rstb.2013.0526

Lee, I. C., Li, S. Y., and Chen, Y. J. (2017). Seizure recurrence in children after stopping antiepileptic medication: 5-year follow-up. *Pediatr. Neonatol.* 58, 338–343. doi: 10.1016/j.pedneo.2016.08.005

Lee, W. H., Rodrigue, A., Glahn, D. C., Bassett, D. S., and Frangou, S. (2019). Heritability and cognitive relevance of structural brain controllability. *Cereb. Cortex.* 30, 3044–3054. doi: 10.1093/cercor/bhz293

Liu, Y. Y., Slotine, J. J., and Barabasi, A. L. (2011). Controllability of complex networks. *Nature* 473, 167–173. doi: 10.1038/nature10011

Malone, L. A., and Felling, R. J. (2020). Pediatric stroke: unique implications of the immature brain on injury and recovery. *Pediatr. Neurol.* 102, 3–9. doi: 10.1016/j.pediatrneurol.2019.06.016

Marder, E., and Goaillard, J. M. (2006). Variability, compensation and homeostasis in neuron and network function. *Nat. Rev. Neurosci.* 7, 563–574. doi: 10.1038/nrn1949

Marder, E., Gutierrez, G. J., and Nusbaum, M. P. (2017). Complicating connectomes: electrical coupling creates parallel pathways and degenerate circuit mechanisms. *Dev. Neurobiol.* 77, 597–609. doi: 10.1002/dneu.22410

Mattson, M. P. (2008). Glutamate and neurotrophic factors in neuronal plasticity and disease. *Ann. N. Y. Acad. Sci.* 1144, 97–112. doi: 10.1196/annals.1418.005

Mikellidou, K., Arrighi, R., Aghakhanyan, G., Tinelli, F., Frijia, F., Crespi, S., et al. (2019). Plasticity of the human visual brain after an early cortical lesion. *Neuropsychologia* 128, 166–177. doi: 10.1016/j.neuropsychologia.2017.10.033

Murphy, T. H., and Corbett, D. (2009). Plasticity during stroke recovery: from synapse to behaviour. *Nat. Rev. Neurosci.* 10, 861–872. doi: 10.1038/nrn2735

Murrough, J. W., and Russo, S. J. (2019). The neurobiology of resilience: complexity and hope. *Biol. Psychiatry* 86, 406–409. doi: 10.1016/j.biopsych.2019.07.016

Olmi, S., Petkoski, S., Guye, M., Bartolomei, F., and Jirsa, V. (2019). Controlling seizure propagation in large-scale brain networks. *PLoS Comput. Biol.* 15:e1006805. doi: 10.1371/journal.pcbi.1006805

Park, B., Kim, D.-S., and Park, H.-J. (2014). Graph independent component analysis reveals repertoires of intrinsic network components in the human brain. *PLoS ONE* 9:e82873. doi: 10.1371/journal.pone.0082873

Park, H. J., and Friston, K. (2013). Structural and functional brain networks: from connections to cognition. *Science* 342:1238411. doi: 10.1126/science.1238411

Park, H. J., Pae, C., Friston, K., Jang, C., Razi, A., Zeidman, P., et al. (2017). Hierarchical dynamic causal modeling of resting-state fmri reveals longitudinal changes in effective connectivity in the motor system after thalamotomy for essential tremor. *Front. Neurol.* 8:346. doi: 10.3389/fneur.2017.00346

Park, H. J., Park, B., Kim, H. Y., Oh, M. K., Kim, J. I., Yoon, M., et al. (2015). A network analysis of (1)(5)O-H(2)O PET reveals deep brain stimulation effects on brain network of parkinson's disease. *Yonsei Med. J.* 56, 726–736. doi: 10.3349/ymj.2015.56.3.726

Park, S., Park, H. J., Kyeong, S. H., Moon, I. S., Kim, M., Kim, H. N., et al. (2013). Combined rTMS to the auditory cortex and prefrontal cortex for tinnitus control in patients with depression: a pilot study. *Acta Otolaryngol.* 133, 600–606. doi: 10.3109/00016489.2012.763181

Potkin, S. G., Kane, J. M., Correll, C. U., Lindenmayer, J. P., Agid, O., Marder, S. R., et al. (2020). The neurobiology of treatment-resistant schizophrenia: paths to antipsychotic resistance and a roadmap for future research. *NPJ Schizophr.* 6:1. doi: 10.1038/s41537-019-0090-z

Proix, T., Bartolomei, F., Guye, M., and Jirsa, V. K. (2017). Individual brain structure and modelling predict seizure propagation. *Brain* 140, 641–654. doi: 10.1093/brain/awx004

Rabinovich, M. I., and Varona, P. (2011). Robust transient dynamics and brain functions. *Front. Comput. Neurosci.* 5:24. doi: 10.3389/fncom.2011.00024

Russo, S. J., Murrough, J. W., Han, M. H., Charney, D. S., and Nestler, E. J. (2012). Neurobiology of resilience. *Nat. Neurosci.* 15, 1475–1484. doi: 10.1038/nn.3234

Saur, D., Lange, R., Baumgaertner, A., Schraknepper, V., Willmes, K., Rijntjes, M., et al. (2006). Dynamics of language reorganization after stroke. *Brain* 129, 1371–1384. doi: 10.1093/brain/awl090

Schreglmann, S. R., Krauss, J. K., Chang, J. W., Martin, E., Werner, B., Bauer, R., et al. (2018). Functional lesional neurosurgery for tremor: back to the future? *J. Neurol. Neurosurg. Psychiatr.* 89, 727–735. doi: 10.1136/jnnp-2017-316301

Siegel, J. S., Seitzman, B. A., Ramsey, L. E., Ortega, M., Gordon, E. M., Dosenbach, N. U. F., et al. (2018). Re-emergence of modular brain networks in stroke recovery. *Cortex* 101, 44–59. doi: 10.1016/j.cortex.2017.12.019

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., et al. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. U.S.A.* 106, 13040–13045. doi: 10.1073/pnas.0905267106

Sporns, O., and Betzel, R. F. (2016). Modular brain networks. *Annu. Rev. Psychol.* 67, 613–640. doi: 10.1146/annurev-psych-122414-033634

Stiso, J., Khambhati, A. N., Menara, T., Kahn, A. E., Stein, J. M., Das, S. R., et al. (2019). White matter network architecture guides direct electrical stimulation through optimal state transitions. *Cell Re.* 28, 2554–2566.e2557. doi: 10.1016/j.celrep.2019.08.008

Tang, E., Giusti, C., Baum, G. L., Gu, S., Pollock, E., Kahn, A. E., et al. (2017). Developmental increases in white matter network controllability support a growing diversity of brain dynamics. *Nat. Commun*. 8:1252. doi: 10.1038/s41467-017-01254-4

Tavor, I., Jones, O. P., Mars, R. B., Smith, S. M., Behrens, T. E., and Jbabdi, S. (2016). Task-free MRI predicts individual differences in brain activity during task performance. *Science* 352, 216–220. doi: 10.1126/science.aad8127

Todorov, E. (2009). Efficient computation of optimal actions. *Proc. Natl. Acad. Sci. U.S.A*. 106, 11478–11483. doi: 10.1073/pnas.0710743106

Tognoli, E., and Kelso, J. A. (2014). The metastable brain. *Neuron* 81, 35–48. doi: 10.1016/j.neuron.2013.12.022

Van Essen, D. C., Ugurbil, K., Auerbach, E., Barch, D., Behrens, T. E., Bucholz, R., et al. (2012). The human connectome project: a data acquisition perspective. *Neuroimage* 62, 2222–2231. doi: 10.1016/j.neuroimage.2012.02.018

Vorovenci, R. J., Biundo, R., and Antonini, A. (2016). Therapy-resistant symptoms in Parkinson's disease. *J. Neural Transm.* 123, 19–30. doi: 10.1007/s00702-015-1463-8

Watanabe, T., Hirose, S., Wada, H., Imai, Y., Machida, T., Shirouzu, I., et al. (2013). A pairwise maximum entropy model accurately describes resting-state human brain networks. *Nat. Commun*. 4:1370. doi: 10.1038/ncomms2388

Watanabe, T., Hirose, S., Wada, H., Imai, Y., Machida, T., Shirouzu, I., et al. (2014a). Energy landscapes of resting-state brain networks. *Front. Neuroinform* 8:12. doi: 10.3389/fninf.2014.00012

Watanabe, T., Kan, S., Koike, T., Misaki, M., Konishi, S., Miyauchi, S., et al. (2014b). Network-dependent modulation of brain activity during sleep. *Neuroimage* 98, 1–10. doi: 10.1016/j.neuroimage.2014.04.079

Watanabe, T., Masuda, N., Megumi, F., Kanai, R., and Rees, G. (2014c). Energy landscape and dynamics of brain activity during human bistable perception. *Nat. Commun*. 5:4765. doi: 10.1038/ncomms5765

Yeh, F. C., Tang, A. N., Hobbs, J. P., Hottowy, P., Dabrowski, W., Sher, A., et al. (2010). Maximum entropy approaches to living neural networks. *Entropy* 12, 89–106. doi: 10.3390/e12010089

Yeo, B. T. T., Krienen, F. M., Eickhoff, S. B., Yaakub, S. N., Fox, P. T., Buckner, R. L., et al. (2015). Functional specialization and flexibility in human association cortex. *Cerebr. Cortex* 25, 3654–3672. doi: 10.1093/cercor/bhu217

Yu, R., Park, H. J., Cho, H., Ko, A., Pae, C., Oh, M. K., et al. (2018). Interregional metabolic connectivity of 2-deoxy-2[(18) F]fluoro-D-glucose positron emission tomography in vagus nerve stimulation for pediatric patients with epilepsy: a retrospective cross-sectional study. *Epilepsia* 59, 2249–2259. doi: 10.1111/epi.14590

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.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership