UNDERSTANDING THE IMPORTANCE OF TEMPORAL COUPLING OF NEURAL ACTIVITIES IN INFORMATION PROCESSING UNDERLYING ACTION AND PERCEPTION

EDITED BY: Daya Shankar Gupta and Andreas Bahmer

PUBLISHED IN: Frontiers in Neuroscience, Frontiers in Computational Neuroscience, Frontiers in Integrative Neuroscience and Frontiers in Psychology



frontiers Research Topics



frontiers

Frontiers eBook Copyright Statement

The copyright in the text of individual articles in this eBook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this eBook is the property of Frontiers.

Each article within this eBook, and the eBook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this eBook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or eBook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-88971-512-1
DOI 10.3389/978-2-88971-512-1

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

UNDERSTANDING THE IMPORTANCE OF TEMPORAL COUPLING OF NEURAL ACTIVITIES IN INFORMATION PROCESSING UNDERLYING ACTION AND PERCEPTION

Topic Editors:
Daya Shankar Gupta, Husson University, United States
Andreas Bahmer, University Hospital Frankfurt, Germany

Citation: Gupta, D. S., Bahmer, A., eds. (2021). Understanding the Importance of Temporal Coupling of Neural Activities in Information Processing Underlying Action and Perception. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88971-512-1

Table of Contents

05 ***Editorial: Understanding the Importance of Temporal Coupling of Neural Activities in Information Processing Underlying Action and Perception***
Daya Shankar Gupta and Andreas Bahmer

09 ***Memory Load Influences Taste Sensitivities***
Pei Liang, Jiayu Jiang, Qingguo Ding, Xiaoyan Tang and Soumyajit Roy

16 ***Goal-Directed Processing of Naturalistic Stimuli Modulates Large-Scale Functional Connectivity***
Zhenfu Wen, Tianyou Yu, Xinbin Yang and Yuanqing Li

28 ***Masking the Integration of Complementary Shape Cues***
Andrew Geoly and Ernest Greene

39 ***Modulation of Functional Connectivity Between Dopamine Neurons of the Rat Ventral Tegmental Area* in vitro**
Luuk van der Velden, Martin A. Vinck, Taco R. Werkman and Wytse J. Wadman

53 ***The Orientation Selectivity of Spike-LFP Synchronization in Macaque V1 and V4***
Zhaohui Li, Mengyu Gao and Yongtian Wang

61 ***The Lateralization of Speech-Brain Coupling is Differentially Modulated by Intrinsic Auditory and Top-Down Mechanisms***
M. F. Assaneo, J. M. Rimmele, J. Orpella, P. Ripollés, R. de Diego-Balaguer and D. Poeppel

72 ***Corrigendum: The Lateralization of Speech-Brain Coupling is Differentially Modulated by Intrinsic Auditory and Top-Down Mechanisms***
M. F. Assaneo, J. M. Rimmele, J. Orpella, P. Ripollés, R. de Diego-Balaguer and D. Poeppel

73 ***The Inter-Regional Connectivity Within the Default Mode Network During the Attentional Processes of Internal Focus and External Focus: An fMRI Study of Continuous Finger Force Feedback***
Zhi-Wei Zhou, Xia-Qing Lan, Yan-Tong Fang, Yun Gong, Yu-Feng Zang, Hong Luo and Hang Zhang

86 ***Musical Role Asymmetries in Piano Duet Performance Influence Alpha-Band Neural Oscillation and Behavioral Synchronization***
Auriel Washburn, Irán Román, Madeline Huberth, Nick Gang, Tysen Dauer, Wisam Reid, Chryssie Nanou, Matthew Wright and Takako Fujioka

104 ***Tracking the Dynamic Functional Network Interactions During Goal-Directed Auditory Tasks by Brain State Clustering***
Gaoyan Zhang, Yuexuan Li and Jinliang Zhang

114 ***Temporal Discrimination Thresholds and Proprioceptive Performance: Impact of Age and Nerve Conduction***
Thorsten M. Odorfer, Teresa Wind and Daniel Zeller

123 ***A Spike Train Distance Robust to Firing Rate Changes Based on the Earth Mover's Distance***
Duho Sihn and Sung-Phil Kim

138 ***Neural Synchrony Gives Rise to Amplitude- and Duration-Invariant Encoding Consistent With Perception of Natural Communication Stimuli***
Michael G. Metzen, Volker Hofmann and Maurice J. Chacron

159 ***Acoustic Change Responses to Amplitude Modulation in Cochlear Implant Users: Relationships to Speech Perception***
Ji-Hye Han and Andrew Dimitrijevic

170 ***Resonance in the Mouse Ventral Tegmental Area Dopaminergic Network Induced by Regular and Poisson Distributed Optogenetic Stimulation*** **in-vitro**
Luuk van der Velden, Martin A. Vinck and Wytse J. Wadman

182 ***Decoding Adaptive Visuomotor Behavior Mediated by Non-linear Phase Coupling in Macaque Area MT***
Mohammad Bagher Khamechian and Mohammad Reza Daliri

198 ***A Model of Memory Linking Time to Space***
Hubert Löffler and Daya Shankar Gupta

209 ***Top-Down Control of Inhibitory Granule Cells in the Main Olfactory Bulb Reshapes Neural Dynamics Giving Rise to a Diversity of Computations***
Zhen Chen and Krishnan Padmanabhan

226 ***The Low-Frequency Fluctuation of Trial-by-Trial Frontal Theta Activity and Its Correlation With Reaction-Time Variability in Sustained Attention***
Yao-Yao Wang, Li Sun, Yi-Wei Liu, Jia-Hui Pan, Yu-Ming Zheng, Yu-Feng Wang, Yu-Feng Zang and Hang Zhang

238 ***Oscillatory Bursting as a Mechanism for Temporal Coupling and Information Coding***
Idan Tal, Samuel Neymotin, Stephan Bickel, Peter Lakatos and Charles E. Schroeder

frontiers in Computational Neuroscience

EDITORIAL
published: 01 September 2021
doi: 10.3389/fncom.2021.729296

Editorial: Understanding the Importance of Temporal Coupling of Neural Activities in Information Processing Underlying Action and Perception

Daya Shankar Gupta[1]* *and Andreas Bahmer*[2,3]

[1] Biology Department, Camden County College, Blackwood, NJ, United States, [2] Ernst Strüngmann Institut, Frankfurt, Germany, [3] Brain Imaging Center Frankfurt, University of Frankfurt, Frankfurt, Germany

Keywords: binding, entropy, decorrelation, mutual information, affordance

Editorial on the Research Topic

Understanding the Importance of Temporal Coupling of Neural Activities in Information Processing Underlying Action and Perception

INTRODUCTION

The aim of this study was to understand the role of the temporal coupling of neural events in information processing underlying action and perception. Action and perception are optimized for a successful interaction of an organism with its environment to carry out tasks needed for its survival.

Information processing in the brain during an interaction with the external environment leads to an increase in mutual information and surprisal information (Gupta and Bahmer, 2019). Mutual information is a general measure of the strength of the association between two variables (Gupta and Bahmer, 2019). Pairs of variables, underlying the changes in the mutual information in the brain, may be represented by the spiking activities of pairs of nodes in a brain network, i.e., low-frequency vs. high-frequency neural oscillations during cross-frequency interaction or spiking vs. local field potential (LFP). Task-induced association between these variables due to temporal coupling can increase the mutual information and reduce the surprisal information in the brain that results from the sensory processing of stimuli, leading to a successful interaction with the external environment. Previous experimental studies have also supported an important role of temporal coupling in different perceptual tasks (Bahmer and Gupta, 2018). Moreover, the temporal coupling of neural events during information processing underlying action and perception can be of different degrees, from a less tight to a more tight degree (Gupta et al., 2020). Furthermore, binding between two brain areas or a brain activity and an external stimulus feature can emerge from temporal coupling (von der Malsburg, 1995).

Many studies in the Frontiers' Research Topics shed light on the role of the temporal coupling in information processing in the brain. Several lines of evidence presented in many contributions indicate that there is temporal coupling and an increase in mutual information during information processing in the central nervous system.

Furthermore, various parameters to study the association between neural activities, reflecting mutual information, are reported by the contributing manuscripts, which include spike–gamma LFP coherence, paired phase consistency (PPC), spike train distance, and bicoherence. In

OPEN ACCESS

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

***Correspondence:**
Daya Shankar Gupta
dayagup@gmail.com

Received: *22 June 2021*
Accepted: *29 June 2021*
Published: *01 September 2021*

Citation:
Gupta DS and Bahmer A (2021) Editorial: Understanding the Importance of Temporal Coupling of Neural Activities in Information Processing Underlying Action and Perception. Front. Comput. Neurosci. 15:729296. doi: 10.3389/fncom.2021.729296

addition, many studies have highlighted the importance of synchronization in information transfer from one region of the brain to the other. Neural oscillations of slow frequency may represent the synchronization of brain networks involved in information transfer whereas high-frequency oscillations may predominantly reflect local neuronal activities (Herrmann et al., 2016). Information transfer would occur if the phase of the excitability of neurons matches with the arrival of inputs (Fries, 2015). Note that evidence from experiments based on LFP (MEG, EEG, fMRT) must be cautiously interpreted because synaptic activity from the sender projecting to the receiver network can result in false-positive coherence measures (Schneider et al., 2020). Moreover, synchronous states increase the probability of joint activity of pairs of neurons in networks, which contributes to the increase in mutual information. The synchronous activity of specific pairs of neurons, in neuron populations, modulated by tasks also reduces the surprise associated with the sensory stimuli, which contributes to a successful interaction of the brain with the environment *via* optimal perception and action (Gupta and Bahmer, 2019).

CONTRIBUTIONS

Interaction With the Environment

Wen et al. investigated how task goals modulate whole brain functional connectivity when human subjects were asked to classify either taxonomic type or behavior type of behaving animals under naturalistic conditions shown in video clips. To study whole brain functional connectivity, the authors used inter-subject functional correlation. This method eliminates intrinsic signals by calculating the inter-regional correlations between different subjects performing the same task. Their findings show that whole brain functional connectivity was modulated by different task goals.

In a study of interpersonal interaction, a piano duet task was designed with three types of performer roles in the duet, namely, starting vs. joining, musical task similarity, and performer animacy (human vs. computer) (Washburn et al.). The authors noted that there are lasting effects of musical ensemble performance on attentiveness, perceptual-motor coordination, and empathy.

In a method article, Sihn and Kim studied a measure of synchrony and temporal similarity between spike trains called spike train distance. They used a method called Earth Mover's Distance (EMD) to compute spike train distance. EMD is sensitive to the temporal pattern but robust to firing rate changes. Since many of the cognitive functions are dependent on temporal patterns (Bahmer and Gupta, 2018) rather than firing rate, EMD may be used to study the effect of similarity or dissimilarity of spike patterns of pairs of neurons on cognitive functions of the brain. Smaller values of EMD in spiking of neurons suggest a greater similarity. Using the actual data from recording in the monkey motor cortex, the authors found that EMD increased as the angle was becoming orthogonal to the preferred direction. This finding underscores the importance of the temporal pattern of neuron firing in coding directional sensitivity.

In a species of weakly electric fishes (*Apteronotus leptorhynchus*), Metzen et al. investigated the coding of natural electrocommunication signals, called chirps. *A. leptorhynchus* are known to robustly give chirp echo responses when stimulated with chirps, which is used for the study of perception in these animals. Using the characteristics of chirp stimuli that are common during the interaction between same-sex conspecifics, the authors showed that synchrony in neuronal activities at all levels increased transiently in a similar fashion in response to these chirps. Furthermore, synchrony at the population level was a much better detector of the chirp stimulus than the single afferent activity. The increase in population synchrony, which promotes information transfer, was invariant to chirp attributes, namely, duration and amplitude (i.e., transient increase in frequency). The behavioral response to chirps was studied by chirp echo response rates, which was also invariant to variations in chirp attributes. During natural interactions between same-sex conspecifics, a simple behavior, giving a chirp echo response, which excludes complex behaviors, such as mating and escape, is sufficient. Thus, the invariance to chirp attributes in current conditions will not be detrimental but would only enhance their interactions with the same-sex conspecifics.

The interaction with the physical world depends on various functions of the nervous system. These include the proprioceptive performance and the discrimination of two sensory stimuli and separate movements by the shortest intervals. Odorfer et al. studied the confounds of aging and polyneuropathy on these functions by studying somatosensory temporal discrimination threshold (STDT), temporal discrimination movement threshold (TDMT), and behavioral measures of proprioception of upper and lower limbs. This study shows that aging resulted in higher STDT and TDMT but had no influence on proprioceptive performance. However, polyneuropathy resulted in higher STDT and TDMT with poor proprioceptive performance. These findings provide the objective basis for the decline in cognitive functions in older individuals, requiring interaction with their environment, such as sports activities.

Attentional Modulation

Earlier studies showed that the default mode network (DMN) shows higher activity at rest compared with tasks involving attention. However, recent studies have suggested that DMN is also engaged during tasks involving attentional modulation. Consistent with the recent trend, Zhou et al. showed that there is a greater connectivity between two nodes of DMN, namely, posterior cingulate cortex and left inferior parietal cortex/angular gyrus in attentional tasks involving external focus in comparison with internal focus. The task paradigm required maintaining a pressure of 20 cm of water between the right index finger and thumb. External focus recruited attentional process based on the direct feedback about the pressure levels. Internal focus involved attentional modulation based on tactile sensory input and memory.

Capacity for sustained attention is important for the interaction of the brain with the environment. Wang et al. studied variability in reaction time and trial-by-trial frontal theta activity in individuals performing sustained attention tasks to understand the electrophysiological underpinnings of attention. Variability in reaction time and trial-by-trial frontal theta activity were assessed by SD and the amplitude of low-frequency fluctuation (ALFF).

The authors reported that the ALFF of reaction time variability has a significant correlation with the ALFF of the trial-by-trial frontal theta activity in a frequency-dependent manner.

Temporal Coupling

In a study by Geoly and Greene, when two complementary subsets of sparse dots are presented to represent a shape, separated by an interval between 0 and 500 ms, the probability of a correct match with the target shape remained above chance. When the complementary subsets of dots were presented simultaneously, the probability of match recognition was the highest. The probability of match recognition decreased with increasing the interval between complementary subsets of dots. These results suggest the importance of the temporal coupling of information from two subsets of dot patterns, where data from two subsets of dots are separated by the duration, ranging from 0 to 500 ms (Gupta et al., 2020). Another experiment displayed the complementary subsets with 200 ms of separation, and the match recognition was disrupted when the random dot mask was displayed midway between the two subsets. When the random dot pattern mask was presented midway between the two subsets of complementary information, it affected both subsets of information to the same extent by temporal coupling, which resulted in a complete disruption of the match recognition.

Li et al. studied spike–gamma LFP coherence by placing chronically implanted multielectrode arrays over V1 and V4 in two macaques performing a selective visual attention task. The visual stimulus used was sinusoidal gratings with orientation, between 0° and 360°, presented in pseudorandom order. The authors found that the spike–LFP synchronization strength between V1 and V4 shows orientation selectivity to drifting gratings. Li et al. further argued that synchronization between different regions not only reflects the basic features of visual stimulation but also describes the orientation tuning characteristics of neurons. This is consistent with the argument that the increase in mutual information, which results from the synchronization between different regions of the brain, can represent complex stimulus characteristics, such as orientation (Gupta and Bahmer, 2019).

van der Velden, Vinck, Werkman et al. simultaneously recorded spontaneous extracellular spikes from 10 to 30 dopamine neurons in acute slices from the lateral ventral tegmental area (VTA) of the rat. The functional connectivity between pairs of neurons was analyzed by PPC, which estimates the square of phase-locking value. Manipulating excitability with high extracellular potassium reduced the PPC, but the application of glutamate did not have any effect on PPC. Since the application of glutamate would affect excitability *via* synaptic connections, the reduced PPC after the application of high K+ is partly due to the uncorrelated activity of pairs of neurons that have no synaptic connections.

It is currently known that during speech comprehension, there is an alignment of the neuronal excitability phase of slow oscillations in the auditory cortex with slow energy fluctuations in the speech or the attributes of the speech envelope (Assaneo et al.), which is referred to as speech tracking. This alignment, indicating temporal coupling, was estimated by computing the phase-locking value between the brain activity and the cochlear envelope (Ding et al., 2017) of the perceived stream of syllables. Assaneo et al. showed that the asymmetry of the auditory tracking is reversed by the presence of semantic information even though the acoustic properties of the stimuli are similar. Hence, their findings reveal the importance of temporal binding between the auditory stimulus and the brain activity in speech perception.

Khamechian and Daliri analyzed non-linear neuronal synchronization in LFPs recorded from the middle temporal lobe signals during a visuomotor task by employing the bicoherence method to examine how non-linear neuronal synchronization in the MT area is involved in the processing of visuomotor information. Bicoherence, a study of two frequencies in a single signal, gives a maximum value if there is a perfect phase-locking and a minimum value if there is a random overlap between the phases of two frequencies. Notably, information transfer depends on phase-locking. Thus, the study of the characteristics of non-linear neuronal synchronization would be important for understanding the complex dynamics of information transfer in cognitive tasks.

In a study employing midbrain slices, van der Velden, Vinck and Wadman. combined the optogenetic stimulation of dopaminergic neurons in the VTA with the recording of action potentials. After stimulation with regular optogenetic pulses, the authors found the highest resonance at 2.9 Hz, which is the intrinsic frequency of the VTA neurons. However, after stimulation with stochastically distributed pulses, maximum resonance was noted at a subharmonic frequency of 1.5 Hz. As authors have noted, wide-band stochastic inputs to dopaminergic neurons can induce synchronous states. Also, it is noted that the synchronous states promote information flow. This is particularly interesting in the study of the response of the VTA in unpredicted rewards (Morales and Margolis, 2017). The stochastic inputs to dopaminergic neurons, resulting from unexpected rewards, can also induce synchronous states for information processing.

Information Processing in the Brain

Using a biologically plausible simulation study, Löffler and Gupta showed that input patterns can be encoded by the coincidence detection in dendrites. When 100 Hz synchronized inputs, from I-neuron (source of activity pattern to be encoded) and A-neuron with 100 Hz regular discharge rate, coincide with the peak of the 8.33 Hz subthreshold membrane potential oscillations at one of the dendritic branches, this results in a dendritic spike leading to a somatic spike. In this model, a single dendritic spike increases the synaptic weight by ∼37% at corresponding synapses. An increase in the synaptic weight at specific synapses is responsible for reproducing the same pattern of activity alone by a 100 Hz regular input even in the absence of I-neuron activity. Since synaptic weights are increased at specific synapses, depending on the temporal pattern of the input from I-neuron, this may provide a biologically plausible basis for the temporal processing of information.

In an fMRI study of dynamic functional network connectivity in subjects performing auditory discrimination and working memory tasks, the authors used the independent component analysis to extract networks, which included the auditory, the visual, the sensorimotor, the cerebellar, the frontoparietal, the

default mode, and the salience networks (Zhang et al.). The sliding window analysis, with each step of 30 s and a total of 178 steps, was used to identify four states, resulting from various configurations of seven networks. The analysis of four states in two tasks revealed distinct dynamic functional connectivity of the networks. It should be noted that there are changes in mutual and surprisal information, which are due to distinctive dynamic connectivity in two tasks. This suggests that distinct dynamic connectivity rather than quantitative changes in mutual and surprisal information underlies the differences in the task goals.

A human behavioral study by Liang et al. found that the memory load, i.e., remembering an increasing visually presented list of alphanumerical items, reduces sweet and bitter taste sensitivities, which are consistent with other studies of the effects of memory load on perception.

Han and Dimitrijevic studied (1) the effect of amplitude modulation (AM) depth on the detection of amplitude modulated white noise and (2) the interaction between cortical N1 responses, hearing performance, and AM changes (4, 40, 100, and 300 Hz) in postlingually deafened subjects with bilateral cochlear implants (CIs) for speech perception in CI users. They showed that AM change stimuli can elicit robust N1 acoustic change complex responses (4 and 40 Hz) in CI users. The N1 latency to 40 Hz (not 4 Hz despite robustness) relates to speech perception measures, and temporal modulation transfer function relates well to speech perception in CI users.

In a dynamic three-node network model studied by Chen and Padmanabhan, consisting of local excitatory mitral/tufted cells and inhibitory granule cells in the olfactory bulb and excitatory cells in the piriform cortex with feedback connections to the granule cells, different computations are possible with the changes in the weight of top-down connections to the granule cells. Different computations produced by the changes in the weight of top-down connections include (1) separating two external (olfactory) stimuli based on rate coding and (2) synchronizing the oscillatory activities of two separate stimuli. The authors also suggested that weight changes can be biologically initiated by neuromodulators. Notably, while synchronization can increase the mutual information, the pattern separation based on rate coding would increase the surprisal information.

Tal et al. have reviewed the study supporting that neural oscillation may occur as transient burst-like events in addition to a sustained increase in power. Burst-like oscillations are transient events, which are separated by temporal gaps. Due to averaging over many trials to increase the signal-to-noise ratio, it was difficult to detect burst-like oscillations previous studies. However, the authors argued that the trial-by-trial changes in oscillatory dynamics must be studied to understand their role in perception and action. The authors also offered advice about choosing suitable analyses to study neural rhythms in single trials.

SUMMARY

A review of various contributions in the Frontiers' Research Topics suggests that synchronous activities can have the following consequences: information flow, temporal coupling, and an increase in mutual information. The synchronization would also reduce the surprise by increasing the coactivation of pairs of neurons. Notably, an initial increase in surprisal information would result from information flow. Both the increase in the mutual information and surprisal information are the important underpinnings of action and perception, subserving the interaction of the brain with the environment (Gupta and Bahmer, 2019).

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

We are grateful to the contributors for their excellent work submitted to this Research Topic.

REFERENCES

Bahmer, A., and Gupta, D. S. (2018). Role of oscillations in auditory temporal processing: a general model for temporal processing of sensory information in the brain? *Front. Neurosci.* 12:793. doi: 10.3389/fnins.2018.00793

Ding, N., Patel, A. D., Chen, L., Butler, H., Luo, C., and Poeppel, D. (2017). Temporal modulations in speech and music. *Neurosci. Biobehav. Rev.* 81(Pt B), 181–187. doi: 10.1016/j.neubiorev.2017.02.011

Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235. doi: 10.1016/j.neuron.2015.09.034

Gupta, D. S., and Bahmer, A. (2019). Increase in mutual information during interaction with the environment contributes to perception. *Entropy* 21:365. doi: 10.3390/e21040365

Gupta, D. S., Banerjee, A., Roy, D., and Piras, F. (2020). Editorial: temporal structure of neural processes coupling sensory, motor and cognitive functions of the brain. *Front. Comput. Neurosci.* 14:73. doi: 10.3389/fncom.2020.00073

Herrmann, C. S., Struber, D., Helfrich, R. F., and Engel, A. K. (2016). EEG oscillations: from correlation to causality. *Int. J. Psychophysiol.* 103, 12–21. doi: 10.1016/j.ijpsycho.2015.02.003

Morales, M., and Margolis, E. B. (2017). Ventral tegmental area: cellular heterogeneity, connectivity and behaviour. *Nat. Rev. Neurosci.* 18, 73–85. doi: 10.1038/nrn.2016.165

Schneider, M., Dann, B., Sheshadri, S., Scherberger, H., and Vinck, M. (2020). A general theory of coherence between brain areas. *bioRxiv [Preprint]*. doi: 10.1101/2020.06.17.156190

von der Malsburg, C. (1995). Binding in models of perception and brain function. *Curr. Opin. Neurobiol.* 5, 520–526. doi: 10.1016/0959-4388(95)80014-x

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

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

Copyright © 2021 Gupta and Bahmer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 11 December 2018
doi: 10.3389/fpsyg.2018.02533

Memory Load Influences Taste Sensitivities

Pei Liang[1,2,3,4], Jiayu Jiang[1,2], Qingguo Ding[3], Xiaoyan Tang[4] and Soumyajit Roy[5]**

[1] Department of Psychology/Facuty of Education, Hubei University, Hubei, China, [2] Brain and Cognition Research Center (BCRC), Faculty of Education, Hubei Univeristy, Hubei, China, [3] The No. 2 Peoples' Hospital of Changshu, Changshu, China, [4] Changshu Institute of Technology, Changshu, China, [5] Eco-Friendly Applied Materials Laboratory, Materials Science Centre, Department of Chemical Sciences, Indian Institute of Science Education and Research, Kolkata, India

Previous literature reports have demonstrated that taste perception would be influenced by different internal brain status or external environment stimulation. Although there are different hypotheses about the cross-modal interactive process, it still remains unclear as of how the brain modulates and processes taste perception, particularly with different memory load. Here in this study we address this question. To do so we assign the participants different memory loads in the form of varying lengths of alphanumerical items, before tasting different concentrations of sweet or bitter tastants. After tasting they were asked to recall the alphanumerical items they were assigned. Our results show that the memory load reduces sweet and bitter taste sensitivities, from sub-threshold level to high concentration. Higher the memory load, less is the taste sensitivity. The study has extended our previous results and supports our previous hypothesis that the cognitive status, such as the general stress of memory load, influences sensory perception.

Keywords: cross-modal, memory load, cognitive status, sweetness perception, bitterness perception

OPEN ACCESS

Edited by:
Andreas Bahmer, Universitätsklinikum Würzburg, Germany

Reviewed by:
Anna Magnusson, Karolinska Institutet (KI), Sweden
Amael Arguel, Macquarie University, Australia

****Correspondence:***
Pei Liang
liangpei0108@126.com
Soumyajit Roy
roy.soumyajit@gmail.com

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Psychology

Received: *30 August 2018*
Accepted: *27 November 2018*
Published: *11 December 2018*

Citation:
Liang P, Jiang J, Ding Q, Tang X and Roy S (2018) Memory Load Influences Taste Sensitivities. Front. Psychol. 9:2533. doi: 10.3389/fpsyg.2018.02533

INTRODUCTION

In today's fast paced society, on one side, many people have quick meals "on the go." Their cognitive brain is still busily processing something related with work while chewing and swallowing meals. It has been suggested that cognitive load would distract the attention and reduce sensory perception. Several studies have demonstrated that taste perception may be influenced by internal brain state such as attention and awareness (Elder and Krishna, 2009). The literature has been very mixed regarding the influence's effect between attention and multi-sensory integration (Odegaard et al., 2016). For instance, some investigations have found that attention has no effect on multisensory integration (Bertelson et al., 2000; Shore and Simic, 2005), while other studies have reported that selective attention can increase integration (Alsius et al., 2005), or reduce integration (Mozolic et al., 2008; Talsma et al., 2010). Yantis (2000) has proposed the competition of limited attention resources. When attention is distracted by other information resources like TVs, the taste perception would get less sensitive and the hedonic rating of the food would be reduced. Similarly, people experience tastants less intense when the environment has loud noise (Spence and Shankar, 2010). Stafford et al. (2012, 2013) have demonstrated that music and distraction may alter taste perception of alcohol. How the brain processes the sensory perception, particularly involved with the neural network of higher level of cognition, remains unclear. Kahneman (2011) has proposed two cognitive processes. Accordingly, the cognitive states are divided into "cognitive ease" and "cognitive strain." The cognitive ease means people feel effortless and comfortable.

When the brain is at cognitive ease state, the response is fast and intuitive. On the contrary, when the brain is at cognitive strain state, people feel less comfortable or stressed. It leads people to increase attention and to invest more effort and the corresponding response is more critical. Our previous studies have demonstrated that visual information affects taste sensitivities (Liang et al., 2013, 2016). For instance, the visual stimuli representing color, shapes, or symbols induce attentive mechanisms, which are difficult to calibrate with respect to the cognitive ease and load concept. And such components might contribute differently to affecting the gustatory perception. Motivated by Kahneman's theory and the previous research, we hypothesized that the cognitive status plays a key role in gustatory perception. More specifically, we would like to focus this study on cognitive strain and observe how taste perception thresholds are affected by a simple memory load task.

Recently van der Wal and van Dillen (2013) have tested whether task load influences the sweet, sour, and salty perception. In their study they demonstrated that the task load reduces not only aversive tastants, such as sourness of lemon juice, but also pleasant tastants like grenadine syrup and salty butter. However, in their study, the tastants are more complex, and the tastants concentrations are relatively higher compared with taste threshold level. It would be therefore important to observe how the basic taste sensitivity changes at lower concentrations to avoid the saturation of the taste receptors or central habituation. Hence, our present paper is to study how the sweetness and bitterness sensitivity is influenced with different degrees of memory load, particularly, at low concentration level (around taste threshold).

To test if cognitive strain influences the perception thresholds, we manipulated the degree of memory load by varying the length of alphanumerical items. Two basic but hedonically opposite tastes, i.e., sweetness and bitterness were applied and calibrated against the memory load. The results confirm the hypothesis that cognitive load reduces taste perception.

MATERIALS AND METHODS

Subjects

Twenty-six student volunteers (sixteen females and ten males) from Changshu Institute of Technology (CIT), China were chosen for the experiments. They were all self-reported right-handed and had normal eyesight or at least were corrected to normal by glasses. None of them was color blind and their ages were between 21 and 30 years old (average 25 $\pm$ 3 years). They did not have any taste- or smell-related disease before. All the participants were well briefed about the details of the experiments and of their performance. They all agreed and signed on the written informed consent declaration to volunteer as subjects in these experiments. The study was approved by the Ethics Committee (IEC) of the No. 2 Peoples' Hospital of Changshu (license number 20151101), according to the Ethics Guidelines.

Tastants Preparation

We applied sucrose solution as sweet tastants. Sucrose was dissolved in distilled water to prepare the sugar solution with concentration of 0, 1.5, 3.1, 3.9, 4.7, and 5.5 g/L, respectively (Liang et al., 2013). The phenylthiourea solution was prepared with concentration of 0, 0.02, 0.04, 0.08, 0.16, and 0.24 mM/L, respectively. All the sweet and bitter solutions were prepared one night before and kept on the table at the room temperature between 20 and 25°C. During the experiments the solutions were provided to the subjects in a series of half-filled odorless white paper cups (25 ml).

Memory Load on Display

Before tasting experiments, the participants were visually exposed to a list of alphanumerical items on the monitor for 2 s. They were requested to remember and to repeat the information given to them after each block of taste experiments (refer to **Figure 1** Experiment flow). The memory load task was modified by changing the length of the list of alphanumerical items. For example, "1C," "L1G0," "6C1A8Z," "G4S3J1Z8" represent the four groups of memory tasks with increasing cognitive load. The number and alphabet are in a staggered random arrangement without any possibility of association and remembrance by the volunteers, during repetition of the experiments. The information of memory task was displayed on 17-inch LCD Monitor with 60 Hz refresh rate.

Experiment Training

The experiments are carried out at the sensory science laboratory, Changshu Institute of Technology, Jiangsu, China. There are two purpose of the experiment training: (1) to pre-test the sweet and bitter taste threshold level and (2) to confirm that tastants concentrations are suitable for the experiments. To measure the sweetness or bitterness threshold, the subjects performed the sweet or bitter taste series of six different concentrations randomly without being exposed to visual stimuli. The sweetness and bitterness thresholds were measured by the staircase method and forced choice tracking procedures (refer to the Pepino and Mennella, 2010). The participants were trained for 2–4 times to perform the memory load task and taste experiments, till they were confident and comfortable with the experiments. To avoid the influence of hunger status of the subjects, the experimental data of the subjects were collected at a fixed time of the day (around 1 h after food intake) while repeating the experiments on different days.

Procedures

During the tasting experiments, the subject sat in front of the table. Six cups of sugar solution with different concentrations were placed on the table next to the subject. Before tasting experiments, the monitor would display the list of alphanumerical items, i.e., randomly generated numbers and alphabets, for 2 s. The tasting experiments followed. For the sugar solution experiments, the subjects sipped the sugar solution from the paper cup (around 12 ml) into the mouth and moved the tip of the tongue slightly, keeping the solution in the mouth for 5 s and spitted it out. During the following 50 s pause, the subjects rinsed mouth twice with distilled water, answered the questionnaire whether they detected sweet taste from the corresponding solution. The bitter taste experiments were carried

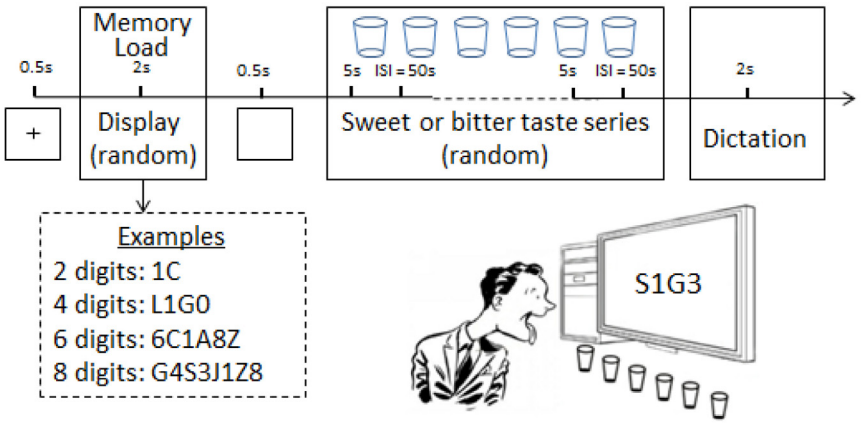

FIGURE 1 | The upper part represents the time schematics of experiment flow. "+"is displayed on the center of monitor for 0.5 s at the beginning of each trial. The below left side describes the examples of lists of alphanumerical items, which consist of alphabets and numbers in random order and are displayed on the center of the monitor in front of the subjects. Six cups represent the six different concentrations of sweet or bitter tastants. ISI means the 50 s rest for the subjects after tasting each cup. In the end of the whole blocks, the subject has 2 s to recall the list of alphanumerical items. The below right side depicts the experiment setup.

out similarly. The subjects needed to finish all the different taste concentrations in the sweet or bitter taste blocks. We used "1," "−1" and "0" to record the results of the taste experiment. When the participant detected the sweetness from the solution, "1" is recorded. When participant detected bitterness from the solution, "−1" is recorded. If the participant did not detect any sweet nor bitter taste, "0" is recorded. In the block experiment, the sweet or bitter solutions of different concentrations were provided to the subjects in random order, respectively. The participants were requested to recall the visual information after each trial. The recall feedbacks would not be reported to the participants. And the next trial will restart after 2 s. The memory task with different length of digital inputs for the subjects was generated randomly in a complete block design. Every participant needed to perform the whole set of the blocks and the trials were repeated ten times. All the experiments were carried out at room temperature 20–25°C.

Data Analysis

All the data were recorded and saved in the computer (Window system 7) and were analyzed offline with MATLAB 7.9 (The MathWorks, Natick, MA, United States). We calculated *the sweetness or bitterness detection ratio* of each memory load for each person. *The taste detection ratio = the number of times when sweet or bitter taste detected with each memory task/the number of times of total experiments repeated with the corresponding task*. For each memory load, the taste detection ratio of each tastants concentration is first analyzed per person. The average detection ratio and the standard deviation across 26 persons are calculated accordingly. For each concentration of sweet or bitter taste, the within group one way repeated ANOVA was used to test the significance of differences in the taste detection ratios with different memory load. And the *post hoc* test Bonferroni adjusted was used to perform the pairwise comparison within the group.

RESULTS

With manipulation of different degrees of memory load (five types), we tested the sweet and bitter taste detection of participants with a series of different concentrations of sugar and phenylthiourea (each taste includes six groups). The results show that when the memory load increases, both the sweet and bitter taste detection ratios decrease significantly (detail data refer to **Tables 1**, **2**). During the training phase, the threshold of sweetness was found between 3.1 and 3.9 g/L, and bitterness threshold between 0.04 and 0.08 mM of phenylthiourea solution. These data were similar to our previous lab observations (Liang et al., 2013). For the taste experiments with cognitive load, 2 (Tastants: Sweet vs. bitter taste) × 6 (Groups: six different concentrations for each tastants) × 5 (Types: five different lengths of memory load tasks) within-group repeated ANOVAs were done separately for the six groups of each tastants, and the levels compared pairwise in *post hoc* tests. The averaged taste detection ratios of different concentrations of both sweet and bitter taste are illustrated in **Figure 2**, and supplemented by **Tables 1**, **2**.

With the still water, as expected, almost no taste was detected with whatever memory load (**Figure 2A**). At very low concentration, the detection ratio decreases significantly with increasing memory load (**Figure 2B**, gray line). Although the averaged bitterness detection has similar tendency as the sweetness detection line, no significant effect has been found (**Figure 2B**, black line). Repeated ANOVA result shows that when the bitter solution concentration is at 0.04 mM, the bitterness detection ratio decreases significantly with increasing memory load [**Figure 2C**, black line, $F(4,105) = 17.070$, $p < 0.001$]. When bitter taste concentration got even higher (above the threshold level), the decreasing trend of taste detection ratio became more obvious [**Figures 2D,E**, black lines, D: $F(4,105) = 33.041$, $p < 0.001$; E: $F(4,105) = 11.498$, $p < 0.001$]. Repeated ANOVA results shows that when sweet solution is 3.1 g/L, the sweetness detection ratio decreases significantly with increasing memory

TABLE 1 | Sweetness detection ratios of different sugar concentration under variant memory load.

	Solution concentration					
Task load	**0 g/L**	**1.5 g/L**	**3.1 g/L**	**3.9 g/L**	**4.7 g/L**	**5.5 g/L**
0	0.01 ± 0.02	0.24 ± 0.10	0.58 ± 0.20	0.89 ± 0.10	0.95 ± 0.06	1.00 ± 0.01
2	0.00 ± 0.00	0.13 ± 0.14	0.49 ± 0.18	0.73 ± 0.12	0.88 ± 0.09	1.00 ± 0.00
4	0.00 ± 0.00	0.11 ± 0.09	0.44 ± 0.16	0.68 ± 0.14	0.84 ± 0.11	0.99 ± 0.03
6	0.00 ± 0.00	0.09 ± 0.08	0.33 ± 0.22	0.57 ± 0.21	0.83 ± 0.17	0.98 ± 0.04
8	0.00 ± 0.00	0.06 ± 0.08	0.27 ± 0.15	0.52 ± 0.16	0.69 ± 0.19	0.95 ± 0.08

Horizontal list labels represent the sugar concentration and the vertical list labels the length of alphanumerical memory load.

TABLE 2 | Bitterness detection ratios of different concentration under variant memory load.

	Solution concentration					
Task load	**0 mM**	**0.02 mM**	**0.04 mM**	**0.08 mM**	**0.16 mM**	**0.24 mM**
0	0.00 ± 0.00	0.18 ± 0.07	0.55 ± 0.11	0.93 ± 0.66	1.00 ± 0.00	1.00 ± 0.00
2	0.00 ± 0.00	0.26 ± 0.16	0.39 ± 0.16	0.70 ± 0.18	0.80 ± 0.12	0.96 ± 0.06
4	0.00 ± 0.00	0.16 ± 0.11	0.31 ± 0.12	0.68 ± 0.15	0.77 ± 0.22	0.92 ± 0.11
6	0.00 ± 0.00	0.13 ± 0.09	0.25 ± 0.13	0.46 ± 0.20	0.75 ± 0.15	0.84 ± 0.10
8	0.00 ± 0.00	0.09 ± 0.13	0.24 ± 0.17	0.43 ± 0.18	0.68 ± 0.21	0.79 ± 0.14

Horizontal list label represent the phenylthiourea concentration and the vertical labels the length of alphanumerical memory load.

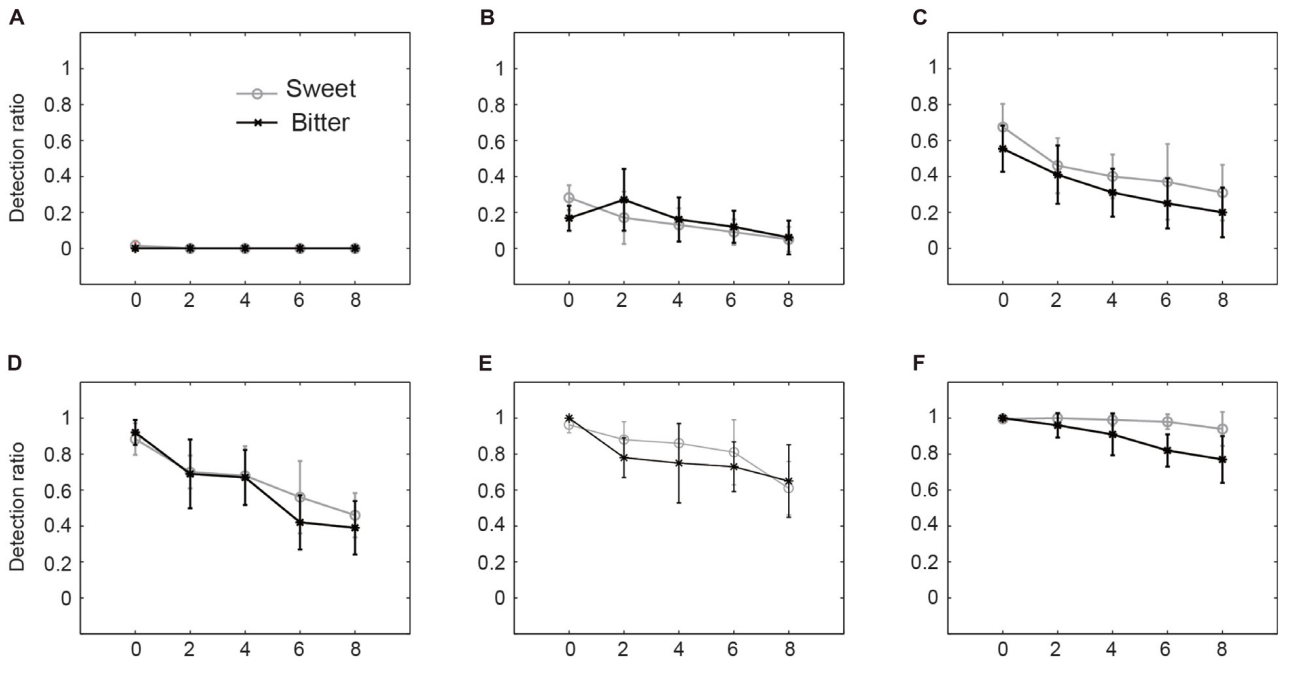

FIGURE 2 | Both sweet and bitter taste detections decrease with memory load. **(A–F)** Present the sweetness detection with sucrose concentration of 0, 1.5, 3.1, 3.9, 4.7, and 5.5 g/L, and bitterness detection with phenylthiourea concentration of 0, 0.02, 0.04, 0.08, 0.16, and 0.24 mM, respectively. The gray circles and black stars represent the sweet and bitter taste detection ratios averaged across all the subjects and the error bars denote the standard deviations across all the subjects. *Y*-axis represents the sweetness and bitterness detection ratio, from 0 to 1, 1 represents 100 percent detection. *X*-axis represents the length of the lists of alphanumerical items as the memory load during taste experiments, 0 represents the detection ratio without memory task.

load [**Figure 2C**, gray line, $F(4,125) = 1.774$, $p < 0.001$]. It is worth to note that **Figure 2D** represents the maximal decrease of both sweetness and bitterness taste detection ratio from around 90% (no memory load) to around 40% (maximal memory load) [sweetness: $F(4,125) = 22.058$, $p < 0.001$]. At highest taste concentration of our experiments (sugar 5.5 g/L), the sweetness detection was all close to 100% [**Figure 2F**, gray line, $F(4,125) = 4.664$, $p < 0.01$]. This may be explained

by the sweetness response saturation. However, the bitterness detection ratio still reduces when the memory load increases [**Figure 2F**, black line, $F(4,105) = 17.635$, $p < 0.001$]. It should be noted that the sweetness and bitterness concentrations are differently prepared. We would expect that at very high bitterness concentration, we might observe similar saturation effect as the sweetness perception.

DISCUSSION

This study has systematically tested the changes of sweetness and bitterness sensitivities with different degrees of memory load. The sweetness and bitterness sensitivities both decrease significantly when memory load increases. These results are consistent along the lines of the hypothesis, that cognitive status influences the taste perception. Our previous publications (Liang et al., 2013, 2016) have shown that the cognitive ease induced by visual inputs such as circular shapes, familiar words may enhance the sweetness sensitivity. The cognitive strain induced by angular shapes, unfamiliar words may reduce the sweetness perception. The cognitive ease or strain induced by the visual shapes or words in the previous studies is mostly referred to affective or hedonic cognition, i.e., cognitive positive and ease or cognitive negative and strain. The stimuli were passive and unfocused for the participants. Different from the previous paper, this study has applied the phonological loop as memory task, which request the participants to repetitively recall the visual alphanumerical information. These task leads active and focused attention of the participants and make them distracted from the taste itself, and thus reduces the taste perception. Moreover, our observation is consistent with other literature that the task load influences the sweet, sour, and salty perception, where much complex and higher concentration tastants have been studied (van der Wal and van Dillen, 2013). All these studies are consistent with the observation that task load, specifically negative or strain load, reduces the taste perception.

Different from the common food in the market with usually relatively high taste concentration, the low concentration of tastants are not much studied. This study has extended the previous literature finding and studied the sweet and bitter taste from sub-threshold to beyond threshold level. When the tastant concentration is extremely low (close to zero), there is no significant effect of memory load on taste perception (**Figure 2B**). When the tastants concentrations increase to around threshold level, the effect of memory load becomes more prominent (**Figure 2C**). As the tastant concentration increases and is more beyond the threshold level, we still observe the significant effect of memory load. Both sweetness and bitterness sensitivities decrease with increasing memory load (**Figures 2D,E**). It is worth to mention that the taste concentrations applied in our experiments are far below the market food or drinks (Huang et al., 2018). The reason is that we are focused to observe the taste perception around the taste threshold level, which has been suggested in our previous paper (Liang et al., 2013, 2016), the uncertain zone in our brain perception around threshold level is easier to be influenced by external stimuli. Marks and Wheeler (1998) have found that the thresholds are lower for attended tastants of sucrose and are critical than unattended ones. However, the attention in their study is limited to the expectation by giving subject the cue of tastants. In our study we extend the investigation to the more general memory load, such as food irrelevant memory load, and observe how are the sweet- and bitter- taste sensitivities influenced by the memory load.

Regarding the internal brain status induced in our experiments, the active and attentive process was generated by memory task, which would lead the subjects to cognitive strain. Under such cognitive state, the sensitivity of the subjects toward taste detection has been observed to be reduced. Higher the cognitive strain, lesser is the sensitivity toward the taste. Recently van der Wal and van Dillen (2013) have demonstrated similar observation that task load reduces sweet, salt, and sour taste perception. In their experiments, the cognitive load was to instruct the subjects to remember seven-digit number (high load) or one-digit number (low load). Such induced brain load is affective neutral and non-food related. Similarly, the task applied in our study was also without affective bias, helped us to observe the effect in a more general and systematic pattern.

Several previous studies have examined the effect of distracting stimuli on food choice (Shiv and Nowlis, 2004; Nowlis and Shiv, 2005). Earlier research shows due to cognitive load the taste perception reduces, and thus people tend to have more food to retain the same preferred taste levels in an attempt to preserve the enjoyment level of the food as compared to relaxed food intake conditions. On the other side, the cognitive status induced by emotion (negative or positive) may influence the taste perception as well. Noel and Dando (2015) found that sour taste was enhanced with negative emotion. The brain mechanisms of taste perception under different cognitive states remain unclear. When the attention of the subject is focused on taste pleasantness, the medial orbitofrontal and pregenual cingulate cortex are greater active than when attention is instructed to taste intensity. The taste detection in a tasteless solution involves insula and overlying operculum (Veldhuizen et al., 2007). Such finding might indirectly support our hypothesis that variant cognitive status affects taste perception differently.

Moreover, our data in this study may be explained from evolutionary biology (Wilson, 2014). We note that gustatory perception of taste sensitivity as that of bitterness is related to survival (Wooding, 2005) and has lower threshold (than sweetness for instance). When cognitive load is applied, cognitive processing takes precedence and gustatory perception sensitivity may reduce, even with bitterness sensitivity which is related to survival (Diamond, 1998). This might be explained by that cognitive beings survived due to cognitive calculations and cautionary steps taken rather than gustatory explorations. Hence when under cognitive stress the gustatory sensory processing manifest as bitterness taste sensitivity takes a backseat registering a decrease in bitterness taste sensitivity. It further leads us to another factor that influences gustatory perception, the affective component. We relate cognitive ease status to this affective component in the matter of gustatory perception. Previously, we have shown that visual information from external environment such as shapes influences taste sensitivities (Liang et al., 2013).

These experiments explored the affective aspect of gustatory taste perception. The circular shapes inducing cognitive ease (positive emotion) were shown to enhance the sweetness sensitivity. The sweetness detection in our previous experiment is more associated with affective positive and cognitive ease stimuli. The bitterness detection as described in this experiment is more associated with cognitive strain stimuli (Kahneman, 2011). The present experiment shows when under cognitive stress the gustatory sensory processing manifest as bitterness/sweetness taste sensitivity takes a backseat registering a decrease in bitterness taste sensitivity. Hence it is reasonable to speculate that the influence of cognitive over the affective in the context of gustatory sensory processing would be a weighted average of both affective and cognitive components. However, how affective and attentive cognitive components associate with each other and influence gustatory perception is not possible to shed light on here; although this may be an interesting topic for future studies with our calibrated model.

Here we infer that the different cognitive (for instance, attentive) states might be the key factor, which contribute to the modification of sweetness and bitterness perceptions. This study show that the memory load influences both sweet and bitter taste in a similar pattern (**Figures 2B–E**). Both taste detections reduce dramatically when the memory load increases. In our experiments, only the neutral cognitive load was applied, no affective component was induced, thus, the sweet and bitter taste sensitivities were influenced similarly. On the other hand, in our previous experiments, the visual inputs have both affective and cognitive components and where affective components were also allowed to exercise their influence alone by the choice of the sweet as the only tastants.

Although this study has extended the task stimuli from our previous studies, still the phonological loop as memory load here is a specific task, which involves silent repetition of verbal information coded from visual information (the list of items displayed on the monitor). The observation is limited to support completely the hypothesis of the study. To reach a more general cognitive load effect, it might be useful to design the articulation suppression paradigm with this material, or to use a non-verbal type of material in future. Moreover, regarding the potential difference of sweetness and bitterness perception from the evolutionary aspect, one should be aware that there may be an evolutionary twist to the human sugar intake. Fructose may be considered natural to us, and might be interesting to test the effect of fructose in our future experiments.

CONCLUSION

To conclude, the current work has demonstrated that the sweetness and bitterness detection ratios decrease with increasing memory load, especially around threshold concentration. At higher concentration, the both taste detection ratios are unaffected by the memory load. It is consistent with previous observation from other laboratories and extends our understanding to a more systematic pattern. Higher the memory load, lesser is the taste sensitivities. This study supports the hypothesis that the cognitive states (positive-ease or negative-strain) influences taste perception, and which of course still has a long way to go before we understand it completely. On a lighter note our work suggests that stress and enjoyment of food do not gel well with each other, and the cognitive process induced by different eating life style may modify the taste perception and lead to the acquired taste.

AUTHOR CONTRIBUTIONS

PL and SR designed the experiments and wrote the manuscript. JJ and QD performed the experiments. PL, JJ, and XT performed the data analysis.

FUNDING

The authors gratefully acknowledge the financial supports from NSFC (Grant No. NSF 61703058) and Jiangsu Province Funding 2017SJB1396 of China.

REFERENCES

Alsius, A., Navarra, J., Campbell, R., and Soto-Faraco, S. (2005). Audiovisual integration of speech falters under high attention demands. *Curr. Biol.* 15, 839–843. doi: 10.1016/j.cub.2005.03.046

Bertelson, P., Vroomen, J., De Gelder, B., and Driver, J. (2000). The ventriloquist effect does not depend on the direction of deliberate visual attention. *Percept. Psychophys.* 62, 321–332. doi: 10.3758/BF03205552

Diamond, J. M. (1998). Guns, germs, and steel. *Softw. Practiti.* 22, 829–831.

Elder, R. S., and Krishna, A. (2009). The effects of advertising copy on sensory thoughts and perceived taste. *J. Consum. Res.* 36, 748–756. doi: 10.1086/605327

Huang, X., Zhang, J., Li, J., Zhao, S., Xiao, Y., Huang, Y., et al. (2018). Daily intake of soft drinks and moderate-to-severe acne vulgaris in Chinese Adolescents. *J. Pediatr.* doi: 10.1016/j.jpeds.2018.08.034 [Epub ahead of print]. doi: 10.1016/j.jpeds.2018.08.034

Kahneman, D. (2011). *Thinking, Fast and Slow*. New York, NY: Farrar, Straus and Giroux.

Liang, P., Biswas, P., Vinnakota, S., Fu, L., Chen, M., Quan, Y., et al. (2016). Invariant effect of vision on taste across two asian cultures: India and China. *J. Sens. Stud.* 31, 416–422. doi: 10.1111/joss.12225

Liang, P., Roy, S., Chen, M.-L., and Zhang, G.-H. (2013). Visual influence of shapes and semantic familiarity on human sweet sensitivity. *Behav. Brain Res.* 253, 42–47. doi: 10.1016/j.bbr.2013.07.001

Marks, L. E., and Wheeler, M. E. (1998). Attention and the detectability of weak taste stimuli. *Chem. Senses* 23, 19–29. doi: 10.1093/chemse/23.1.19

Mozolic, J. L., Hugenschmidt, C. E., Peiffer, A. M., and Laurienti, P. J. (2008). Modality-specific selective attention attenuates multisensory integration. *Exp. Brain Res.* 184, 39–52. doi: 10.1007/s00221-007-1080-3

Noel, C., and Dando, R. (2015). The effect of emotional state on taste perception. *Appetite* 95, 89–95. doi: 10.1016/j.appet.2015.06.003

Nowlis, S. M., and Shiv, B. (2005). The influence of consumer distractions on the effectiveness of food-sampling programs. *J. Market. Res.* 42, 157–168. doi: 10.1509/jmkr.42.2.157.62287

Odegaard, B., Wozny, D. R., and Shams, L. (2016). The effects of selective and divided attention on sensory precision and integration. *Neurosci. Lett.* 614, 24–28. doi: 10.1016/j.neulet.2015.12.039

Pepino, M. Y., and Mennella, J. A. (2010). Effects of cigarette smoking and family history of alcoholism on sweet taste perception and food cravings in women. *Alcohol. Clin. Exp. Res.* 31, 1891–1899. doi: 10.1111/j.1530-0277.2007.00519.x

Shiv, B., and Nowlis, S. M. (2004). The effect of distractions while tasting a food sample: the interplay of informational and affective components in subsequent choice. *J. Consum. Res.* 31, 599–608. doi: 10.1086/425095

Shore, D. I., and Simic, N. (2005). Integration of visual and tactile stimuli: top-down influences require time. *Exp. Brain Res.* 166, 509–517. doi: 10.1007/s00221-005-2391-x

Spence, C., and Shankar, M. U. (2010). The influence of auditory cues on the perception of, and responses to, food and drink. *J. Sens. Stud.* 25, 406–430. doi: 10.1111/j.1745-459X.2009.00267.x

Stafford, L. D., Agobiani, E., and Fernandes, M. (2013). Perception of alcohol strength impaired by low and high volume distraction. *Food Q. Pref.* 28, 470–474. doi: 10.1016/j.foodqual.2012.12.005

Stafford, L. D., Fernandes, M., and Agobiani, E. (2012). Effects of noise and distraction on alcohol perception. *Food Q. Pref.* 24, 218–224. doi: 10.1016/j.foodqual.2011.10.012

Talsma, D., Doty, T. J., Strowd, R., and Woldorff, M. G. (2010). Attentional capacity for processing concurrent stimuli is larger across sensory modalities than within a modality. *Psychophysiology* 43, 541–549. doi: 10.1111/j.1469-8986.2006.00452.x

van der Wal, R. C., and van Dillen, L. F. (2013). Leaving a flat taste in your mouth: task load reduces taste perception. *Psychol. Sci.* 24, 1277–1284. doi: 10.1177/0956797612471953

Veldhuizen, M. G., Bender, G., Constable, R. T., and Small, D. M. (2007). Trying to detect taste in a tasteless solution: modulation of early gustatory cortex by attention to taste. *Chem. Senses* 32:569.

Wilson, E. O. (2014). *The Meaning of Human Existence*. New York, NY: Liveright.

Wooding, S. (2005). Evolution: a study in bad taste? *Curr. Biol.* 15, R805–R807.

Yantis, S. (2000). Goal-directed and stimulus-driven determinants of attentional control. *Attent. Perform.* 18, 73–103.

Conflict of Interest Statement: 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.

Copyright © 2018 Liang, Jiang, Ding, Tang and Roy. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 29 January 2019
doi: 10.3389/fnins.2018.01003

Goal-Directed Processing of Naturalistic Stimuli Modulates Large-Scale Functional Connectivity

Zhenfu Wen[1,2], *Tianyou Yu*[1,2], *Xinbin Yang*[3] *and Yuanqing Li*[1,2]*

[1] Center for Brain Computer Interfaces and Brain Information Processing, South China University of Technology, Guangzhou, China, [2] Guangzhou Key Laboratory of Brain Computer Interaction and Application, Guangzhou, China, [3] Department of Surgical Thoracic Oncology, Affiliated Cancer Hospital and Institute of Guangzhou Medical University, Guangzhou, China

Humans selectively process external information according to their internal goals. Previous studies have found that cortical activity and interactions between specific cortical areas such as frontal-parietal regions are modulated by behavioral goals. However, these results are largely based on simple stimuli and task rules in laboratory settings. Here, we investigated how top-down goals modulate whole-brain functional connectivity (FC) under naturalistic conditions. Analyses were conducted on a publicly available functional magnetic resonance imaging (fMRI) dataset (OpenfMRI database, accession number: ds000233) collected on twelve participants who made either behavioral or taxonomic judgments of behaving animals containing in naturalistic video clips. The task-evoked FC patterns of the participants were extracted using a novel inter-subject functional correlation (ISFC) method that increases the signal-to-noise ratio for detecting task-induced inter-regional correlation compared with standard FC analysis. Using multivariate pattern analysis (MVPA) methods, we successfully predicted the task goals of the participants with ISFC patterns but not with standard FC patterns, suggests that the ISFC method may be an efficient tool for exploring subtle network differences between brain states. We further examined the predictive power of several canonical brain networks and found that many within-network and across-network ISFC measures supported task goals classification. Our findings suggest that goal-directed processing of naturalistic stimuli systematically modulates large-scale brain networks but is not limited to the local neural activity or connectivity of specific regions.

Keywords: top-down goals, naturalistic condition, inter-subject functional correlation, multivariate pattern analysis, large-scale brain networks

OPEN ACCESS

Edited by:
Andreas Bahmer, Universitätsklinikum Würzburg, Germany

Reviewed by:
Wei Liao, University of Electronic Science and Technology of China, China
Ruiwang Huang, South China Normal University, China

***Correspondence:**
Yuanqing Li
auyqli@scut.edu.cn

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *20 August 2018*
Accepted: *12 December 2018*
Published: *29 January 2019*

Citation:
Wen Z, Yu T, Yang X and Li Y (2019) Goal-Directed Processing of Naturalistic Stimuli Modulates Large-Scale Functional Connectivity. Front. Neurosci. 12:1003. doi: 10.3389/fnins.2018.01003

1. INTRODUCTION

Selective processing of information according to behavioral goals is crucial for our interaction with the complex environment. However, the organizational basis underlying this goal-directed behavior is unclear. Electrophysiological and functional imaging studies have suggested that task goals modulate the neural representation of a stimulus (Mirabella et al., 2007; Ptak and Schnider, 2010; Gilbert and Li, 2013). Recently, using powerful multivariate pattern analysis (MVPA) methods (Norman et al., 2006; Haxby et al., 2014), many studies have found that top-down behavioral goals can be decoded with distributed activities across frontoparietal and sensory regions (Chiu et al., 2011; Waskom et al., 2014; Loose et al., 2017; Long and Kuhl, 2018).

Despite numerous studies, most neural investigations of goal-directed behavior have employed simple stimuli, such as moving dots, static faces and object images, which are functionally well-characterized (Nastase et al., 2017). However, many common perceptual tasks require combining low-level features of stimuli to represent abstract semantic information in our brains. With the recent observation that less-controlled naturalistic stimuli such as movies evoke reliable neural responses across individuals (Hasson et al., 2004, 2010; Simony et al., 2016), a few studies have used natural paradigms to investigate how task contexts modulate the neural representation of high-level visual and semantic information (Cukur et al., 2013; Nastase et al., 2017, 2018). These pioneering studies have found that goal-directed processing of different objects or semantic features in natural movies modulates distributed cortical areas. However, these studies mainly focused on the modulation effects of behavioral goals on neural activity in certain brain regions but ignored the interaction between distributed cortical areas, which are increasingly recognized as the biological basis for cognition and behaviors (Fox et al., 2005; Sporns, 2014; Mišić and Sporns, 2016).

Evidence from electrophysiology studies indicates the important role of neuronal synchronization in goal-directed behavior (Von Stein and Sarnthein, 2000; Engel et al., 2001; Womelsdorf et al., 2007). Similar findings have also been reported in recent functional magnetic resonance imaging (fMRI) studies (Spreng et al., 2010; Al-Aidroos et al., 2012). For example, activity in V4 is correlated more strongly with activity in the fusiform face area in a face attention task and with activity in the parahippocampal place area in a scene attention task (Al-Aidroos et al., 2012). Notably, most of these studies used rudimentary visual stimuli and primarily focused on connections between a limited number of brain regions that were selected based on prior anatomical knowledge or on activation patterns during tasks. These preselecting methods may lead to some regions being ignored, given that goal-directed processing of external information recruits a wide variety of brain regions (Corbetta and Shulman, 2002; Petersen and Posner, 2012; Vaziripashkam and Xu, 2017). Accordingly, examining whole-brain functional connectivity (FC) may provide new insights into top-down goals representation.

A standard method to characterize whole-brain FC is to calculate a Pearson correlation between the time series of all pairs of regions within each subject (van den Heuvel and Hulshoff Pol, 2010; Zalesky et al., 2012). However, a potential limitation of this method is that the calculated FC measures consist of task-evoked correlations, within-subject intrinsic neural fluctuations, and non-neuronal artifacts, and these types of signals cannot be reliably separated (Hasson et al., 2004; Simony et al., 2016). Given that the FC structure during task performance has been shown to be highly correlated with the intrinsic FC structure (Cole et al., 2014), it would be difficult to reliably detect differences in FC patterns across task contexts. Recently, a novel method termed inter-subject functional correlation (ISFC) has been proposed (Simony et al., 2016). By calculating inter-regional correlations between subjects that are performing the same task, the ISFC method increases the signal-to-noise ratio (SNR) for detecting task-evoked FC, making it an effective method for examining subtle differences between cognitive states (Simony et al., 2016; Rosenthal et al., 2017).

In the present study, we applied the ISFC method to a publicly available dataset to investigate how behavioral goals modulate whole-brain FC. Dynamic video clips of animals behaving in natural environments were used as stimuli. During the fMRI experiment, participants were required to made either behavioral or taxonomic judgments when exposed to identical naturalistic video clips. We used MVPA methods to explore task modulation of whole-brain FC. We show that ISFC patterns support successful task classification and that task goals modulate connections between large-scale brain regions that can be assigned to a variety of canonical functional networks.

2. MATERIALS AND METHODS

2.1. Subjects

A publicly available dataset was used in this study (Nastase et al., 2017, 2018). This dataset was obtained from the OpenfMRI database (http://www.openfmri.org), and the accession number was ds000233. A total of 12 right-handed healthy adults (7 females; mean age $= 25.4 \pm 2.6$ SD years) provided informed consent and participated in the main experiment. The study was approved by the Institutional Review Board of Dartmouth College.

2.2. Experimental Design

The experimental paradigm was described clearly in the original paper of Nastase et al. (2017, 2018). We briefly describe the most relevant aspects of the experimental design here for completeness. A total of 80 naturalistic clips of behaving animals (each lasting 2 s), collected from the Internet, were used in the experiment. Semantically, these clips could be partitioned into five groups based on taxonomic categories (primates, ungulates, birds, reptiles, and insects) or four groups based on behavioral categories (eating, fighting, running, and swimming). Each participant completed 10 experimental runs (each lasting 392 s) while viewing these clips under two task contexts. In half of the runs, participants were instructed to pay attention to taxonomy types in the presented clips (taxonomy task runs), and in the other half of the runs, participants were instructed to pay attention to the behavioral types of the stimuli (behavior task runs). These 5 taxonomic attention runs and 5 behavior attention runs were presented in a counterbalanced order across participants. Note that the appearance order of movie clips in each experimental run of each subject was randomized. Therefore, the appearance orders of clips in the two tasks were irregular, it is unlikely that the following MVPA results were contributed by differences of stimulus sequences between the two tasks.

In the taxonomy task runs, participants were asked to press a button if two sequential clips contained the same taxonomic category. In the behavior task runs, participants were asked to press a button if two sequential trials contained the same behavioral category. There were 4 repetition trials in each run that required a response. These tasks required participants to attend to

the taxonomic or behavioral features of clips in a corresponding task context.

2.3. Image Acquisition

Functional and structural images were acquired on a 3 T Philips Intera Achieva MRI scanner with a 32-channel head coil. Functional images were obtained using single-shot gradient-echo echo-planar imaging with a SENSE reduction factor of 2 (TR/TE = 2,000/35 ms, flip angle = 90°, resolution = 3 mm isotropic, matrix size = 80 × 80, FOV = 240 × 240 mm^2, 42 transverse slices in an interleaved fashion). Each participant completed 10 experimental runs in a scanning session, with an additional structural scan obtained at the end of the session using a high-resolution T1-weighted 3D turbofield echo sequence (TR/TE = 8.2/3.7 ms, flip angle = 8°, resolution = 0.938 × 0.938 × 1.0 mm^3, matrix size = 256 × 256 × 220, FOV = 240 × 240 × 220 mm^3).

2.4. Image Preprocessing

Imaging data were preprocessed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm) and DPARSF (Chao-Gan and Yu-Feng, 2010). Functional data were slice-time adjusted, motion-corrected, and normalized to the Montreal Neurological Institute (MNI) space using a segmented high-resolution gray matter structural image and a gray matter template. The resulting images were detrended to abandon linear trends. The nuisance time series, including motion, white matter, CSF and their derivatives, were regressed out using linear regressions. Low-frequency signals were removed using a high-pass filter (> 0.08 Hz). We did not use a low-pass filter, as in resting-state fMRI studies, as this allowed us to retain potentially informative task signals at higher frequencies (Shirer et al., 2012; Cole et al., 2013). Signals corresponding to stimulus presentation were further removed using standard general linear regression models of task events (Cao et al., 2014; Cole et al., 2014). Specifically, task events were modeled by convolving stimulus onsets with the standard hemodynamic response function. These regressors were then regressed out from voxel activities. The resultant residual time series were used for the following functional network analyses.

2.5. Definition of Nodes

A 264-node brain atlas was used for FC analysis. This atlas was derived from both resting and task FC meta-analyses (Power et al., 2011) and has been widely used in network analyses (Vatansever et al., 2015; Schultz and Cole, 2016). Each of the 264 nodes was assigned to one of the 14 subnetworks in the original publication (Cole et al., 2013). Among these 14 subnetworks, we focused on 10 well-established subnetworks, including the frontoparietal, cingulo-opercular, salience, dorsal attention, ventral attention, default mode, somatomotor (hand and mouth), auditory, visual, and subcortical networks. The other three networks, including the cerebellum network, the memory retrieval network, and a network of uncertain function, were also involved in our analyses, but they were treated as a single subnetwork (the others network) for convenience. Therefore, the 264 nodes were assigned to 11 subnetworks in this study. The nodal-mean time series were extracted by averaging the time series over all voxels in each of the 264 nodes, resulting in a neural signal matrix X, which has the form of a $P \times N$ matrix containing time series from P nodes over N time points. The neural signal matrix of each subject and each experimental run was used for the following network constructions.

2.6. Inter-subject Functional Correlation

We used the recently proposed ISFC to assess task-evoked FC (Simony et al., 2016). The ISFC method effectively eliminates intrinsic signals by calculating the inter-regional correlations between different subjects who perform the same task. Assuming we have a neural signal matrix X_k for each subject $k, k = 1, ..., K$ with each regional time series normalized to a zero mean and unit variance. In contrast to the standard FC measure, which is calculated within each neural signal matrix, the ISFC of subject k is defined as the Pearson correlation between this subject and the average of all other subjects:

$$\hat{C_k} = \frac{1}{N} X_k [\frac{1}{K-1} \sum_{q \neq k} X_q^T] \quad (1)$$

which is a $P \times P$ correlation matrix where each element (i, j) represents a correlation between node i of subject k and the mean series of node j of the other subjects. To increase the normality of the distribution of correlation values, each correlation coefficient was converted to a z-score using Fisher's r-to-z transformation. To further impose symmetry, the final ISFC matrix of subject k was given by $(\hat{C_k} + \hat{C_k}^T)/2$. The group-based ISFC matrix was calculated by averaging the ISFC matrixes across subjects:

$$\hat{C} = \frac{1}{K} \sum_k \hat{C_k} \quad (2)$$

2.7. Similarity Analysis

A key question of this study was whether goal-directed visual processing modulates whole-brain FC. Conceptually, if the task goal modulates ISFC, similarities between ISFC patterns from the same task should be higher than those from different tasks. To confirm this hypothesis, we performed a similarity analysis as follows. First, the neural signal matrices of each subject were averaged across the 5 behavioral task runs and the 5 taxonomic task runs, resulting in two neural signal matrices (one for the behavioral task and one for the taxonomic task) for ISFC pattern estimations. Then, the 12 subjects were randomly split into two independent groups of 6 subjects, and group-based ISFC matrices were calculated for each group and each task according to Equation (2). Next, we calculated the between-task ISFC similarity and within-task ISFC similarity across the two groups. Specifically, the between-task ISFC similarity was defined as the spatial Pearson correlation between the ISFC matrices from different groups and different tasks (e.g., group 1 task 1 vs. group 2 task 2). The within-task ISFC similarity was defined as the spatial Pearson correlation between the ISFC matrices from different groups and the same task (e.g., group 1 task 1 vs. group 2 task 1). We repeated this procedure 462 times (all possible situations with the 12 subjects divided into two

groups of 6 subjects) and compared the mean within-task ISFC similarities and mean between-task ISFC similarities across all situations.

2.8. ISFC Classification of Attention Task

We further used MVPA methods to examine whether the task goals of subjects could be predicted using whole-brain ISFC patterns. A leave-one-subject-out-cross-validation (LOSOCV) procedure was employed to assess the classification performance. In each iteration of the LOSOCV, we left out the data of one subject as the test set and use the data of the other subjects as the training set. A template-matching method was used for task label predictions (Simony et al., 2016). Similar to the similarity analysis, neural signal matrices were first averaged across tasks for each subject within the training set. Then, the group-based ISFC matrices were calculated according to Equation (2) for each attention task based on the corresponding neural signal matrices. Therefore, based on the training data, we obtained one ISFC matrix, C_{beh}, for the behavioral task and one ISFC matrix, C_{tax}, for the taxonomic task. These two matrices were used as ISFC templates for the two attention tasks. Note that the test data were never used to derive the ISFC templates.

For each run r of the left-out subject, we attempted to predict its task label by comparing its ISFC matrix with the two ISFC templates. We calculated an ISFC matrix, $C_{beh,r}$, between the run r and the average neural signal matrix corresponding to the behavior task from the training set. Similarly, we also obtained an ISFC matrix, $C_{tax,r}$, for the taxonomic task. The predicted label for this run was then given by the attention task $m \in \{beh, tax\}$ that maximized the Pearson correlation between $C_{m,r}$ and the templates C_m:

$$\hat{m}_r = \underset{m\in\{beh,tax\}}{\arg\min} \{Corr(C_{m,r}, C_m)\} \tag{3}$$

This procedure was repeated for each subject and each run, and the classification accuracy was then computed as the proportion of times that an experimental run was assigned to the correct task context.

2.9. FC Classification of Attention Task

For comparison, we also used standard FC to classify task goals. This procedure was similar as the ISFC classification procedure described above except the FC matrixes were calculated within subjects. We obtained a correlation matrix for each subject and each experimental run by calculating the Pearson correlation coefficient between every pair of nodes. The correlation matrixes of the same task were further averaged within each subject. Averaging correlation matrixes of each task increased the signal-to-noise ratio (SNR) of estimated FC templates. Previous MVPA studies have suggested that this average step often improve classification performance to some degree (Isik et al., 2013; Hebart et al., 2018). For the employed template-matching method in this study, the testing sample was assigned the label of the FC template with which it is maximally correlated. Therefore, we would expect a better classification performance by increasing the SNR of estimated FC templates. Then, using training data, we obtained two FC templates C_{beh} and C_{tax} by averaging the correlation matrixes of the behavioral tasks and taxonomic tasks across subjects, respectively. For a run r of the test subject, we obtained its within-subject FC matrix C_r, and the label of this FC matrix was predicted as the task that maximized the Pearson correlation between C_r and the templates C_m:

$$\hat{m}_r = \underset{m\in\{beh,tax\}}{\arg\min} \{Corr(C_r, C_m)\}. \tag{4}$$

2.10. Identifying Discriminative Connections

Connections contribute differently to classification. To determine discriminative connections that contributed more to task classification, we performed an edge-based analysis similar to a previous study (Finn et al., 2015). Computationally, the Pearson correlation of two normalized vectors (zero mean, unit variance) was calculated as the sum of the element-wise products. Thus, an element with a large positive product contributes more to the correlation coefficient. In this classification procedure, we calculated Pearson correlation coefficients between the ISFC matrix derived from the test data and the templates derived from the training data, and the task label was chosen as the one that resulted in the largest correlation coefficient. Conceptually, the product of a discriminative connection should be large when the test data and the template are from the same task. In contrast, the product should be small when the test data and the template are from different tasks. Therefore, given a test ISFC matrix $C_{m,r}$ for run r and templates C_m, $m \in \{beh, tax\}$, we defined the discriminative measure of edge e as:

$$\phi_{m,r}(e) = Comp(C_m(e) * C_{m,r}(e), C_u(e) * C_{m,r}(e))\ u \in \{beh, tax\}, u \neq m \tag{5}$$

where $Comp(a, b) = 1$ if $a > b$, otherwise $Comp(a, b) = 0$. The first term of the function, $Comp()$, is the within-task edge-wise product, and the second term is the between-task edge-wise product. The discriminative measures were then averaged across all iterations of the LOSOCV to obtain a single value, $\phi(e)$, for each connection e. A connection with a large $\phi(e)$ is thought to be discriminative.

2.11. Subnetwork-Based Classification

To further assess the classification ability of individual canonical subnetworks, we conducted the same LOSOCV procedure as described above. However, this time, only within-network connections calculated between regions from a specific subnetwork or across-network connections calculated between regions from two different canonical subnetworks were used for task classification.

2.12. Effects of Scan Length on Classification

Since an experimental run consisted of a relatively long time series (196 time points), we further explored whether task goals could be predicted using fewer time points. We varied the number of time points n that were used to calculate the ISFC measures between 20 and 180 in increments of 10. Following previous studies (Finn et al., 2015; Greene et al., 2018), for each number of time points n, we randomly chose the start time point, and then extracted n continues time points beginning

with that starting point to calculate the ISFC measures for task classification. We did not extract time points randomly from the whole series, because this strategy may ignore the temporal autocorrelation of fMRI time series (Woolrich et al., 2001) and thus bias the estimation of functional connectivity. This procedure was repeated 10 times, and the mean classification accuracies for these times were obtained.

2.13. Effects of Atlas on Classification

To test whether the task classification accuracy based on ISFC measures was sensitive to the specific choice of atlas and network assignments, we conducted the aforementioned classification analyses using an additional 268-node atlas provided by Shen et al. (2010). This atlas functionally divides the brain into 268 regions by maximizing the similarities of the voxel-wise time series within each node and assigns each region to one of the following subnetworks: subcortical-cerebellum, frontoparietal, default mode, medial frontal, motor, visual 1, visual 2, or visual association.

2.14. Statistical Analysis

We used a non-parametric permutation test (Nichols and Holmes, 2002) to assess whether the difference between the mean within-task similarity and the mean between-task similarity was significant. We first combined the 462 within-task similarities and the 462 between-task similarities into one group. Then, we randomly split the group into two equal-size groups and calculated the difference between the means of these two groups. We repeated this procedure 1,000 times and obtained a null distribution of the differences. The p-value was then calculated as the number of null-hypothesis differences that were equal to or greater than the observed true difference divided by 1,000. With this approach, the smallest p-value that can be reported is $1/1,000 = 0.001$.

A non-parametric permutation test was also used to assess the statistical significance of the task classification accuracy. In brief, we first shuffled the labels of all the experimental runs and then performed the aforementioned LOSOCV procedure for ISFC features to obtain a classification accuracy. We repeated this procedure 1,000 times, resulting in a null distribution of accuracies. The p-values were calculated as described above.

For subnetwork-based classifications, to control for the presence of multiple comparisons (Nichols and Holmes, 2002), we obtained the maximum classification accuracy across all subnetworks in each iteration of the permutation. These maximum values were used to construct the null distribution of accuracies. Similarly, for classifications based on variations in scan length, the maximum accuracies for each assessed time point were used to construct a null distribution.

To determine the discriminative connections, we obtained the maximum discriminative measure for all connections at each iteration of the permutation. The null distribution was constructed using these maximum discriminative measures. Connections with true discriminative measures larger than the 95*th* percentile of the null distribution were considered to be discriminative (i.e., $p < 0.05$).

3. RESULTS

3.1. Behavioral Results

As stated in the original publication of the dataset (Nastase et al., 2017), participants performed very well in both the behavior task (mean accuracy: 0.994, SD: 0.005) and the taxonomy task (mean accuracy: 0.993, SD: 0.005). A paired t-test revealed no significant task-related difference in detection accuracy [$t_{(11)} = 0.469, p = 0.91$]. In addition, response times were also not significantly different between the two conditions [paired t-test: $t_{(11)} = 0.015, p = 0.99$), though the small number of response trials might hinder a robust estimation of response times. Therefore, it is unlikely that the subsequent classification analyses were influenced by differential behavioral responses.

3.2. Modulation of Whole-Brain FC

We used a similarity analysis to examine the top-down modulation of whole-brain FC. As shown in **Figure 1A**, when the ISFC method was used to extract task-evoked FC patterns, the within-task similarities were larger than the between-task similarities, with most of the data points falling below the diagonal. This difference was significant, as indicated by a permutation test ($p = 0.001$). This finding indicates that whole-brain FC was modulated when humans processed the same stimulus with different task goals. In contrast, the standard FC method resulted in very similar values for within-task similarity and between-task similarity (**Figure 1B**), with no significant difference observed between the two types of measures ($p = 0.19$). This result is consistent with recent findings that connections observed during different tasks are highly correlated (Cole et al., 2014). The successful detection of attentional modulation of FC using the ISFC method may be attributed to the effectiveness of the ISFC method in eliminating intrinsic signals (Simony et al., 2016; Kim et al., 2017).

3.3. Classification of Task Contexts

Having confirmed that task contexts modulated whole-brain FC, we further applied MVPA methods to investigate the possibility of task goals prediction. We used the ISFC method to extract task-evoked FC for each of the tasks and used a template-matching method to predict the task label of each experimental run of a left-out subject. The LOSOCV procedure showed an accuracy of 90% (**Figure 2**), which was significantly higher than chance (50%), as indicated by a permutation test ($p = 0.001$). We also performed the same LOSOCV procedure using standard FC patterns (**Figure 2A**). In this case, the accuracy was much lower and did not reach significance (54.17%, p = 0.13). Together with the similarity analyses, these results suggest that the ISFC method is powerful in detecting subtle differences between cognitive states. Since the classification accuracy was much higher with ISFC patterns than with standard FC patterns, we focused on ISFC patterns in the following analyses.

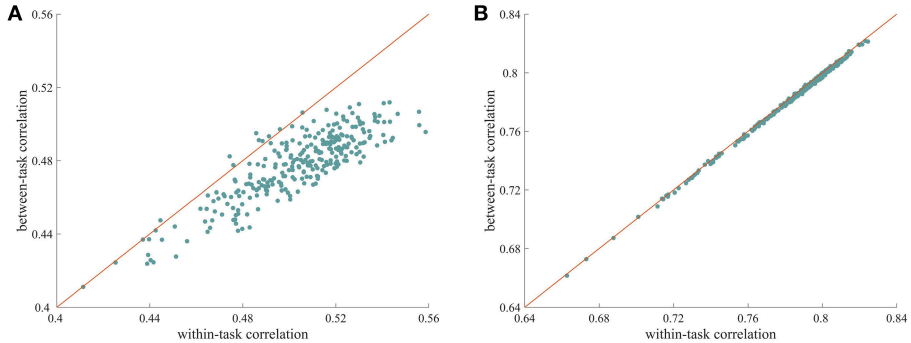

FIGURE 1 | Modulation of whole-brain FC under different tasks. **(A)** Scatter plot of within-task correlation values vs. between-task correlation values calculated using ISFC patterns. The within-task correlation value was significantly higher than the between-task correlation value ($p = 0.001$, permutation test), indicating that ISFC patterns are more similar when subjects perform the same task than when they perform different tasks. **(B)** Scatter plot of within-task correlation values vs. between-task correlation values calculated using standard FC patterns. The two types of measures were highly correlated, with no significant differences observed ($p = 0.63$), suggesting that standard FC patterns are less sensitive at detecting modulated connections. Each data point corresponds to a random partition of the subjects.

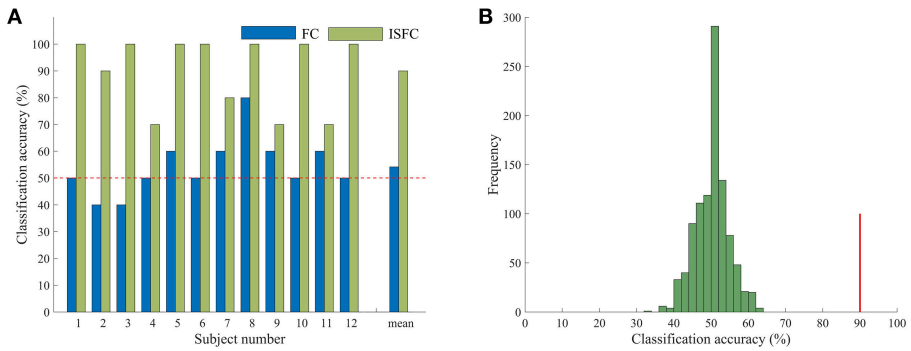

FIGURE 2 | Accuracy in task classification. **(A)** Accuracy of each single subject and mean accuracy across the LOSOCV using ISFC and FC patterns. The horizontal dashed line shows the chance level (50%). **(B)** Permutation test of accuracy when ISFC patterns were used for classification. The histogram shows the null distribution of accuracy values when task labels were randomly permuted, and the solid red line indicates the accuracy obtained for the true task labels. The classification accuracy (90%) was significantly higher than the chance level ($p = 0.001$).

3.4. Connections Contribution to Classification

The discriminative connections that largely contributed to the classification were determined using edge-wise analysis. Specifically, for each edge, a mean discriminative measure was calculated across the LOSOCV procedure and compared to a null distribution constructed from 1,000 random permutations. We then identified discriminative edges as those that possessed discriminative measures larger than the 95th percentile of the null distribution. We found 383 discriminative connections among all 34980 possible connections. These discriminative connections are displayed in a circle plot (**Figure 3**) and projected to a surface rendering of a human brain (**Figure 4**) using the BrainNet viewer (Xia et al., 2013). The majority of the discriminative connections are within the visual network and between the visual network and other networks, mainly including the dorsal attention, somatomotor, and default mode networks.

3.5. Subnetwork-Based Classification

We have shown that the top-down behavior goals could be reliably classified using whole-brain ISFC patterns and found that discriminative connections were distributed across several networks. However, whether a specific subnetwork (e.g., the dorsal attention network) supports task classification was not clear. To explore this possibility, we performed classification analyses using the within-network ISFC measures from each of the 11 canonical networks separately. As shown in **Figure 5**, the classification accuracies of the somatomotor (74.16%, $p = 0.001$), cingulo-opercular (67.50%, $p = 0.003$), visual (89.170%, $p = 0.001$), frontoparietal (68.33%, $p = 0.001$), salience (70.83%, $p = 0.001$), dorsal attention (78.33%, $p = 0.001$) and the others network (75.83%, $p = 0.001$) were significantly higher than chance (50%).

We also tested whether the across-network ISFC measures between two subnetworks (e.g., connections between the visual and dorsal attention networks) would support the classification

(**Figure 5**). This analysis revealed many discriminative across-network measures, mainly related to the visual, dorsal attention, frontoparietal, and somatomotor networks. To further assess the classification ability of each specific network, we averaged the accuracies obtained using across-network ISFC measures between each network and every other network. As shown in **Figure 5**, all across-network measures showed significant accuracies. These results suggest that the modulated connections are distributed across the brain and not limited to specific subnetworks.

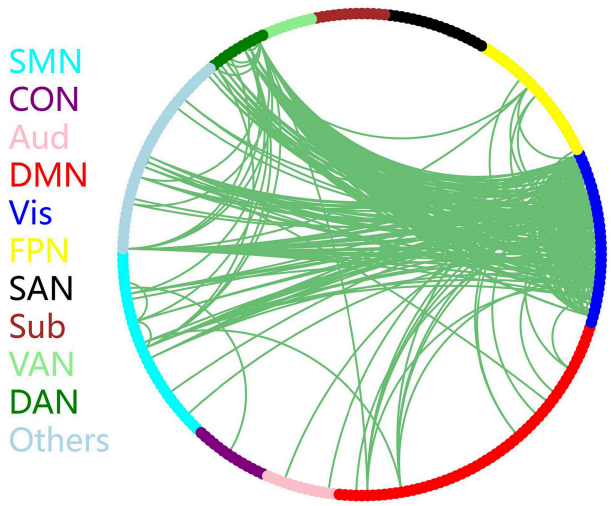

FIGURE 3 | Distribution of discriminative connections identified by edge-wise analysis. Brain regions are arranged and color-coded according to 11 canonical subnetworks: frontoparietal (FPN), cingulo-opercular (CON), salience (SAN), dorsal attention (DAN), ventral attention (VAN), default mode (DMN), somatomotor (SMN), auditory (Aud), visual (Vis), subcortical (Sub), and a network with other regions (Others).

3.6. Effects of Scan Length on Classification

To explore how the number of time points used for ISFC estimation influenced the classification accuracy, we performed the classification with ISFC calculated using a varying number of time points between 10 and 180. We observed accuracies ranging from 57.94 to 88.33%, with higher accuracies obtained using larger numbers of time points (**Figure 6**). Permutation testing revealed that the accuracies were significantly higher than chance with scan lengths as short as 20 time points (40 s), suggesting that attentional modulation of ISFC can be reliably detected using relatively short scan lengths.

3.7. Effects of Atlas on Classification

We repeated the classification analyses using ISFC measures calculated from a 268-node atlas that divides the human brain into 8 networks. As expected, the subnetwork-based analysis revealed a large number of within-network and across-network ISFC measures that were discriminative in task classification (**Figure 5B**). In addition, the average classification accuracies found using across-network ISFC measures between a certain subnetwork and every other subnetwork were significantly higher than chance. For classifications using shorter time courses, the use of this 268-node atlas led to similar accuracies as the 264-node atlas, with accuracies ranging from 59.39 to 81.11% (**Figure 6**). These results suggest that our main findings have robust reproducibility.

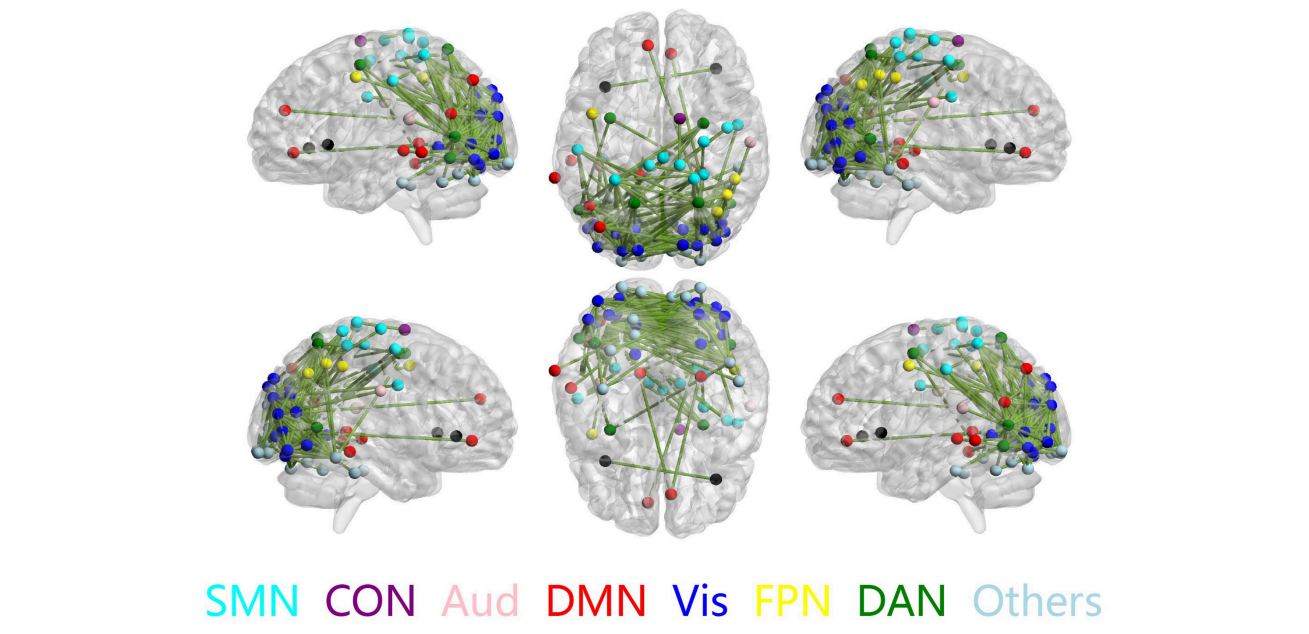

FIGURE 4 | Discriminative connections shown in sagittal (left/right), axial (top/bottom), and coronal (front/back) views. Nodes indicate brain regions, and edges represent connections between regions. Only regions that formed discriminative connections are shown.

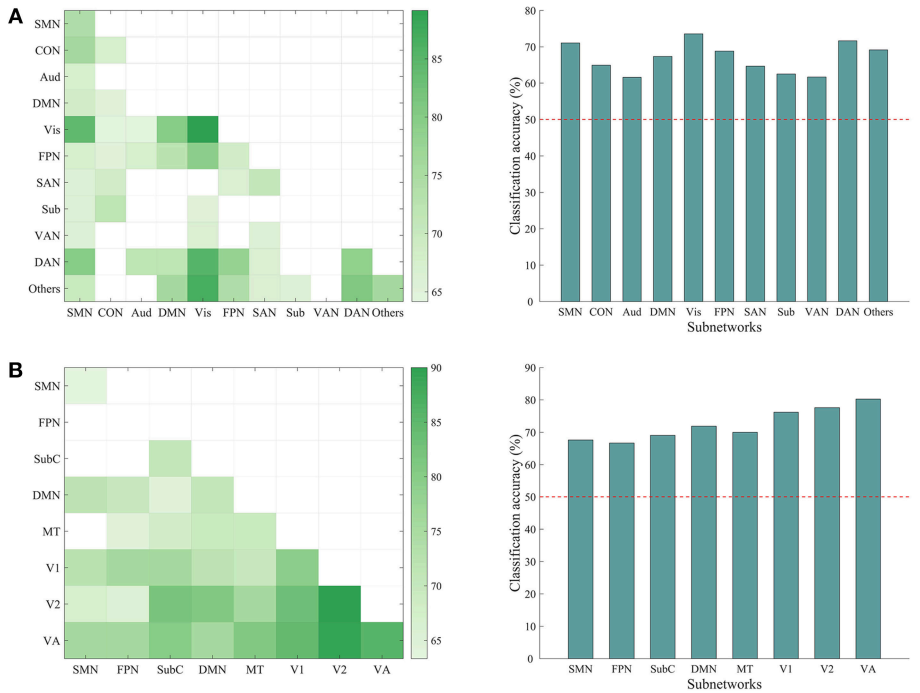

FIGURE 5 | Classification accuracy using ISFC patterns of subnetworks. **(A)** Left: matrix representing the classification accuracy based on ISFC measures within and between canonical subnetworks. Rows and columns represent predefined subnetworks based on the 264-node atlas. Only accuracies significantly higher than the chance level (50%) are shown in the matrices ($p < 0.05$, corrected). The abbreviations for each subnetwork are the same as in **Figure 3**. Right: average accuracy based on across-network ISFC measures between a specific subnetwork on the x-axis and all the other subnetworks. The horizontal dashed line shows the chance level. Each of the subnetworks reached significance (all $p = 0.001$, corrected). **(B)** Similar information as presented in A, but with a different 268-node atlas that contained 8 subnetworks: medial frontal (SMN), frontoparietal (FPN), default mode (DMN), subcortical-cerebellum (SubC), motor (MT), visual 1 (V1), visual 2 (V2), and visual association (VA).

3.8. Effects of Head Motion

Previous studies have suggested that head motion can influence the estimation of FC (Power et al., 2012; Van Dijk et al., 2012). However, motion artifacts are unlikely to have contributed to the observed successful task context classification using ISFC, given that the classification accuracy determined using classic FC measures was much lower than that of the ISFC measures. To remain conservative, however, we further examined the motion estimates for the two attention tasks. Using the motion parameters generated during the motion correction procedure during preprocessing, we calculated the average frame-to-frame motion for each experimental run and each subject (Power et al., 2012). This process resulted in a total of 60 (12 subjects, each performed 5 runs per attention task) values for each task. These values were assessed using paired t-tests to compare the head motion difference between the two attention tasks. We found that the difference between the two attention tasks was not significant ($t_{(59)} = 1.568, p = 0.12$). Therefore, the successful classification of task goals is unlikely to be based on motion artifacts.

4. DISCUSSION

In this study, we investigated how behavioral goals modulate the whole-brain FC of subjects. In contrast to previous goal representation studies that used rudimentary stimuli designed for a laboratory setting, dynamic, complex naturalistic stimuli that conveyed rich information were used in this experiment. This naturalistic paradigm suitably mimicked goal-directed behavior in a real-life context. Given that top-down information processing recruits wide swaths of brain regions, we examined large-scale FC across the whole brain. We employed a novel ISFC method to isolate task-evoked FC from intrinsic neural fluctuations and non-neuronal artifacts. We first conducted a similarity analysis and showed that ISFC patterns were more efficient in representing specific task context than standard FC patterns. We then employed MVPA methods to examine whether attention tasks could be predicted from the corresponding ISFC patterns. We found that ISFC reliably distinguished one attention task from another with a high classification accuracy, even with relatively short scan lengths. We further identified many within-network and across-network ISFC measures that enabled task classification, suggesting a global modulation of connectivity patterns by task contexts.

Multivariate approaches ensure high sensitivity to fine-grained discriminative patterns (Norman et al., 2006; Zeng et al., 2012; Haxby et al., 2014), and recent MVPA research investigating task representation has shown that distributed patterns of activity in the parietal, medial and lateral prefrontal cortex (PFC) represent top-down tasks (goals) (Chiu et al., 2011;

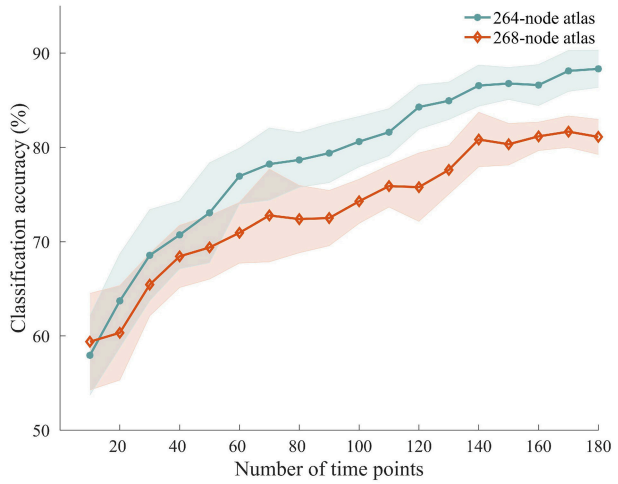

FIGURE 6 | Classification accuracy using shorter time series. ISFC patterns were calculated with shorter time series and processed through the same classification procedure. Solid lines and shaded areas indicate the mean accuracies and SEMs across the 10 randomizations, respectively. Two different atlases were used for this analysis.

Waskom et al., 2014; Long and Kuhl, 2018). However, these findings were largely based on simple stimuli and task rules and overlooked the interactions between regions. Our results extend these studies by showing that selective processing of complex visual information conveyed by naturalistic stimuli modulated large-scale brain networks and that this modulation contained highly predictive information on the task contexts. We employed a LOSOCV procedure to estimate task classification performance. This across-subject MVPA can be challenging, given high levels of interindividual variability (Finn et al., 2015). For example, in a previous study with a LOSOCV setting, standard FC patterns successfully predicted which task a subject was performing, but the classification accuracies were relatively low (Cole et al., 2013). In the present study, the accuracy obtained using standard FC patterns was not significant, which is consistent with the similarity analysis showing that the standard FC patterns of the two tasks were highly correlated. This negative result may be attributed to the complex information contained in the naturalistic stimuli, which drives complex neural responses and thus hinders the detection of subtle differences between two tasks. On the other hand, the modulation of FC measures was possibly overwhelmed by intrinsic FC patterns, as previous studies have found that resting-state FC largely matches the FC during task performance (Cole et al., 2014; Kim et al., 2017). In contrast, we obtained a high level of accuracy when using ISFC patterns for the classification, and the accuracies remained significant even when very short time courses were used for ISFC pattern estimation. By repeating the MVPA procedure using another brain atlas, we have also shown that the performance was not specific to the choice of atlas. Along with previous ISFC studies (Simony et al., 2016; Kim et al., 2017; Rosenthal et al., 2017), the current successful classification of attention tasks shows promise for the utilization of the ISFC method in other contexts to investigate subtle differences between task-evoked FC patterns.

The edge-based analysis and subnetwork-based analysis found that ISFC measures within the visual network and between this network and many other networks largely contributed to task classification, indicating that connections with visual regions are extensively modulated by behavioral goals. The activity of visual regions is modulated in a variety of attention tasks, possibly reflects the differentiated representation of visual stimuli under specific task context (Mirabella et al., 2007; Reynolds and Heeger, 2009; Jehee et al., 2011). Recent MVPA studies have also found that activity in the visual cortex provides discriminative information on which visual dimension of a stimulus the subjects are processing (Waskom et al., 2014). In addition to biased neural activity, many neuroimaging studies have found that interactions with visual regions are modulated by behavioral goals (Maunsell and Treue, 2006; Al-Aidroos et al., 2012). Attention to different visual categories modulates connections between the occipital and ventral temporal cortexes (Al-Aidroos et al., 2012). And interactions between primary visual regions and frontoparietal regions were enhanced when visual stimuli were attended (Griffis et al., 2015). We provide additional evidence that connections within the visual network and across-network connections between the visual network and many other networks, including the dorsal attention, frontoparietal, and default mode networks, support the reliable discrimination of tasks under naturalistic conditions. Interactions between the frontoparietal and sensory regions are widely thought to play crucial roles in the biased processing of goal-relevant sensory information (Miller and Cohen, 2003; Vossel et al., 2014). Furthermore, enhancements in the connections between the dorsal attention network and the visual network have been observed during natural movie watching, with the possible function of controlling attention to the display (Kim et al., 2017). Our findings are well aligned with these studies. Although there is evidence that the biasing of sensory areas emerges from the frontoparietal regions (Bressler et al., 2008; Baldauf and Desimone, 2014), we cannot investigate this causal relationship because of the use of Pearson correlations to represent interactions between regions. Methods such as dynamic causal modeling (Friston et al., 2003) or Granger causality (Roebroeck et al., 2005) may be employed for future explorations of the direction of influences between regions.

The employed MVPA methods also revealed many other within- and across-network ISFC measures that were modulated by top-down goals. Indeed, we found that almost every canonical subnetwork formed discriminative connections with other subnetworks. Within-network connections in the dorsal attention network and across-network connections between the dorsal attention network and other networks such as the default mode, somatomotor, and frontoparietal networks resulted in high classification accuracies. Recent human neuroimaging experiments and studies in stroke patients have suggested that the dorsal attention network is largely involved in mediating the top-down guided voluntary allocation of attention to locations or features (Ptak and Schnider, 2010; Vossel et al., 2014). Regions from the dorsal attention and frontoparietal networks have also been consistently highlighted in task context representations

(Chiu et al., 2011; Waskom et al., 2014; Long and Kuhl, 2018). Our MVPA results are consistent with these findings, suggesting that functional connections with frontoparietal regions are differentially modulated by behavioral goals. The successful task classification based on across-network connections with the somatomotor network may be due to the biased processing of action information in the behavioral attention task, since action observations activate the somatomotor regions (Buccino et al., 2001) and engage a network of sensorimotor brain regions called the action observation network (Gardner et al., 2015). Natural vision has been shown to modulate large-scale network interactions, and recent attention studies using naturalistic paradigms have demonstrated that attention to complex semantic information changes the neural activity of widely distributed regions (Cukur et al., 2013; Nastase et al., 2017). Our current findings of distinct changes in broadly distributed within-network and across-network connectivity suggest that goal-directed behavior under naturalistic conditions is reflected not solely by local changes in specific activations or connectivity but likely by systematic changes across large-scale brain networks. One limitation of this study is that the sample size of the publicly available dataset we used is relatively small (12 subjects). This relatively small sample size may compromise the reliability of experimental results to some degree. In the future, we should collect more fMRI data by ourselves to enhance the experimental results.

In summary, using the novel ISFC method, we show that selective processing of complex visual information under naturalistic conditions modulates large-scale FC and that this modulation supports the reliable discrimination of top-down task goals. We identified a large number of within-network and across-network discriminative connections, suggesting that goal-directed processing of naturalistic stimuli modulates the coordination of wide swaths of brain regions that belong to different canonical functional networks. This analysis based on large-scale brain networks extends previous studies of goal-directed behavior that focused on changes in local neural activity by showing that the modulation of connectivity between brain regions is broadly distributed. Our study may shed light on the role of large-scale brain networks in goal-directed behavior and suggests that the ISFC may provide an efficient method for identifying task-evoked networks.

ETHICS STATEMENT

All participants provided written, informed consent prior to participating in the study in compliance with the Committee for the Protection of Human Subjects at Dartmouth College, including a provision for data to be shared with other researchers around the world or on a publicly available data archive. The study was approved by the Institutional Review Board of Dartmouth College, and participants received monetary compensation for their participation.

AUTHOR CONTRIBUTIONS

ZW and YL designed research and performed research. ZW, TY, XY, and YL analyzed data. ZW, TY, and YL wrote the paper.

FUNDING

This work was supported by the National Key R&D Program of China under Grant 2017YFB1002505, the National Natural Science Foundation under Grants 61633010 and 61876064, the Guangdong Natural Science Foundation under Grant 2014A030312005, and the Guangzhou Medical and Health Science and Technology Project 20151A 010110.

REFERENCES

Al-Aidroos, N., Said, C. P., and Turk-Browne, N. B. (2012). Top-down attention switches coupling between low-level and high-level areas of human visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 109, 14675–14680. doi: 10.1073/pnas.1202095109

Baldauf, D., and Desimone, R. (2014). Neural mechanisms of object-based attention. *Science* 344, 424–427. doi: 10.1126/science.1247003

Bressler, S. L., Tang, W., Sylvester, C. M., Shulman, G. L., and Corbetta, M. (2008). Top-down control of human visual cortex by frontal and parietal cortex in anticipatory visual spatial attention. *J. Neurosci.* 28, 10056–10061. doi: 10.1523/JNEUROSCI.1776-08.2008

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* 13, 400–404. doi: 10.1046/j.1460-9568.2001.01385.x

Cao, H., Plichta, M. M., Schäfer, A., Haddad, L., Grimm, O., Schneider, M., et al. (2014). Test-retest reliability of fMRI-based graph theoretical properties during working memory, emotion processing, and resting state. *Neuroimage* 84, 888–900. doi: 10.1016/j.neuroimage.2013.09.013

Chao-Gan, Y. and Yu-Feng, Z. (2010). DPARSF: a matlab toolbox for "pipeline" data analysis of resting-state fMRI. *Front. Syst. Neurosci.* 4:13. doi: 10.3389/fnsys.2010.00013

Chiu, Y.-C., Esterman, M., Han, Y., Rosen, H., and Yantis, S. (2011). Decoding task-based attentional modulation during face categorization. *J. Cogn. Neurosci.* 23, 1198–1204. doi: 10.1162/jocn.2010.21503

Cole, M., Bassett, D., Power, J., Braver, T., and Petersen, S. (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., Reynolds, J. R., Power, J. D., Repovs, G., Anticevic, A., and Braver, T. S. (2013). Multi-task connectivity reveals flexible hubs for adaptive task control. *Nat. Neurosci.* 16, 1348–1355. doi: 10.1038/nn.3470

Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3:201. doi: 10.1038/nrn755

Cukur, T., Nishimoto, S., Huth, A. G., and Gallant, J. L. (2013). Attention during natural vision warps semantic representation across the human brain. *Nat. Neurosci.* 16:763. doi: 10.1038/nn.3381

Engel, A. K., Fries, P., and Singer, W. (2001). Dynamic predictions: oscillations and synchrony in top–down processing. *Nat. Rev. Neurosci.* 2:704. doi: 10.1038/35094565

Finn, E. S., Shen, X., Scheinost, D., Rosenberg, M. D., Huang, J., Chun, M. M., et al. (2015). Functional connectome fingerprinting: Identifying individuals based on patterns of brain connectivity. *Nat. Neurosci.* 18:1664. doi: 10.1038/nn.4135

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U.S.A.* 102, 9673–9678. doi: 10.1073/pnas.0504136102

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

Gardner, T., Goulden, N., and Cross, E. S. (2015). Dynamic modulation of the action observation network by movement familiarity. *J. Neurosci.* 35, 1561–1572. doi: 10.1523/JNEUROSCI.2942-14.2015

Gilbert, C. D., and Li, W. (2013). Top-down influences on visual processing. *Nat. Rev. Neurosci.* 14, 350–363. doi: 10.1038/nrn3476

Greene, A. S., Gao, S., Scheinost, D., and Constable, R. T. (2018). Task-induced brain state manipulation improves prediction of individual traits. *Nat. Commun.* 9:2807. doi: 10.1038/s41467-018-04920-3

Griffis, J. C., Elkhetali, A. S., Burge, W. K., Chen, R. H., and Visscher, K. M. (2015). Retinotopic patterns of background connectivity between v1 and fronto-parietal cortex are modulated by task demands. *Front. Hum. Neurosci.* 9:338. doi: 10.3389/fnhum.2015.00338

Hasson, U., Malach, R., and Heeger, D. J. (2010). Reliability of cortical activity during natural stimulation. *Trends Cogn. Sci.* 14, 40–48. doi: 10.1016/j.tics.2009.10.011

Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., and Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science* 303, 1634–1640. doi: 10.1126/science.1089506

Haxby, J. V., Connolly, A. C., and Guntupalli, J. S. (2014). Decoding neural representational spaces using multivariate pattern analysis. *Annu. Rev. Neurosci.* 37, 435–456. doi: 10.1146/annurev-neuro-062012-170325

Hebart, M. N., Bankson, B. B., Harel, A., Baker, C. I., and Cichy, R. M. (2018). The representational dynamics of task and object processing in humans. *Elife* 7:e32816. doi: 10.7554/eLife.32816

Isik, L., Meyers, E. M., Leibo, J. Z., and Poggio, T. (2013). The dynamics of invariant object recognition in the human visual system. *J. Neurophysiol.* 111, 91–102. doi: 10.1152/jn.00394.2013

Jehee, J. F., Brady, D. K., and Tong, F. (2011). Attention improves encoding of task-relevant features in the human visual cortex. *J. Neurosci.* 31:8210. doi: 10.1523/JNEUROSCI.6153-09.2011

Kim, D., Kay, K., Shulman, G. L., and Corbetta, M. (2017). A new modular brain organization of the bold signal during natural vision. *Cereb. Cortex* 28, 3065–3081. doi: 10.1093/cercor/bhx175

Long, N. M., and Kuhl, B. A. (2018). Bottom-up and top-down factors differentially influence stimulus representations across large-scale attentional networks. *J. Neurosci.* 38, 2495–2504. doi: 10.1523/JNEUROSCI.2724-17.2018

Loose, L. S., Wisniewski, D., Rusconi, M., Goschke, T., and Haynes, J.-D. (2017). Switch independent task representations in frontal and parietal cortex. *J. Neurosci.* 37, 8033–8042. doi: 10.1523/JNEUROSCI.3656-16.2017

Maunsell, J. H. R., and Treue, S. (2006). Feature-based attention in visual cortex. *Trends Neurosci.* 29, 317–322. doi: 10.1016/j.tins.2006.04.001

Miller, E. K., and Cohen, J. D. (2003). An integrative theory of prefrontal cortex function. *Annu. Rev. Neurosci.* 24, 167–202. doi: 10.1146/annurev.neuro.24.1.167

Mirabella, G., Bertini, G., Samengo, I., Kilavik, B. E., Frilli, D., Della Libera, C., and Chelazzi, L. (2007). Neurons in area v4 of the macaque translate attended visual features into behaviorally relevant categories. *Neuron* 54, 303–318. doi: 10.1016/j.neuron.2007.04.007

Mišić, B., and Sporns, O. (2016). From regions to connections and networks: new bridges between brain and behavior. *Curr. Opin. Neurobiol.* 40, 1–7. doi: 10.1016/j.conb.2016.05.003

Nastase, S. A., Connolly, A. C., Oosterhof, N. N., Halchenko, Y. O., Guntupalli, J. S., Visconti di Oleggio Castello, M., et al. (2017). Attention selectively reshapes the geometry of distributed semantic representation. *Cereb. Cortex* 27, 4277–4291. doi: 10.1093/cercor/bhx138

Nastase, S. A., Halchenko, Y. O., Connolly, A. C., Gobbini, M. I., and Haxby, J. V. (2018). Neural responses to naturalistic clips of behaving animals in two different task contexts. *Front. Neurosci.* 12:316. doi: 10.3389/fnins.2018.00316

Nichols, T. E. and Holmes, A. P. (2002). Nonparametric permutation tests for functional neuroimaging: a primer with examples. *Hum. Brain Mapp.* 15, 1–25. doi: 10.1002/hbm.1058

Norman, K. A., Polyn, S. M., Detre, G. J., and Haxby, J. V. (2006). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends Cogn. Sci.* 10, 424–430. doi: 10.1016/j.tics.2006.07.005

Petersen, S. E., and Posner, M. I. (2012). The attention system of the human brain: 20 years after. *Annu. Rev. Neurosci.* 35, 73–89. doi: 10.1146/annurev-neuro-062111-150525

Power, J. D., Barnes, K. A., Snyder, A. Z., Schlaggar, B. L., and Petersen, S. E.(2012). Spurious but systematic correlations in functional connectivity MRI networks arise from subject motion. *Neuroimage* 59, 2142–2154. doi: 10.1016/j.neuroimage.2011.10.018

Power, J. D., Cohen, A. L., Nelson, S. M., Wig, G. S., Barnes, K. A., Church, J. A., et al. (2011). Functional network organization of the human brain. *Neuron* 72, 665–678. doi: 10.1016/j.neuron.2011.09.006

Ptak, R., and Schnider, A. (2010). The dorsal attention network mediates orienting toward behaviorally relevant stimuli in spatial neglect. *J. Neurosci.* 30, 12557–12565. doi: 10.1523/JNEUROSCI.2722-10.2010

Reynolds, J. H., and Heeger, D. J. (2009). The normalization model of attention. *Neuron* 61, 168–185. doi: 10.1016/j.neuron.2009.01.002

Roebroeck, A., Formisano, E., and Goebel, R. (2005). Mapping directed influence over the brain using granger causality and fMRI. *Neuroimage* 25, 230–242. doi: 10.1016/j.neuroimage.2004.11.017

Rosenthal, G., Tanzer, M., Simony, E., Hasson, U., Behrmann, M., and Avidan, G. (2017). Altered topology of neural circuits in congenital prosopagnosia. *Elife* 6:e25069. doi: 10.7554/eLife.25069

Schultz, D. H., and Cole, M. W. (2016). Higher intelligence is associated with less task-related brain network reconfiguration. *J. Neurosci.* 36, 8551–8561. doi: 10.1523/JNEUROSCI.0358-16.2016

Shen, X., Papademetris, X., and Constable, R. T. (2010). Graph-theory based parcellation of functional subunits in the brain from resting-state fMRI data. *Neuroimage* 50, 1027–1035. doi: 10.1016/j.neuroimage.2009.12.119

Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., and Greicius, M. D. (2012). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cereb. Cortex* 22, 158–165. doi: 10.1093/cercor/bhr099

Simony, E., Honey, C. J., Chen, J., Lositsky, O., Yeshurun, Y., Wiesel, A., et al. (2016). Dynamic reconfiguration of the default mode network during narrative comprehension. *Nat. Commun.* 7:12141. doi: 10.1038/ncomms12141

Sporns, O. (2014). Contributions and challenges for network models in cognitive neuroscience. *Nat. Neurosci.* 17, 652–660. doi: 10.1038/nn.3690

Spreng, R. N., Stevens, W. D., Chamberlain, J. P., Gilmore, A. W., and Schacter, D. L. (2010). Default network activity, coupled with the frontoparietal control network, supports goal-directed cognition. *Neuroimage* 53, 303–317. doi: 10.1016/j.neuroimage.2010.06.016

van den Heuvel, M. P., and Hulshoff Pol, H. E. (2010). Exploring the brain network: a review on resting-state fMRI functional connectivity. *Eur. Neuropsychopharmacol.* 20, 519–534. doi: 10.1016/j.euroneuro.2010.03.008

Van Dijk, K. R., Sabuncu, M. R., and Buckner, R. L. (2012). The influence of head motion on intrinsic functional connectivity MRI. *Neuroimage* 59:431. doi: 10.1016/j.neuroimage.2011.07.044

Vatansever, D., Menon, D. K., Manktelow, A. E., Sahakian, B. J., and Stamatakis, E. A. (2015). Default mode dynamics for global functional integration. *J. Neurosci.* 35, 15254–15262. doi: 10.1523/JNEUROSCI.2135-15.2015

Vaziripashkam, M., and Xu, Y. (2017). Goal-directed visual processing differentially impacts human ventral and dorsal visual representations. *J. Neurosci.* 37, 3392–3316. doi: 10.1523/JNEUROSCI.3392-16.2017

Von Stein, A., and Sarnthein, J. (2000). Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization. *Int. J. Psychophysiol.* 38, 301–313. doi: 10.1016/S0167-8760(00)00172-0

Vossel, S., Geng, J. J., and Fink, G. R. (2014). Dorsal and ventral attention systems: distinct neural circuits but collaborative roles. *Neuroscientist* 20, 150–159. doi: 10.1177/1073858413494269

Waskom, M. L., Kumaran, D., Gordon, A. M., Rissman, J., and Wagner, A. D. (2014). Frontoparietal representations of task context support the flexible control of goal-directed cognition. *J. Neurosci.* 34, 10743–10755. doi: 10.1523/JNEUROSCI.5282-13.2014

Womelsdorf, T., Schoffelen, J.-M., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., et al. (2007). Modulation of neuronal interactions through neuronal synchronization. *Science* 316, 1609–1612. doi: 10.1126/science.1139597

Woolrich, M. W., Ripley, B. D., Brady, M., and Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of fMRI data. *Neuroimage* 14, 1370–1386. doi: 10.1006/nimg.2001.0931

Xia, M., Wang, J., and Yong, H. (2013). Brainnet viewer: A network visualization tool for human brain connectomics. *PLoS ONE* 8:e68910. doi: 10.1371/journal.pone.0068910

Zalesky, A., Fornito, A., and Bullmore, E. (2012). On the use of correlation as a measure of network connectivity. *Neuroimage* 60, 2096–2106. doi: 10.1016/j.neuroimage.2012.02.001

Zeng, L. L., Shen, H., Liu, L., Wang, L., Li, B., Fang, P., et al. (2012). Identifying major depression using whole-brain functional connectivity: a multivariate pattern analysis. *Brain* 135(Pt 5):1498. doi: 10.1093/brain/aws059

Conflict of Interest Statement: 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.

Copyright © 2019 Wen, Yu, Yang and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 13 March 2019
doi: 10.3389/fnins.2019.00178

Masking the Integration of Complementary Shape Cues

*Andrew Geoly and Ernest Greene**

Laboratory for Neurometric Research, Department of Psychology, University of Southern California, Los Angeles, CA, United States

Retinal and cortical mechanisms provide for persistence of visual information across intervals of many hundreds of milliseconds, which supports the integration of partial shape cues. The present experiments displayed unknown shapes in a match recognition task, wherein a target shape was quickly followed by a comparison shape; the task was to specify whether the comparison shape was the same or different from the target. The target and comparison shapes were displayed as sparse dots that marked boundary locations. The first experiment successively displayed the target shape as two complementary subsets and found that the probability of correct match remained above chance with up to 500 ms of subset separation. The second experiment demonstrated masking of the target by a random pattern of dots when the target and mask were displayed simultaneously, but with much less or no masking when the two were separated by 100 ms. The third experiment displayed the target subsets with 200 ms of separation and found that match recognition was disrupted when the random-dot mask was displayed midway between the two subsets. Much less masking of an intact target was produced with that amount of temporal separation, which suggests that mechanisms for integration of shape cues have a special vulnerability to masking. The third experiment also found very little impairment of match recognition when the mask was displayed simultaneous with one of the subsets. We hypothesize that there is embedding of the subset pattern within the mask pattern, but additional display of the other subset effectively disembeds the buried partial shape cues.

Keywords: visual persistence, visual masking, working memory, shape perception, shape integration

OPEN ACCESS

Edited by:
Daya Shankar Gupta, Camden County College, United States

Reviewed by:
Christian Wallraven, Korea University, South Korea
Gideon Paul Caplovitz, University of Nevada, Reno, United States

***Correspondence:**
Ernest Greene
egreene@usc.edu

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *13 September 2018*
Accepted: *14 February 2019*
Published: *13 March 2019*

Citation:
Geoly A and Greene E (2019) Masking the Integration of Complementary Shape Cues. Front. Neurosci. 13:178. doi: 10.3389/fnins.2019.00178

INTRODUCTION

> *"[T]here is some support for the view that sensory persistence is produced by the activity of coding mechanisms at the level of feature extraction in visual information processing".*
>
> *Vincent* Di Lollo (1977).

A substantial amount of work has been done to evaluate early stages of shape encoding, with masking and manipulation of neuronal persistence providing some of the most effective research tools. It seems odd, therefore, that we have very little information about the effect of masking on the persistence of stimulus information. Persistence across several 100 ms is thought to mediate working memory, and masking has been used to manipulate the contents of working memory, so it is a reasonable hope that the combined use of both methods would provide useful insights about the nature of working memory.

The research reported below differed from prior research in fundamental ways. First, the targets to be identified were unknown shapes, each seen only once using a match-recognition protocol. A given source shape consisted of a sequence of dots that marked an outer boundary, similar to an outlined figure (see **Figure 1**). The experiments displayed reduced-density versions of these shapes, meaning that a sparse subset of boundary markers provided the target to be identified. Following display of a given target, a sparsely marked comparison shape was displayed that either matched the target or provided a non-matching shape. Because each target shape was shown only once to a given respondent, all identification was based on short-term memory.

Second, the display equipment avoids some of the problems that have bedeviled the study of temporal integration of visual cues, which can also be described as information persistence (Coltheart, 1980). Much of the earlier work has used cathode ray tubes with short-lived phosphors to display the stimuli, so that any persistence could be correctly attributed to the visual system rather than the display itself. A fast phosphor can fall to about 1% residual emission within microseconds, but it can then persist at that level for a second or more. Rayner and Pollatsek (1983) demonstrated that observers could still perceive this weak afterglow, providing a basis for combining successive displays using persistence of the stimulus itself. Jonides et al. (1983) failed to replicate earlier work done in their own lab when the displays were presented with LEDs rather than with phosphor-based images (see also Di Lollo et al., 1997, 2000). The present work avoids these issues by using an LED array to display the shape stimuli. Also, the equipment can display successive dot patterns with microsecond control of timing for pulse duration and interstimulus interval.

Initial experiments provided confirmation that classic information persistence and masking effects would be found with this match-recognition task. Then the influence of masking on temporal integration was evaluated.

The target shape was divided into complementary subsets that were successively displayed, which requires the integration of shape information across an interstimulus interval spanning hundreds of milliseconds. Then the ability of a mask to impair the integration process was examined. The results suggest the possibility of a motion-to-form encoding mechanism as part of working memory, as will be discussed subsequently.

Visibility of a briefly displayed image can be reduced by presentation of a different image, which we describe as visual masking. Generally, one image provides the target, with the task requiring discrimination or recognition, and the mask acts to impair effective performance of the task. The relative timing of target and mask can determine whether visibility is affected. The experimental protocol is designated as *forward masking* if the mask precedes the target, *backward masking* if it follows the target, and *concurrent (simultaneous) masking* if the two are presented at the same moment (Breitmeyer, 2007).

The earliest masking studies used large, spatially uniform increments of luminance (Baxt, 1871; Crawford, 1947; Sperling, 1965). Subsequent work often has used masks that contained image elements, such as shapes, lines, or dots (Rieger et al., 2005). The term *pattern masking* serves as a general classification of masks that provide these components (Enns and Di Lollo, 2000), irrespective of whether there is temporal separation, i.e., backward, forward, or simultaneous display. Impaired discrimination or recognition of the target has been attributed to inhibitory interactions among neurons that register the image elements (Weisstein et al., 1975; Breitmeyer and Ganz, 1976; Macknik and Livingstone, 1998).

There are subcategories relating to the spatial attributes of pattern masks. For instance, noise masks could be made up of dots or boxes that have little in common with the target image (Kinsbourne and Warrington, 1962). Alternatively, structure masks would be those that bear a strong resemblance to the target, or have lines with common orientations. For example, a pattern composed of lines might be used to mask alphabetical letters. Specific pattern masks can vary with respect to contrast, luminance, or other physical parameters, limited only by the creativity and discretion of the investigator.

Three different mechanisms for masking are often invoked, one being the *erasure* of stimulus information, another calling for *integration* (merging) of stimulus information, and the third providing for *interruption* of perceptual processing. Each may be acting within the retina or in cortex. Persistence of retinal activity in *integration* masking can cause the target and mask to be perceived as a unitary pattern that precludes recognition or discrimination of the target itself (Eriksen, 1966; Turvey, 1973; Breitmeyer, 1984). This model seems most plausible when one gets maximal disruption of performance with simultaneous display of target and mask, and one perceives the combined image of the target embedded in the mask pattern. It is common to see the masking effect become nil with about 100 ms of target/mask separation in either direction (Enns and Di Lollo, 2000).

Alternatively, pattern masking can occur with substantial temporal separation of target and mask through *interruption* of information processing (Bachmann and Allik, 1976). Because the mask can act across an extended interval, most attribute the masking to disruption of cortical mechanisms that are required for recognition or discrimination. Alternatively, one might see a U-shaped function wherein there is progressive impairment of performance across an extended range of target/mask separation, followed by recovery of target recognition or discrimination (Bachmann and Allik, 1976; Michaels and Turvey, 1979).

Conditions that produce a delay of target masking have been designated as *metacontrast masking*. Here the mask consists of an annulus that surrounds the target image, and masking only occurs within a narrow temporal range (Enns and Di Lollo, 2001). When the target and mask are simultaneously displayed, or the interval between them is very short, the target is clearly visible and is seen as lying within the interior of the mask. With very long separations the mask is ineffective at impairing shape processing, and both are visible. At intermediate intervals, perception of the target is impaired, yielding a U-shaped function of accuracy (hit rate) as a function of the separation interval (Enns and Di Lollo, 2001). One explanation for the effect, which can be described as a "two-channel" theory, proposes that image information is transmitted by a fast burst of neuronal activity followed by sustained (tonic) activity that conveys fine details about the stimulus attributes. A metacontrast mask is thought to

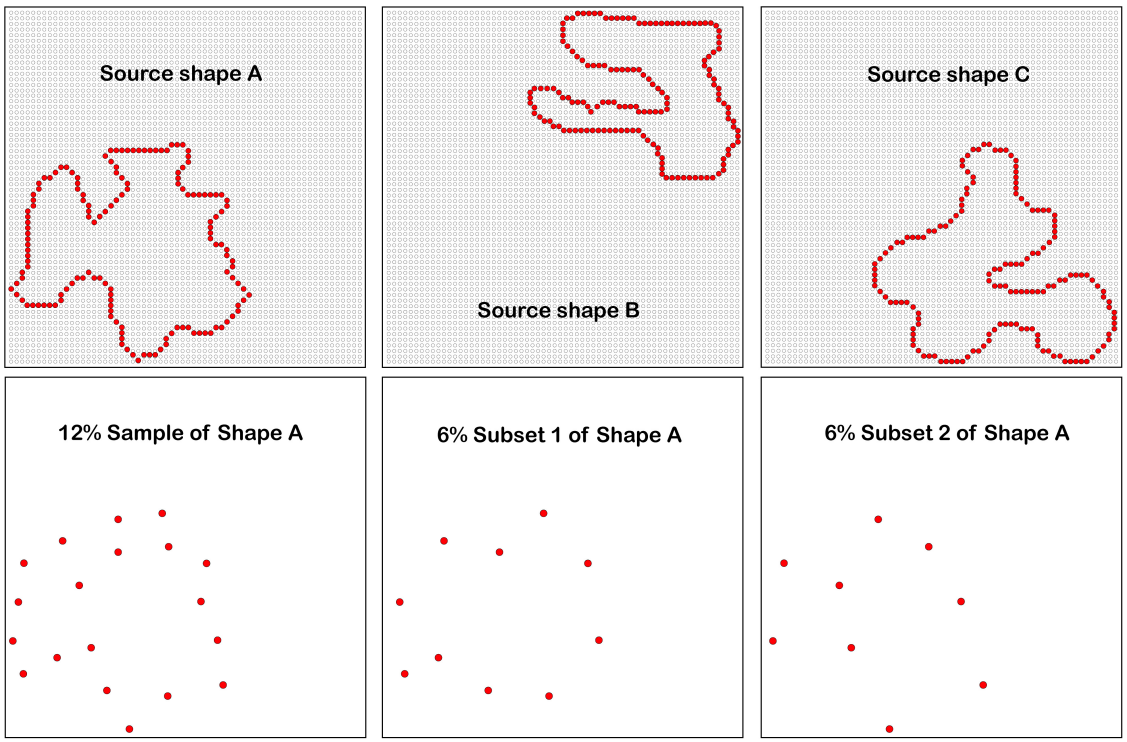

FIGURE 1 | Three examples of source shapes are shown in the upper panels, with the dots of the display board provided as background. The experiments displayed low density patterns that were sampled from source shapes, providing either 12% sets or successive displays of complementary 6% subsets. For ease of discourse, the term "shape" will also be used to describe these patterns. For better visibility, the background dots of the array are not shown in the three lower panels, and dot size has been increased.

occur when the fast-acting signals from the mask's onset interfere with the sustained slower signals of the tonic channel, disrupting the processing information about the earlier target.

The first experiment was a replication of a Greene and Hautus (2018) experiment that examined temporal integration of shape cues for unknown shapes in a match-recognition protocol. The shape cues were provided by sparse dots that marked the outer boundary of a given shape. This is a new experimental approach to the study of shape recognition, so it is appropriate to show that the integration of shape cues over a span of half a second is a reliable finding. This is especially worthwhile given current concerns about reproducibility of results.

The second experiment examined mask interference with the integration of unknown shapes, wherein the mask consisted of a random pattern of dots. The masking stimuli were random dot patterns, which seems especially appropriate for either overwriting the information from sparse markers, or interrupting short-term memory of that information. The experimental results provided evidence for classic disruption of shape recognition when the target and mask were simultaneously displayed, but the mask was relatively ineffective when it was separated from the shape cues by about 100 ms.

The third experiment examined mask interference with temporal integration of the shape cues. Based on findings from Experiment 2, there was an expectation that simultaneous display of the mask with a subset of the shape cues would greatly impair match recognition. Further, there was an expectation that there would be no interference with match recognition where the temporal separation of mask from the shape cues was 100 ms, as was found in Experiment 2. Neither expectation was confirmed, i.e., the results were the opposite of expectation. These findings suggest some new principles for how stimulus information is integrated and stored in working memory, which will be discussed once the experiments have been reported.

MATERIALS AND METHODS

Authorization and Consent of Respondents

This study was carried out in accordance with the recommendations and guidelines of the Psychology Department Subject Pool. The protocol was approved by the USC Institutional Review Board. All respondents gave written informed consent in accordance with the Declaration of Helsinki. A total of 24 undergraduates volunteered and provided data, eight for each of the three experiments reported below.

Source Shapes, Sets, and Subsets

An inventory of 480 unknown shapes provided the source of stimulus patterns that were displayed in each of the three experiments, so hereafter they will be described as "source

shapes." Each source shape consisted of a continuous string of dot locations on a display board (detailed below), forming a shape boundary akin to a silhouette. The number of dots in source shapes ranged from 100 to 269, with the mean being 168 dots. Distance from the centroid to dots ranged from 12.8 to 22.3 dots, mean distance being 16.9 dots. The shapes were constructed to avoid similarity with known shapes and objects, to avoid long-term memory factors and to focus on early sensory encoding. Examples of three of the source shapes are shown in the upper panels of **Figure 1**.

The three experiments displayed low-density samples drawn from source shapes, which can be designated as "sets" and "subsets." We can characterize these dot patterns as "shapes," with the understanding that they are providing various degrees of effective cues relative to the original shapes from which they were derived.

Shape sets were at 12% density and the subsets provided complementary 6% densities. The dots for a given 12% set were chosen by first randomly picking a starting point from among the boundary dots of the source shape, and then proceeding along the boundary, marking every eighth dot to be included for display. The final dot that was chosen at the end of the circuit would commonly leave a span that was shorter than the others. The computer made these selections "on the fly," meaning that the dots displayed to a given respondent were chosen at random on each trial.

For the experiments that displayed complementary 6% subsets, the 12% set was further divided. Beginning at a randomly chosen starting point, dots were successively numbered. The odd numbered dots were designated as "subset 1" and the even numbered dots were assigned to "subset 2." The fact that the two subsets can be combined to provide a 12% set is the basis for describing them as being "complementary."

Mask Stimuli

Experiments 2 and 3 included display of random-dot masks to evaluate the conditions that would impair match recognition of the targets. A 4% dot-density level was chosen to provide approximately the same number of dots in the mask as the mean number of dots among the shapes in the inventory. To be specific, the mean number of dots in the inventory of shapes is 166 (100 at the minimum and 269 maximum), and a 4% random sample from the full LED array provides 164 dots. For trials in which a set or subset with displayed simultaneous with the mask, the random selection of dot locations did not include the locations of set or subset dots. A different random pattern was used on each trial in which a mask was presented.

Figure 2 illustrates how the random-dot pattern is effective at precluding perception of low-density shape samples. The left panel shows a 12% set derived from Shape C of **Figure 1**. The middle panel shows a 4% mask superimposed on the set. Here the mask dots are shown in gray so that one can still pick out the locations of dots in the 12% set. The right panel shows all the dots in red, this being the stimulus that the respondent would see with simultaneous display of the mask and shape set. It is clear from inspection that the 12% set cannot be discriminated in the presence of the random-dot pattern, which assures that the mask would be effective in impairing match recognition. The 4% random-dot pattern would also mask information from display of complementary 6% subsets, which together would be equivalent to the 12% set.

Stimulus Displays

Room illumination was dim (10 lux). Shape and mask patterns were displayed as brief flashes on a 64 × 64 array of LEDs mounted on a display board. All dots of a given pattern were displayed as simultaneous ultra-brief flashes. Each flash had a duration of 10 μs and an intensity of 1000 μW/sr. At a viewing distance of 3.5 m, the visual angle of a given dot of the display board was 4.92 arc°, dot to dot spacing was 9.23 arc°, and the total span of the 64 × 64 array (horizontal and vertical) was 9.80 arc°. The shape patterns that were displayed would therefore have overall dimensions that ranged from 2.0 to 3.5 arc°, with the mean being 2.6 arc°.

Basic Task Conditions

The basic task can be described as requiring a match-to-sample judgment, which for convenience can be described as match recognition. When used with unknown shapes, one can assess the encoding and persistence of shape information without the confounding influence of long-term memory (Greene and Hautus, 2017, 2018).

The present work calls for initial display of a "target" set (or subsets), followed by display of a "comparison" set that might or might not be the same shape as the target. On each trial, target and comparison shapes were chosen at random from the inventory. A given shape was shown only once as a target, or only once as a non-matching comparison shape. On half the trials the comparison shape was the same as the target shape, which was designated as "matching," and on half it was a different shape, designated as "non-matching." In other words, the task is asking whether the cues provided by the target set or subset are sufficient for recognition of the comparison set.

The order of treatments was chosen at random. The corner in which the target was displayed was chosen at random on each trial in each of the experiments. The comparison set was then displayed in one of the other three corners, again chosen at random. Positioning of a given pattern required placement of at least one dot in the outside boundary of the top or bottom, and one dot on a side edge of the array.

A fixation point consisting of four central dots was provided prior to each trial, and respondents were instructed to keep their eyes centered on this location. Following each display sequence, the respondent voiced a decision of whether the comparison shape was the "same" or "different" from the target shape, and this information was entered by the experimenter into a computer file. Neither the experimenter nor the respondent was informed as to which treatment condition was presented on a given trial, or whether the judgment was correct.

Experimental Treatments

Experiment 1 displayed complementary 6% subsets as targets, with six levels of inter-stimulus interval, specifically: 0, 100, 200,

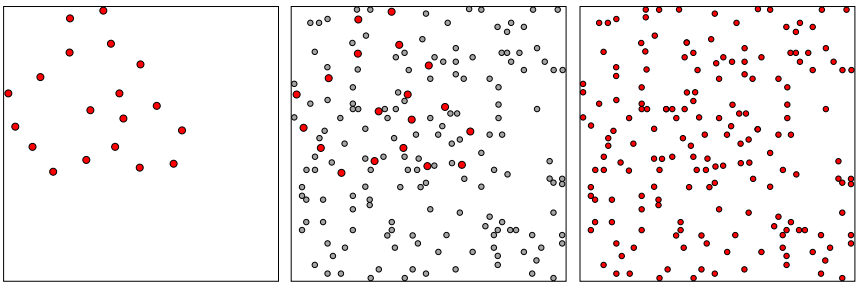

FIGURE 2 | The left panel illustrates a low-density shape set, i.e., a 12% sample of dots from Shape C in **Figure 1**. The middle panel shows a 4% random-dot pattern that could mask of the shape set. Here the mask dots are shown in gray so that one can still see the dots of the 12% set (in red). Dot sizes in the set have been enlarged in both panels to provide for better discrimination of the shape set in relation to mask dots. The right panel has colored all dots red, with mask and shape-set dots rendered as the same size. This illustrates the stimulus that would be seen with simultaneous display of the mask and the 12% shape set. One can see that the mask precludes perception of shape boundary markers, and the same would be true for the 6% subsets, which provide only half the number of markers.

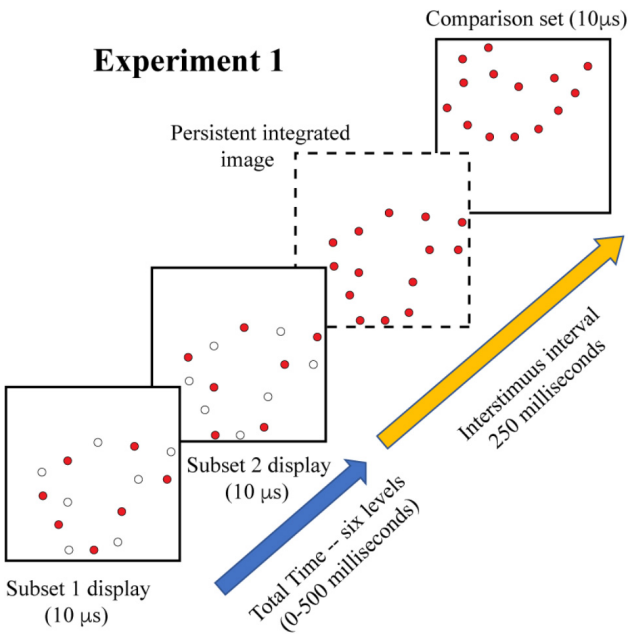

FIGURE 3 | On each trial of Experiment 1, a randomly sampled shape provided the complementary subsets as a target. The frames show the displayed subset as colored dots, and include open dots to show which members of the 12% pattern would not be displayed, i.e., they remain dark at that moment. Display of subset 1 was followed by subset 2 at an inter-stimulus interval that varied from 0 to 500 ms. If the shape information from the two displays were completely integrated by the visual system, the resulting image would contain 12% of the boundary markers, as illustrated in the frame shown with broken lines. This amount of shape information would be expected to provide a moderately high level of shape identification. A comparison shape was shown 250 ms after display of the second subset, providing an opportunity for a shape-matching decision.

300, 400, and 500 ms. The comparison shape. either matching or non-matching, was displayed after an additional interval of 250 ms. The display sequence is illustrated in **Figure 3**. Each subject judged 25 trials for each of these treatment conditions for a total of 300 trials.

Experiment 2 examined masking of target-shape information. All targets were displayed with 12% density, this being to demonstrate effectiveness of masking against integrated 6% subsets (in Experiment 3, to follow). A 4% random-dot mask was added to each target-comparison sequence, with display of the mask coming either before, during, or after display of the target set. The mask/target intervals were: −100, −50, 0, +50, +100 ms – the negative values designating display of the mask prior to the target and positive values designating display of the mask after the target. At 0 ms the mask dots were displayed at the same 10 μs moment, so the dots of both patterns were superimposed, as illustrated in **Figure 2**. The inter-stimulus interval between target shapes and comparison shapes was again 250 ms. **Figure 4** illustrates the display timing. Each subject judged 32 trials for each of the five treatment conditions for a total of 320 trials.

Experiment 3 combined treatments that would require integration of target cues as well as masking of those cues. On each trial, the two 6% subsets were displayed with a temporal separation of 200 ms. A 4% mask was inserted into this sequence to provide the potential for disruption of the integration process. For logging of data and statistical analysis, mask timing was specified relative to the first subset display, i.e., at 0, 50, 100, 150, and 200 ms. However, the mask was expected to produce disruption of performance when simultaneously displayed with either of the subsets, and provide the least influence of judgments at the midpoint between display of the subsets. Therefore, we have re-designated the treatment levels as 0, 50, 100, 50, and 0, providing labels that reflect two symmetrical limbs of mask influence. These display conditions are illustrated in **Figure 5**.

Two additional control treatments that did not include a mask stimulus were added to provide measures that aid in interpreting effects. One control treatment displayed just the two subsets, which provided evidence of performance from simple integration of the two sources of shape information. The other control treatment displayed only one of the subsets, this to establish the level of performance if there was masking of information from only one of the subsets. As in each of the earlier experiments, the comparison shape was presented 250 ms after

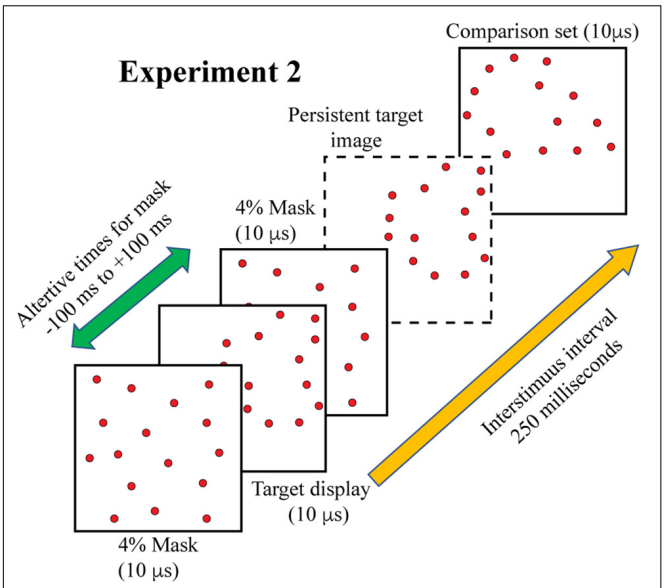

FIGURE 4 | On each trial a full 12% target shape was flashed, followed 250 ms later by the comparison shape. In addition, a 4% random-dot mask was displayed at times that either preceded, followed, or was simultaneous with the target display. The illustration shows the 4% mask twice to represent the two ends of the range at which it could be displayed. The dashed frame shows a target shape that was unaffected by the mask, though this would not be expected if the mask erased the stimulus information, occluded it, or otherwise interfered with shape processing.

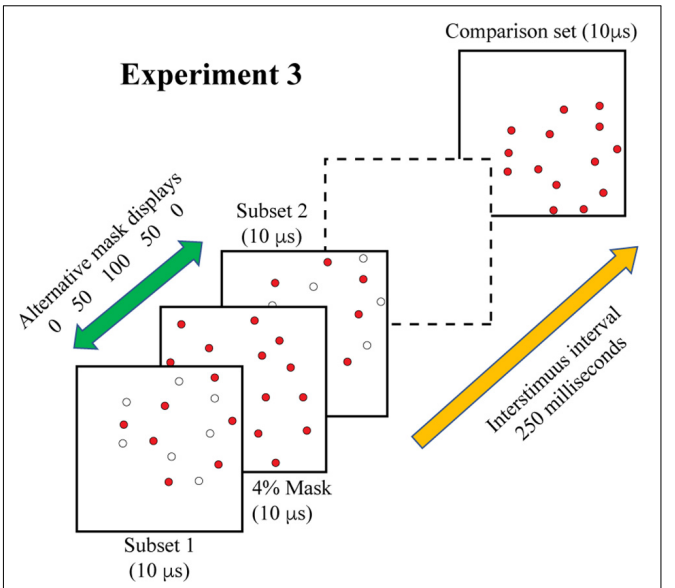

FIGURE 5 | In Experiment 3 the two complementary subsets were displayed 200 ms apart, with the comparison shape being provided 250 ms after the second subset was shown. A 4% random-dot mask was also provided, either simultaneous with subset 1, simultaneous with subset 2, or at three intervals in between the subset displays. We are designating the midpoint of this range as 100, for this would display the mask 100 ms away from each subset. The goal was to determine whether information from the two subsets would be integrated or if the mask would impair this process. This illustration shows the mask precluding effective integration, i.e., impaired match recognition, in that the dashed frame does not contain a persistent image of the combined subsets. Two additional control conditions were included in the experiment (see text).

display of second subset. Each of the five mask and two control conditions were displayed to a given subject for 22 trials, for a total of 308 trials.

Bias Correction of Judgments

Responses were evaluated with signal detection analysis that derived an unbiased index of performance. Signal detection theory provides a framework that corrects for response bias. A method developed by Macmillan and Creelman (2005) was adopted, which uses the bias-correcting formula $d' = \sqrt{2}(z(H) - z(F))$. In this formula, H is the proportion of "same" judgments to matching shapes, F is the proportion of "same" judgments to non-matching shapes, and $z(\bullet)$ is the inverse-normal transform (Green and Swets, 1966/1974) Values of F and H were adjusted for values of 0 or 1 (which would otherwise lead to $d' = \pm\infty$) prior to calculation of d'. We adopted the log-linear correction for this purpose (DeCarlo, 1998). Bias correction requires the combination of response information from both matching and non-matching shapes.

It is more intuitive to express performance as a proportion, which can be done by converting d' into $p(c)_{\max}$ using the formula:

$$p\,(c)_{\max} = \Phi\left(d' \Big/ \left(2\sqrt{2}\right)\right)$$

Here the function $\Phi(\bullet)$ is the cumulative distribution of the normal distribution. The $p(c)_{\max}$ index scales with 0.5 being chance and 1.0 being decisions that are perfectly correct. For convenience, the present discourse will describe this index as "probability of match recognition."

Statistical Analysis

For each of the experiments, linear mixed-model regression was used to test for omnibus treatment effects. Experiment 1 was a replication of earlier work and there exists a substantial body of literature on masking effects, so we had clear expectations about the influence of treatments for each of the three experiments. This justified the use of planned comparisons to test hypotheses about specific treatment effects. For Experiments 1 the only test of interest was whether the mean at the longest temporal separation would be above chance. For Experiment 2 the question was whether performance would be above chance where the mask and shape subset were displayed simultaneously.

Experiment 3 provided results that were the opposite of what was predicted, so *post hoc* tests of mean differences were done instead of planned comparisons. Specifically, we tested whether each mean that was observed with simultaneous display of mask and shape subset differed from the control condition that displayed the subsets with no mask being present. We also tested whether masks presented midway between the two shape subsets, i.e., at 100, differed from the one-subset control condition, and whether it differed from chance.

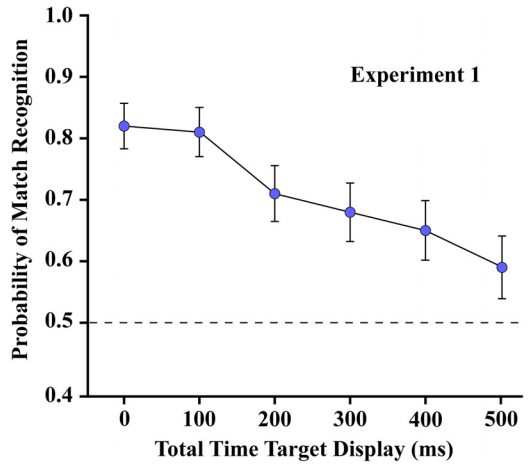

FIGURE 6 | The mean probability of match recognition for the respondents of Experiment 1 are plotted for each of the six levels of temporal separation of the two subsets. Judgments were well above the chance level of 0.5 with up to 500 ms of separation.

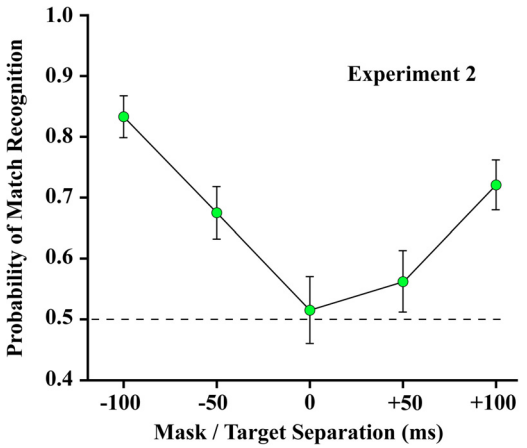

FIGURE 7 | The 4% mask was ineffective at blocking match recognition when it preceded the 12% target by 100 ms and was only partially effective when it led by 50 ms. Match recognition was at chance levels when the mask and target were displayed at the same instant (designated as 0 ms). Judgments were above chance when the mask followed the target by 50 ms or more.

A piece-wise linear regression was calculated to assess the influence of temporal separation of the mask from each subset. Based on the results of Experiment 2, there was an expectation that the mask would block shape information when it was display at the same moment as the subset, and would have little or no effect when it was separated from the subset by 100 ms. Therefore, the plan was to do a separate regression analysis for each leg of the sequence, i.e., from 0 (mask + subset 1) to 100, and then from 100 to 0 (mask + subset 2) – see **Figures 5**, **8**. The regression itself makes no prediction about the direction of effect, so this was still the appropriate analysis even though the results were opposite of what had been expected.

RESULTS

Experiment 1

As shown in **Figure 6**, the probability of correct matching decisions was quite high (above 0.8) where the two subsets were displayed simultaneously, i.e., with temporal separation of zero. Performance dropped as the interval between the two subsets was increased, which reflects a decline of persistence of shape information from the first subset. Regression across the five treatment levels confirmed that the decline was significant at $p < 0.0001$ (slope = −0.0005/ms, $t_7 = -9.37$).

The probability of match recognition was still well above chance at the longest subset separation interval (500 ms). A planned comparison established that the mean at the 500 ms treatment level was significantly different from chance ($t_{35} = 2.92$, one sided $p = 0.0031$). That level of performance was about the same that found in Experiment 3, where a control condition provided display of a single subset (see below). It is likely, therefore, that by 500 ms all shape information from the first subset had completely evaporated, providing a level of match recognition that could be elicited by the second subset, acting alone.

Experiment 2

The results of Experiment 2 are shown in **Figure 7**. Linear regression confirmed a significant decline in forward masking between −100 and 0 ms (slope = −0.0032/ms, $t_{23} = -8.74$, $p < 0.0001$). Match recognition was well above chance when the mask preceded the target by 50 ms, and no masking was evident with a temporal separation of 100 ms. Trends were similar when the mask followed the shape set (slope = 0.0020/ms, $t_{23} = 5.46$, $p < 0.0001$), though masking was less complete at the longest interval. These results suggest that greater proximity of mask and target is needed for forward masking to be effective than is required for backward masking. The mean at 0 ms of mask/target separation was not significantly different from chance ($t_{28} = 0.49$, unadjusted $p = 0.6273$). [Where the mean does not differ significantly from chance, an unadjusted comparison is the more conservative statistic].

Experiment 3

For Experiment 3, piece-wise linear regression found a significant decline in match recognition as the mask was temporally separated from each of the shape subsets (see **Figure 8**). The decline with mask separation from subset 1 (from 0 to 100) was significant at $p = 0.0031$ (slope = −0.0011/ms, $t_{30} = -3.22$), and with separation from subset 2 (from 100 to 0) was significant at $p = 0.0205$ (slope = 0.0008/ms, $t_{30} = 2.45$). Simultaneous display of mask and shape subsets yielded means that were not significantly different from the first control condition (green broken line), wherein the two subsets were displayed without any mask being provided ($t_{42} = -0.52$, unadjusted $p = 0.6054$ and $t_{42} = -1.13$, unadjusted $p = 0.2632$ for subset 1 and subset 2, respectively). The mask at treatment level 100 – midway

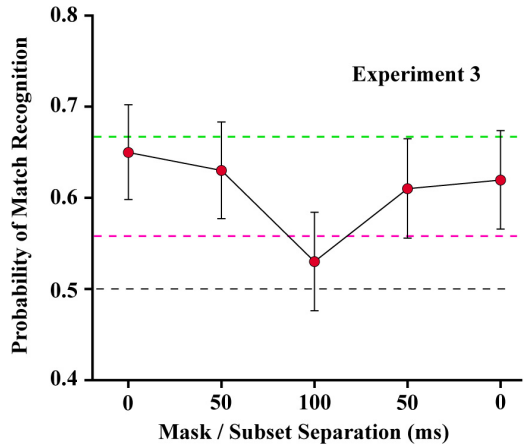

FIGURE 8 | Mean probability of match recognition judgments are plotted for the masking conditions of Experiment 3. The mask was ineffective at blocking match recognition when it was displayed simultaneous with one of the subsets, i.e., at either 0 position. At ±50 ms of separation from a subset the mask was only partially effective. At 100 of separation, midway between the two target subsets, match recognition was not greater than chance. The green dashed line indicates the level of match recognition from successive display of the two subsets, wherein there was no mask in the display sequence. The red dashed line indicates the level of match recognition when only a single (6%) subset was displayed.

between the shape subsets – did not differ significantly from the second control condition (red broken line) that assessed match recognition with display of a single subset ($t_{42} = -1.11$, unadjusted $p = 0.2750$).

DISCUSSION

Visible persistence is the continued perception of a sensory image, like an afterimage (Coltheart, 1980), and is thought to reflect persistence of neuron activity within the retina. It is quite sensitive to the physical parameters of a stimulus, given that increasing the luminance of a stimulus will cause the period of visible persistence to decline (Dick, 1974). Also, it is quite vulnerable to simultaneous masking or backward masking protocols that have short inter-stimulus intervals (Long, 1980). A bright, spatially uniform flash can readily disrupt processing of image content (Baxt, 1871; Crawford, 1947; Sperling, 1965). It is generally thought that masking of visible persistence occurs within a 100 ms window (Coltheart, 1980; Greene and Visani, 2015), so the simultaneous display of the mask with the 12% target (Experiment 2) would have intruded upon visible persistence of the boundary markers.

In Experiment 2 the 4% random-dot mask likely embedded the 12% targets when the two were presented at the same moment, so the shape could no longer be identified as the same or different from the comparison shape that was subsequently displayed. The random-dot mask was ineffective at blocking match-recognition when it preceded the target shape by 100 ms, and performed slightly better when it followed the target shape. These results are consistent with previous findings (Schultz and Eriksen, 1977).

If the stimulus is sufficiently salient, it can go out from the retina very quickly (VanRullen and Thorpe, 2001). If the separation of pattern and mask is 100 ms or longer, one can be fairly sure the image information has been passed to cortex and any masking disrupts retrieval of information from working memory or long term memory (Kolers, 1962; Turvey, 1973; Enns and Di Lollo, 2000; Vogel et al., 2006).

Information persistence is a longer duration process that mediates encoding of stimulus information (Coltheart, 1980; Greene, 2007). The duration of information persistence grows longer as the duration of stimulus display is increased (Irwin and Yeomans, 1986). Numerous laboratories have reported persistence of visual information for many hundreds of milliseconds (Sperling, 1960; Eriksen and Collins, 1968; Hogben and Di Lollo, 1974; Coltheart, 1980; Chun and Potter, 1995; Ward et al., 1997; Vogel et al., 2006). Some of the evidence was based on what can be called the "temporal integration" paradigm, where stimulus information is divided into complementary subsets and the time required to integrate the information is assessed. Greene and Visani (2015) displayed letters composed of dots that were divided into complementary subsets, as in the present work. There was substantial summation of information from the two displays for 200 ms, after which the hit rate remained above the one-subset level across the treatment range (one second being the longest that was tested). Greene (2016) found that complementary-dot subsets provided for persistence of information for recognition of "thin" letters for upward of 600 ms.

Letters are extremely overlearned and the number of potential alternatives are relatively small, thus it should not be surprising that the choices could remain correct on the basis of minimal information, making it possible to observe an extended duration of persistence. Earlier work had suggested that shape information will decay much faster. Greene (2014) used a temporal integration protocol where the task called for recognition of diverse real-world shapes, e.g., animals, plants, vehicles, tools, furniture. Here the boost provided by the temporal integration condition declined to the one-subset level within 100 ms. For this task the information to be retrieved was extremely open-ended, requiring comparison of shape cues against an indeterminate store of shape memories. Singer and Kreiman (2014) found similar results for integration of image patches where the task called for specifying the category of the objects being shown. Asking to identify the outline of a real-world object may require substantially more information, meaning that even modest decay of the information could preclude effective recognition.

The present work used unknown shapes, each being displayed only once, so decisions were not based on retrieval of information from long-term memory. Match recognition declined as a function of time, following a near-linear trajectory, but remaining well above chance across the 500 ms range that was tested. A prior report from this laboratory found similar results (Greene and Hautus, 2018). We are confident that the temporal integration protocol – the display of complementary subsets – calls for integration of shape information in working memory, and this

integration can be provided across intervals of several 100 ms. This finding is critical in interpreting the findings in Experiment 3, which yielded very unexpected results.

Pattern masks, which would include the 4% random-dot mask used here, can work by erasing, embedding, or overwriting the shape information that might otherwise be discriminated or recognized. In Experiment 3 the target pattern was divided into two complementary subsets that were displayed with a 200 ms separation. Displaying the mask in the middle of this interval, with a 100 ms of separation from either subset, yielded performance that was statistically within the chance range. In Experiment 2 the mask was completely ineffective when it followed the 12% target by a 100 ms and was substantially ineffective when it preceded the target by that much. A follow-up experiment using a 6% target found that recognition was at chance levels when the mask preceded the target by 100 ms (unreported data). Therefore, there should have been no impairment of performance in Experiment 3 where the mask was separated from both subsets having 6% density by 100 ms, designated as zero in **Figure 8**. Yet this condition produced the greatest level of masking, with performance being in the chance range.

Apparently, with 100 ms of separation the mask does not greatly impair fully integrated shape information that is being held in working memory (Experiment 2) but does impair the processing of information that is being integrated, as was the case for the temporal integration required in Experiment 3. We infer that the integration of shape information is a special state that is more vulnerable to masking. As a potentially related matter, Greene and Hautus (2018) found the decline in match recognition across a 500 ms interval was more rapid when two subsets were being integrated than when the dots of the target were displayed one at a time. Apparently having a higher complement of boundary dots available at a given moment can foreclose the integration process, and once completed, the summary is less subject to decay or disruption. This result might be attributed to object substitution masking (Enns and Di Lollo, 1997), wherein the information from the first subset is lost and hence no integration is possible between the first and the second subsets.

Experiment 3 also displayed the mask simultaneously with the 6% patterns of subset 1 or subset 2. This should have completely precluded any use of shape information from the masked subset, given that simultaneous masking of a 12% target produced chance performance (Experiment 2). Instead, masking of subset information was relatively weak, and match recognition was well above chance. These results suggest that the subset pattern becomes embedded in the random dot pattern of the mask, and the other subset is able to disembed the subset pattern from the mask. Apparently this can occur in either direction. So where the mask was superimposed on subset 1, the subsequent encounter with subset 2 accomplished disembedding of the target information. And where subset 1 was displayed alone, it persisted across the 200 ms and was able to disembed the subset 2 pattern from the mask. **Figure 9** illustrates this hypothesis.

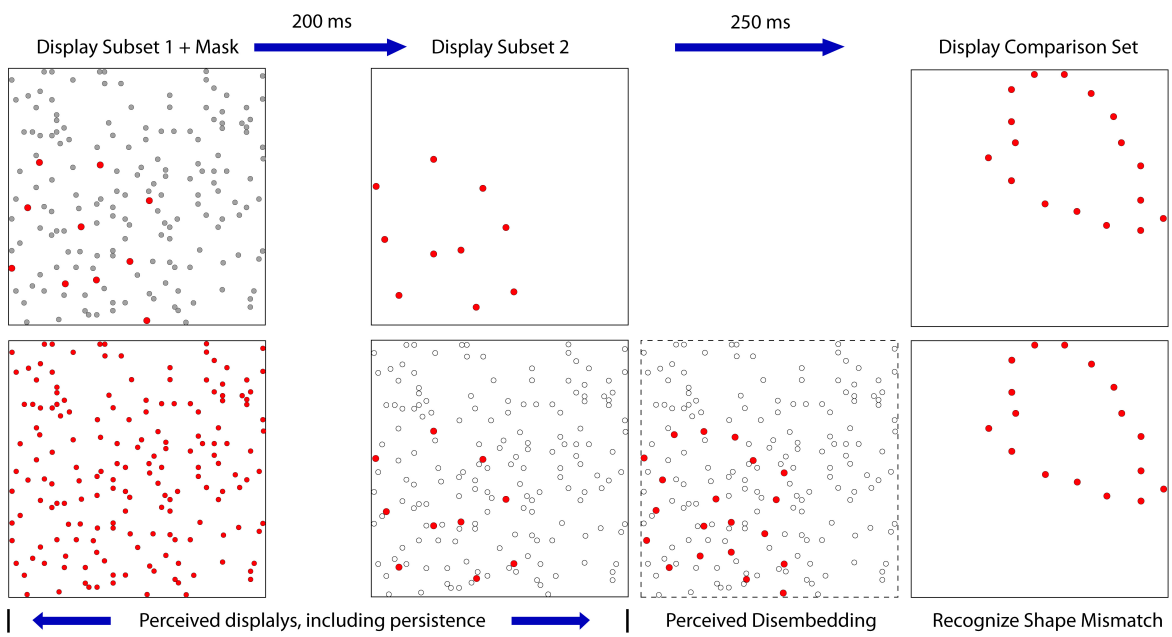

FIGURE 9 | This figure illustrates the disembedding concept where the mask is presented simultaneously with subset 1. The upper panels show the stimulus configurations, providing subset dots as red and the mask dots as gray. The lower panels illustrate the perceptual states that are hypothesized. The first panel on the left shows the first subset as being buried within the mask dots, which prevents the pattern from being perceived. The second panel shows a decreased salience of mask dots due to decay of persistence, so the newly flashed dots of the second subset are conspicuous. The third panel illustrates the second subset disembedding the dots of the first subset, making the full target set available for match recognition. A similar process is assumed for simultaneous presentation of the mask with the second subset.

We previously noted that a mask can impair recognition through erasure, integration (embedding), or disruption of process. We now have evidence that a partial target pattern that is embedded in a noise mask can be disembedded by display of the remaining target information. The process of disembedding a subset from a background might be attributed to feature salience of the subset pattern. But perhaps it is a product of form-based motion processing, which would be consistent with much of the current literature relating to dynamic object perception. Finding that simultaneous masking of a subset did not greatly impair match recognition is consistent with conclusions by McCarthy et al. (2017) that simultaneous events should not be integrated if they do not appear along the same contour. Consequently, when the second subset is presented, it may be interpreted as a rotational or translational continuation of the first subset — irrespective of the mask — and vice versa.

It has been postulated that although an object's identity is maintained during occlusion, its specific features might not be. The continuity seen across objects that disappear behind a source of occlusion and then re-emerge may have more to do with spatio-temporal continuity than spatial continuity (Erlikhman and Caplovitz, 2017). A case can be made that a temporal integration paradigm with an interleaved mask emulates this same type of occlusion and re-emergence, which would explain the results in terms of similar spatio-temporal continuity.

The disembedding concept that is illustrated is **Figure 9** might relate to motion-form cueing. A real-life example might be seen where a dog runs behind a white picket fence. The dog's features are broken up into subset components (in the slits between the boards), none of which would be recognizable as a dog. But on seeing the complements of each subset as a sequence of cues, the viewer perceives a dog running behind the fence. The processes of disembedding may be akin to this kind of motion cueing.

Interestingly, some reports of brain mechanisms are consistent with the current results. An interaction between the dorsal and ventral visual pathways is thought to underlie form motion interactions, in particular the updating of "no-longer-visible" information (McCarthy et al., 2017). Processing form and motion information calls for activity from a number of brain structures, including V3A, V3B, Kinetic Occipital cortex (KO), Medial Temporal cortex (MT), and the inferior parietal sulcus. This system may provide mechanisms for deriving "structure from motion" (Klaver et al., 2008) "biological motion" (Vaina et al., 2001), processing of motion edges (Vinberg and Grill-Spector, 2008) and contour curvature during rotational motion (Caplovitz and Tse, 2007). Moreover, in an experiment to investigate what information is represented during dynamic occlusion, Erlikhman and Caplovitz (2017) measured BOLD fMRI activity across both early (V1–V3) and higher-level cortical areas while observers viewed various shapes passing behind occluding quadrants. They found that that the information represented in early visual cortex during dynamic occlusion is not shape-specific. Rather, it may correspond to the object's position, its motion path, or the path of attention. Further analysis found that shape identity could be decoded in higher visual areas such as VO, LO, TO, LOC, PHC, parahippocampal place area, and hMT. We hope our research may aid in understanding the interaction between dorsal and ventral stream pathways that have a role in dynamic form-object processing. Those who do classical psychophysics studies might aim to replicate our results and members of the neuroimaging community might employ a similar masking paradigm to assess fMRI activation.

AUTHOR CONTRIBUTIONS

EG designed the study, provided custom test equipment, experimental apps, and oversaw drafting of the manuscript. AG tested respondents, evaluated the data, did the literature review, and provided various drafts of the manuscript.

FUNDING

This experiments were run using custom applications written by Jack Morrison of Digital Insight. Statistical analysis was performed by Dr. Wei Wang, Harvard Medical School. Funding for this research was provided by the Neuropsychology Foundation and the Quest for Truth Foundation.

REFERENCES

Bachmann, T., and Allik, J. (1976). Integration and interruption in the masking of form by form. *Perception* 5, 79–97. doi: 10.1068/p050079

Baxt, N. (1871). Ueber die zeit, welche nöthig ist, damit ein gesichtseindruck zum bewusstsein kommt und über die grösse der bewussten wahrnehmung bei einem gesichtseindrucke von gegebener dauer. *P. Archiv Für Die Gesammte Physiologie Des Menschen Und Der Thiere.* 4, 325–336. doi: 10.1007/bf01612494

Breitmeyer, B. G. (1984). *Visual Masking: An Integrative Approach*. Oxford: Clarendon Press.

Breitmeyer, B. G. (2007). Visual masking: past accomplishments, present status, future developments. *Adv. Cogn. Psychol.* 3, 9–20. doi: 10.2478/v10053-008-0010-7

Breitmeyer, B. G., and Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychol. Rev.* 83, 1–36. doi: 10.1037//0033-295x.83.1.1

Caplovitz, G. P., and Tse, P. U. (2007). Rotating dotted ellipses: motion perception driven by grouped figural rather than local dot motion signals. *Vision Res.* 47, 1979–1991. doi: 10.1016/j.visres.2006.12.022

Chun, M., and Potter, M. (1995). A two-stage model for multiple target detection in rapid serial visual presentation. *J. Exp. Psych. Hum. Percept. Perform.* 21, 109–127. doi: 10.1037/0096-1523.21.1.109

Coltheart, M. (1980). Iconic memory and visible persistence. *Percept. Psychophys.* 27, 183–228. doi: 10.3758/bf03204258

Crawford, B. H. (1947). Visual adaptation in relation to brief conditioning stimuli. *Proc. R. Soc. B Biol. Sci.* 134, 283–302. doi: 10.1098/rspb.1947.0015

DeCarlo, L. T. (1998). Signal detection theory and generalized linear models. *Psychol. Methods* 3, 186–205. doi: 10.1037/1082-989X.3.2.186

Di Lollo, V. (1977). Temporal characteristics of iconic memory. *Nature* 267, 241–243. doi: 10.1038/267241a0

Di Lollo, V., Enns, J. T., Yantis, S., and Dechief, L. G. (2000). Response latiencies to the onset and offset of visual stimuli. *Percept. Psychophys.* 62, 218–225. doi: 10.3758/BF03212073

Di Lollo, V., Seiffert, A. E., Burchett, G., Rabeeh, R., and Ruman, T. A. (1997). Phosphor persistence of oscilloscopic displays: a comparison of four phosphors. *Spat. Vis.* 10, 353–360. doi: 10.1163/156856897X00276

Dick, A. O. (1974). Iconic memory and its relation to perceptual processing and other memory mechanisms. *Percept. Psychophys.* 16, 575–596. doi: 10.3758/BF03198590

Enns, J. T., and Di Lollo, V. D. (1997). Object substitution: a new form of masking in unattended visual locations. *Psych. Sci.* 8, 135–139. doi: 10.1111/j.1467-9280.1997.tb00696.x

Enns, J. T., and Di Lollo, V. D. (2000). What's new in visual masking? *Trends. Cogn. Sci.* 4, 345–352. doi: 10.1016/s1364-6613(00)01520-5

Enns, J. T., and Di Lollo, V. D. (2001). "An object substitution theory of visual masking," in *From Fragments to Objects Segmentation and Grouping in Vision*, eds T. Shipley and P. Kellman (New York, NY: Elsevier), 121–143.

Eriksen, C. W. (1966). Temporal luminance summation effects in backward and forward masking. *Percept. Psychophys.* 1, 87–92. doi: 10.3758/bf03210033

Eriksen, C. W., and Collins, J. F. (1968). Sensory traces versus the psychological moment in the temporal organization of form. *J. Exp. Psych.* 77(3 Pt 1), 376–382. doi: 10.1037/h0025931

Erlikhman, G., and Caplovitz, G. P. (2017). Decoding information about dynamically occluded objects in visual cortex. *NeuroImage* 146, 778–788. doi: 10.1016/j.neuroimage.2016.09.024

Green, D. M., and Swets, J. A. (1966/1974). *Signal Detection Theory and Psychophysics*. Huntington, NY: Robert E. Krieger Publishing.

Greene, E. (2007). Information persistence in the integration of partial cues for object recognition. *Percept. Psychophys.* 69, 772–784. doi: 10.3758/BF03193778

Greene, E. (2014). Shape recognition elicited by microsecond flashes is not based on photon quantity. *Iperception* 5, 87–93. doi: 10.1068/i0639rep

Greene, E. (2016). Information persistence evaluated with low-density dot patterns. *Acta Psychol.* 170, 215–225. doi: 10.1016/j.actpsy.2016.08.005

Greene, E., and Hautus, M. J. (2017). Demonstrating invariant encoding of shapes using a matching judgment protocol. *AIMS Neurosci.* 4, 120–147. doi: 10.3934/Neuroscience.2017.3.120

Greene, E., and Hautus, M. J. (2018). Evaluating persistence of shape information using a matching protocol. *AIMS Neurosci.* 5, 81–96. doi: 10.3934/neuroscience.2018.1.81

Greene, E., and Visani, A. (2015). Recognition of letters displayed as briefly flashed dot patterns. *Atten. Percept. Psychophys.* 77, 1955–1969. doi: 10.3758/s13414-015-0913-6

Hogben, J. H., and Di Lollo, V. D. (1974). Perceptual integration and perceptual segregation of brief visual stimuli. *Vis. Res.* 14, 1059–1069. doi: 10.1016/0042-6989(74)90202-8

Irwin, D. E., and Yeomans, J. M. (1986). Sensory registration and informational persistence. *J. Exp Psych. Hum. Percept. Perform.* 12, 343–360. doi: 10.1037/0096-1523.12.3.343

Jonides, J., Irwin, D. E., and Yantis, S. (1983). Failure to integrate information from successive fixations. *Science* 222, 188. doi: 10.1126/science.6623072

Kinsbourne, M., and Warrington, E. K. (1962). The effect of an after-coming random pattern on the perception of brief visual stimuli. *Q. J. Exp. Psychol.* 14, 223–234. doi: 10.1080/17470216208416540

Klaver, P., Lichtensteiger, J., Bucher, K., Dietrich, T., Loenneker, T., and Martin, E. (2008). Dorsal stream development in motion and structure-from-motion perception. *NeuroImage* 39, 1815–1823. doi: 10.1016/j.neuroimage.2007.11.009

Kolers, P. A. (1962). Intensity and contour effects in visual masking. *Vis. Res.* 2, 277–294. doi: 10.1016/0042-6989(62)90037-8

Long, G. (1980). Iconic memory: a review and critique of the study of short-term visual storage. *Psychol. Bull.* 88, 785–820. doi: 10.1037/0033-2909.88.3.785

Macknik, S. L., and Livingstone, M. S. (1998). Neuronal correlates of visibility and invisibility in the primate visual system. *Nat. Neurosci.* 1, 144–149. doi: 10.1038/393

Macmillan, N. A., and Creelman, C. D. (2005). *Detection Theory: A User' s Guide*, 2nd Edn. Mahwah, NJ: Lawrence Erlbaum Associates.

McCarthy, J. D., Erlikhman, G., and Caplovitz, G. P. (2017). The maintenance and updating of representations of no longer visible objects and their parts. *Prog. Brain Res.* 236, 163–192. doi: 10.1016/bs.pbr.2017.07.010

Michaels, C. F., and Turvey, M. T. (1979). Central sources of visual masking: indexing structures supporting seeing at a single, brief glance. *Psychol. Res.* 41, 1–61. doi: 10.1007/bf00309423

Rayner, K., and Pollatsek, A. (1983). Is visual information integrated across saccades? *Percept. Psychophys.* 34, 39–48. doi: 10.3758/BF03205894

Rieger, J. W., Braun, C., Bülthoff, H. H., and Gegenfurtner, K. R. (2005). The dynamics of visual pattern masking in natural scene processing: a magnetoencephalography study. *J. Vis.* 5, 10. doi: 10.1167/5.3.10

Schultz, D. W., and Eriksen, C. W. (1977). Do noise masks terminate target processing? *Mem. Cogn.* 5, 90–96. doi: 10.3758/bf0320919

Singer, J. M., and Kreiman, G. (2014). Short temporal asynchrony disrupts visual object recognition. *J. Vis.* 14:7. doi: 10.1167/14.5.7

Sperling, G. (1960). The information available in brief visual presentations. *Psychol. Monogr.* 74, 1–29. doi: 10.1037/h0093759

Sperling, G. (1965). Temporal and spatial visual masking by impulse flashes. *J. Opt. Soc. Am.* 55, 541. doi: 10.1364/josa.55.000541

Turvey, M. T. (1973). On peripheral and central processes in vision: inferences from an information-processing analysis of masking with patterned stimuli. *Psychol. Rev.* 80, 1–52. doi: 10.1037/h0033872

Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., and Belliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proc. Natl. Acad. Sci. U.S.A.* 98, 11656–11661. doi: 10.1073/pnas.191374198

VanRullen, R., and Thorpe, S. J. (2001). The time course of visual processing: from early perception to decision-making. *J. Cogn. Neurosci.* 13, 454–461. doi: 10.1162/08989290152001880

Vinberg, J., and Grill-Spector, K. (2008). Representation of shapes, edges, and surfaces across multiple cues in the human visual cortex. *J. Neurophysiol.* 99, 1380–1393. doi: 10.1152/jn.01223.2007

Vogel, E. K., Luck, S. J., and Woodman, G. F. (2006). The time course of consolidation in visual working memory. *J. Exp. Psych. Hum. Percept. Perform.* 32, 1436–1451. doi: 10.1037/0096-1523.32.6.1436

Ward, R., Duncan, J., and Shapiro, K. (1997). Effects of similarity, difficulty, and nontarget presentation on the time course of visual attention. *Percept. Psychophys.* 59, 593–600. doi: 10.3758/bf03211867

Weisstein, N., Ozog, G., and Szoc, R. (1975). A comparison and elaboration of two models of metacontrast. *Psychol. Rev.* 82, 325–343. doi: 10.1037/0033-295X.82.5.325

Conflict of Interest Statement: 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.

Copyright © 2019 Geoly and Greene. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Integrative Neuroscience

ORIGINAL RESEARCH
published: 25 June 2019
doi: 10.3389/fnint.2019.00020

Modulation of Functional Connectivity Between Dopamine Neurons of the Rat Ventral Tegmental Area *in vitro*

Luuk van der Velden[1], *Martin A. Vinck*[2], *Taco R. Werkman*[1] *and Wytse J. Wadman*[1]*

[1] Center for Neuroscience, University of Amsterdam, Amsterdam, Netherlands, [2] Ernst Strüngmann Institute for Neuroscience in Cooperation With Max Planck Society, Frankfurt am Main, Germany

Micro Electrode Arrays were used to simultaneously record spontaneous extracellular action potentials from 10 to 30 dopamine neurons in acute brain slices from the lateral Ventral Tegmental Area (VTA) of the rat. The spike train of an individual neuron was used to characterize the firing pattern: firing rate, firing irregularity and oscillation frequency. Functional connectivity between a pair of neurons was quantified by the Paired Phase Consistency (PPC), taking the oscillation frequency as reference. Under baseline conditions the PPC was significantly different from zero and 42 of the 386 pairs of VTA neurons showed significant coupling. Fifty percent of the recorded dopamine neurons were part of the coupled VTA network. Raising extracellular potassium from 3.5 to 5 mM increased the mean firing rate of the dopamine neurons by 45%. The same increase could be induced by bath application of 300 μm glutamate. High potassium reduced the PPC, but it did not change during the glutamate application. Our findings imply that manipulating excitability has distinct and specific consequences for functional connectivity in the VTA network that cannot be directly predicted from the changes in neuronal firing rates. Functional connectivity reflects the spatial organization and synchronization of the VTA output and thus represents a unique element of the message that is sent to the mesolimbic projection area. It adds a dimension to pharmacological manipulation of the VTA micro circuit that might help to understand the pharmacological (side) effects of e.g., anti-psychotic drugs.

Keywords: dopamine, network, functional connectivity, glutamate, potassium, action potentials, multi-electrode array

OPEN ACCESS

Edited by:
Andreas Bahmer, Universitätsklinikum Würzburg, Germany

Reviewed by:
Fatuel Tecuapetla, Institute of Cellular Physiology, National Autonomous University of Mexico, Mexico
Habibeh Khoshbouei, University of Florida, United States

***Correspondence:**
Wytse J. Wadman
w.j.wadman@uva.nl

Received: *14 March 2019*
Accepted: *06 June 2019*
Published: *25 June 2019*

Citation:
van der Velden L, Vinck MA, Werkman TR and Wadman WJ (2019) Modulation of Functional Connectivity Between Dopamine Neurons of the Rat Ventral Tegmental Area in vitro. Front. Integr. Neurosci. 13:20. doi: 10.3389/fnint.2019.00020

1. INTRODUCTION

The ventral tegmental area (VTA) is a midbrain nucleus alongside the substantia nigra. The VTA plays a role in emotional processing, in the reward system and it is implicated in cognitive functions such as associative learning and memory (Lisman and Grace, 2005; Fields et al., 2007; Fujisawa and Buzsáki, 2011; Kim et al., 2012). The most abundant neuron type in the VTA is the dopamine neuron (Nair-Roberts et al., 2008) which projects to mesolimbic and mesocortical structures. Dopamine neurons form local synaptic connections with each other (Bayer and Pickel, 1990) but they are also interconnected with glutamatergic and GABAergic neurons (Omelchenko and Sesack, 2009). Many details of this local microcircuit are still under intense investigation. *In vivo* VTA dopamine neurons exhibit spontaneous activity and generate single action potentials

(spikes) as well as bursts, with consequences for dopamine release (Gonon, 1988; Paladini and Roeper, 2014). Most remarkably they demonstrate a low quite regular firing rate (1–8 Hz) and this spontaneous activity even persists, at a somewhat lower frequency, in the *in vitro* brain slice which is devoid of external input (Grace and Onn, 1989; Bowery et al., 1994; Werkman et al., 2001; Bayer et al., 2007). The firing pattern can theoretically be explained by the composition of the ion conductances in the membrane wherein calcium currents and/or persistent sodium currents (Khaliq and Bean, 2010; Drion et al., 2011) play a prominent role. The VTA network, and in particularly the dopaminergic neurons, seem also sufficiently synchronized to generate a low frequency oscillation in the local field potential that could well be relevant for memory (Lisman and Grace, 2005; Fujisawa and Buzsáki, 2011) and noise correlations that support reward processing (Kim et al., 2012; Moghaddam et al., 2017). Model network studies show that the strength of the local functional connectivity is a major factor in generating such population oscillations (Traub et al., 1989). Anatomical and physiological studies have indicated considerable heterogeneity in cellular properties of DA neurons with consequences for their firing pattern and potentially linked to distinct output connectivity (Lammel et al., 2008). We have restricted our study to the mesolimbic projecting neurons in the lateral part of the VTA, where the classical slow firing, high DAT containing neurons are located (Björklund and Dunnett, 2007; Lammel et al., 2008). Most pharmacological studies in the VTA use the firing rate of individual dopaminergic neurons as their output parameter (Hand et al., 1987; Wang and French, 1993; Werkman et al., 2001) and indeed binding of various antipsychotic drugs to D_2 receptors on these neurons leads to modulation in their firing rate (Pucak and Grace, 1994, 1996). In Substantia Nigra dual patch clamp recordings demonstrated that direct dopaminergic chemical transmission (as well as electrical) results in coupling (Vandecasteele et al., 2008). In previous work we have shown that the VTA neurons are functionally connected (van der Velden et al., 2017) and this connectivity is in part organized by the volume transmission of dopamine. If they are partially synchronized, drug manipulation at the receptor level might affect the organization of the local VTA network. This network effect of pharmacological manipulation can be determined by simultaneous recording from a sufficiently large sample of VTA dopamine neurons and analyzing their mutual relations. Micro Electrode Arrays (MEA) (Taketani and Baudry, 2010), consisting of a grid of 60 electrodes, were used in this study to simultaneously record the activity of at least 10–20 dopamine neurons within the acute VTA midbrain slice. The *in vitro* preparation has limitations compared to the intact complete brain. However, the fact that the neurons in the VTA slice are completely devoid of external input has the great advantage that it eliminates background interference from other brain regions. Sufficient dopamine neurons were spontaneously active in the VTA slice to estimate a population measure of functional connectivity, based on the detailed spike timing in the spike trains. The connectivity was quantified with a proven statistical method: the Paired Phase Consistency (Vinck et al., 2010). Two simple manipulations were used that increase the mean firing rate of VTA dopamine neurons to almost the same level: increasing extracellular potassium from 3.5 to 5 mM or bath application of 300 μM glutamate. Although these two manipulations result in a similar change in firing rate, they surprisingly yield distinct modulations of the PPC, confirming the role of functional connectivity as an emerging network property with potential functional consequences for more complex pharmacological manipulations.

2. METHODS

2.1. Slice Preparation

Male wistar rats (Harlan, Zeist, The Netherlands) between 75 and 100 g (age >P24) were decapitated. The midbrain was dissected and kept in artificial cerebral spinal fluid (ACSF) at 4°C, containing (in mM) $NaHCO_3$ 25, D-glucose 10, $CaCl_2$ 2.5, NaH_2PO_4 1.25, $MgSO_4$ 1.3, KCl 3.5, NaCl 120, which was bubbled with carbogen (95% O_2; 5% CO_2), pH was 7.4. Coronal slices were cut 300 μm thick from caudal to rostral using a vibratome (Leica VT1000S, Wetzlar Germany). The fading of the substantia nigra during progressive slicing was a marker for the caudal-medial part of the VTA. Two to three slices containing the medial to caudal part of the VTA were used for the experiments. Slices were incubated for 30 min at 32°C directly after slicing and were kept at room temperature until the start of the experiment. All experiments and methods were approved by the ethical committee for animal experimentation of the University of Amsterdam.

2.2. Solutions

With exception of bicuculline (Tocris Bioscience, Abbington, UK), all chemicals were obtained from Sigma-Aldrich (Zwijndrecht, NL). Stock solutions of quinpirole-HCl (10 mM), glutamate monosodium salt (100 mM) and bicuculline (20 mM) stock were made in H_2O. All stock solutions were kept at −20° and diluted just before use.

2.3. Electrophysiology

During recording (MEA-1600, Multichannel Systems, Reutlingen Germany) the slice was kept at 32°C and continuously perfused with ACSF bubbled with carbogen. The VTA was identified in the midbrain slice and positioned on top of the 3D MEA (Qwane Biosciences, Lausanne, Switzerland) containing 60 electrodes (8*8 layout) of 30 μm diameter and 100 μm spacing in order to record the spontaneous activity of multiple single-units (Olivier et al., 2002). A 20 min acclimatization time preceded the recordings.

2.4. Data Acquisition

The extracellular recordings with the 60-channel MEA showed identifiable extracellular spikes of 30 to 130 μV amplitude superimposed on a background noise of about 15 μV. The raw signal was high pass filtered at 225 Hz using a second order Butterworth filter and sampled at 20 kHz. Voltage peaks (positive and negative) were detected, with a relatively low threshold to prevent detection failures. The signal around each peak (± 3 ms) was extracted and K-means clustering was used to cluster the

largest two principle components and the maximum amplitudes of the peak waveforms. The auto-correlation and inter-spike-interval distribution of the peaks in the various clusters were examined to identify clusters consisting of neuronal spikes. For electrodes that contained more than one neuron the most reliably recorded neuron was selected, based on the cluster with the largest peak amplitude.

2.5. Experimental Conditions

In baseline experiments VTA activity was recorded for 40 min under standard conditions. The selective dopamine D_2 receptor agonist quinpirole (1 μM) was used to confirm the dopamine sensitivity of the recorded units. It was administered in the last 3 min of the experiment in 40% of the recordings, containing more than 50% of the reported units. Quinpirole induced an unambiguous cessation of action potential firing in all tested slices ($n = 6$) and neurons ($n = 98$). Two manipulations were used to systematically increase the mean firing rate of the DA neurons: (1) in the high potassium experiment $[K^+]_o$ was increased from 3.5 mM (control) to 5 mM and (2) in the glutamate experiments 300 μM glutamate was added to the standard ACSF. Experiments with high extracellular potassium or glutamate consisted of three wash-in and wash-out (25 min each) sequences. The third application was compared to its preceding baseline for the analysis. To determine the potential role of GABA signaling in the observed phenomena, we performed experiments where 20 μM of the GABA antagonist bicuculline was added to the standard ACSF and administered for 20 min. Here the wash-in and wash-out sequence was repeated two times and the second application was compared to its preceding baseline in the analysis.

2.6. Data Analysis and Statistics

The firing properties of the VTA neurons were characterized by classical measures: the spike waveform, the mean firing rate (spike/s) and the inter-spike-interval (ISI) distribution. The spike duration was computed by detecting a level crossing in both the beginning and end of the spike waveform. The firing of DA neurons is controlled by an underlying intrinsic rhythm (Drion et al., 2011). The dominant oscillation frequency of this rhythm was estimated from the auto-correlation function (with a 50 ms bin size to accommodate the firing rates in the 1–5 Hz range, **Figure 2A**). The oscillation frequency was computed from intervals between side-lobes in the auto-correlation function (details are given in **Figures 2A,B**). The irregularity of neuronal firing in the VTA was assessed using a measure of local variation, which quantifies the similarity between consecutive ISIs. The local variation (LV, Shinomoto et al., 2005, 2009) of a spike train ranges from 0 (perfectly regular firing) to 1 (Poisson distributed firing) and above 1 for burst-like firing and is given by:

$$LV = \frac{3}{n-1}\sum_{i=1}^{n-1}\left(\frac{T_i - T_{i+1}}{T_i + T_{i+1}}\right) \quad (1)$$

where T_i is the i-th interval in the spike train that contains n spikes. Two factors contribute to LV in our situation (1) the local irregularity of the consecutive spikes (jitter) and (2) the consequences of cycle skipping in the firing pattern. The fraction of cycle skipping was calculated as the percentage of the ISIs that were larger than 1.5 times the median ISI. Experimentally, *in vivo*, the local VTA field potential is assumed to best reflect the neuronal Population Output (Fujisawa and Buzsáki, 2011). Our MEA recordings did not provide a field potential of sufficient signal-to-noise ratio, thus we decided to emulate it per slice from all recorded spikes. The Population Output signal of a slice contained the spikes of each recorded neuron in that slice, where a neuron's contribution was normalized to its total number of spikes. This signal was convoluted with a Gaussian kernel (standard deviation 60 ms) to convert it into a continuous signal. The spectral properties of the Population Output signal were computed using Welch's method (Welch, 1967; Hunter, 2007). To determine the strength of the functional connectivity between two neurons that produce a spike train, we used the Paired Phase Consistency (PPC) as previously defined (Vinck et al., 2010). The PPC calculates the similarity of the relative phases of the two trains with respect to a chosen reference frequency and estimates the square of the classic Phase Lock Value (Lachaux et al., 1999). The PPC is an unbiased metric of phase-synchronization that scales with the square rather than the square root of the coherence and phase locking value (Vinck et al., 2010). Thus, a value of 0.0023 corresponds to a coherence value of about 0.048. The PPC indicates the consistency of the relative phase between two spike trains across segments. Assuming a unimodal distribution of relative phases, the probability of having the preferred or most common relative phase for a given segment will be a factor of approximately $(1 + 2 * \sqrt{PPC})/(1 - 2 * \sqrt{PPC})$ larger than the probability of having the non-preferred or least common relative phase (Ardid et al., 2015). This equation follows from Taylor expansion of the circular von Mises distribution around $PPC = 0$. Hence, with PPC values of 0.0023, the average peak-to-through modulation of the relative phase distribution is approximately 21%. To compute the PPC, spike trains were binned at 1 ms bins and a windowed (Hanning) Fourier Transform was computed on a series of (at least 80) time segments of the spike train. The length of the time segments was set to contain a fixed number of cycles of the reference frequency of interest (e.g., 5 cycles). The relative phase is defined as the complex argument of the classic spectral coherence (Lachaux et al., 1999; Vinck et al., 2010). From these relative phases the PPC was computed:

$$PPC = \frac{2}{N(N-1)}\sum_{j=1}^{N-1}\sum_{k=j+1}^{N} cos(\theta_j - \theta_k) \quad (2)$$

where there are N time segments, segment j has relative spike phase θ_j and segment k has relative spike phase θ_k, computed in respect to a chosen reference frequency. The PPC was calculated for each unique neuronal pair in the slice and evaluated at three reference frequencies: the oscillation frequency of each neuron in the pair and their mean oscillation frequency. The frequency that yielded the highest PPC value was used for further analysis. The strong auto-correlation of VTA neurons forms a potential bias for the PPC; this was reduced by selecting time segments that were at least 14 s apart. The size of the blind spot (14 s) was

based on the time within which the auto-correlation functions of all neurons decayed to less than 15% of its maximum value (**Figures 2A,B**). The statistical significance of an observed PPC value for a neuron pair was investigated by comparing it with the calculation over the shuffled dataset. The bootstrapping for the PPC value included 1,200 computations, where time segments were temporally shuffled each time. The experimentally found PPC value was tested against this shuffled distribution (one-sided, $\alpha = 0.05$). The relation between LV and PPC of the neurons was investigated using mutual information analysis (MacKay, 2003). The mutual information was computed over all neuron pairs in the baseline experiments (368 pairs) between the LV values of the neurons in a pair (2-dimensional variable) vs. their shared PPC value (one-dimensional variable). As the mutual information can have a strong positive bias (Panzeri et al., 2007), a shuffle correction was applied (Ince et al., 2009). Unless otherwise mentioned, all values reported in this study are given as mean and standard error of the mean (SE).

3. RESULTS

3.1. Baseline Activity

The activity of spontaneously active, mesolimbic projecting, dopamine neurons in the lateral VTA was first recorded under baseline conditions (68 neurons from 6 experiments, 8–14 neurons per slice, each from a different animal) using the 60 channel MEA. The dopamine neurons fire action potentials that can extracellularly be recognized by their characteristic broad tri-phasic waveforms and have a spike duration of 2.44 (SE 0.04) ms. **Figure 1A** illustrates the mean waveform of all 68 neurons superimposed and normalized to their first peak. The classification of all recorded neurons as principal dopamine neurons was confirmed by measuring the quinpirole sensitivity (see methods). The identical waveform of all spikes allows an easy transformation of the recording into a point process, which has been done for all data that follows. The mean firing rate of all observed neurons under baseline condition was 0.93 (SD 0.45) spikes/s (n = 68, **Figure 1B**). **Figures 1C,D** illustrate two subclasses of firing types that were encountered: (1) quite regularly firing neurons (examples 1–4) and irregularly firing neurons (examples 5–8). For each neuron the irregularity of the firing pattern was quantified by the local variance (Equation 1, LV: mean 0.27 (SD 0.21), n = 68), indicating a large variability over the neurons (compare **Figures 1C,D**). The LV distribution (**Figure 1E**) did not deviate from a unimodal distribution as tested with the Hartigan's diptest. The firing irregularity correlated negatively with the mean firing rate of the neurons (Spearman's $\rho = -0.46, p = 6.7 * 10^{-5}$, **Figure 1F**), indicating that dopamine neurons with a higher baseline firing rate fired more regularly.

Figures 1G,H illustrate the inter-spike-interval (ISI) distribution for a regular firing neuron (LV = 0.15) and an irregular firing one (LV = 0.34). At least two factors contribute to the LV value: (1) the width of the dominant ISI peak, which is much narrower for the neuron in **Figure 1G** than the one in **Figure 1H**) and (2) the fraction of intervals that are multiples of the median interval and reflect "cycle skipping". This was corroborated by the correlation found between the LV and the percentage of ISIs exhibiting cycle skipping (Spearman's $\rho = 0.82, p = 7.6 * 10^{-18}$). These two factors can also be distinguished in the return maps that were made from the same neurons (**Figures 1I,J**). The intervals for the neuron in **Figures 1G–I** were sharply clustered, while that for the neuron in **Figures 1H–J** were more diffuse. The LV and mean firing rate were not different across slices (ANOVA, $\alpha = 0.05, n = 6$), showing that the observed variation originated at the level of individual neurons.

3.2. Neuronal Oscillation Frequency

Most VTA neurons demonstrated a sharp dominant peak in their ISI distribution (e.g., **Figures 1G,H**) suggesting a preferred spike interval (cycle time), associated with equidistant side lobes in their auto-correlation function. As expected, the side lobes were more prominent in regular firing neurons (**Figure 2A**, side lobes indicated by dots) than the ones in irregular firing neurons (**Figure 2B**). The oscillation frequency of the neuronal activity was determined from the time intervals between the side-lobes in the auto-correlation (see markers in **Figures 2A,B**). The mean oscillation frequency of the recorded neurons was 1.53 (SD 0.47) Hz (n = 68) and was either equal but often considerable higher than the mean firing rate (**Figure 2C**). The difference between the oscillation frequency and the mean firing rate was more pronounced at low firing rates (**Figure 2C**). In contrast to the mean firing rate, the oscillation frequency did not correlate with the firing irregularity (LV) (Spearman rank regression, $p = 0.33$, **Figure 2D**), which confirms that the oscillation frequency is less sensitive to cycle skipping than the firing rate and therefore a better and preferred estimator of the intrinsic rhythm of the activity of the VTA dopamine neuron.

3.3. Population Output

In vivo, the local field potentials in the VTA contain slow oscillations that are thought to reflect synchronized population activity. Our MEA recordings are not able to provide the equivalent of such a signal and therefore we decided to construct and investigate an alternative Population Output signal based on the joint spike output of all the neurons in the slice. This Population Output signal should be able to indicate signs of underlying neuronal synchrony. The experiment that contained, under baseline conditions, the largest number of spiking VTA neurons (n = 14) was analyzed and the power spectrum was computed over a 300 s period. It exhibited a prominent peak oscillation at 1.9 Hz (**Figure 3**, red line); this frequency was only slightly higher than the mean of the oscillation frequencies calculated for the contributing neurons (1.70 (SE 0.11) Hz). The Population Output spectrum was statistically tested against the spectrum calculated from the shuffled data (**Figure 3**, blue line, gray band indicates the SD). The shuffling conserved the auto-correlation of the individual neurons, but broke the temporal relationship between the spike trains. The oscillation at 1.9 Hz was the only frequency where the baseline spectrum was significantly different from the shuffled spectrum. The sharply peaked Population Output spectrum suggested an appreciable degree of synchrony between the VTA dopamine neurons.

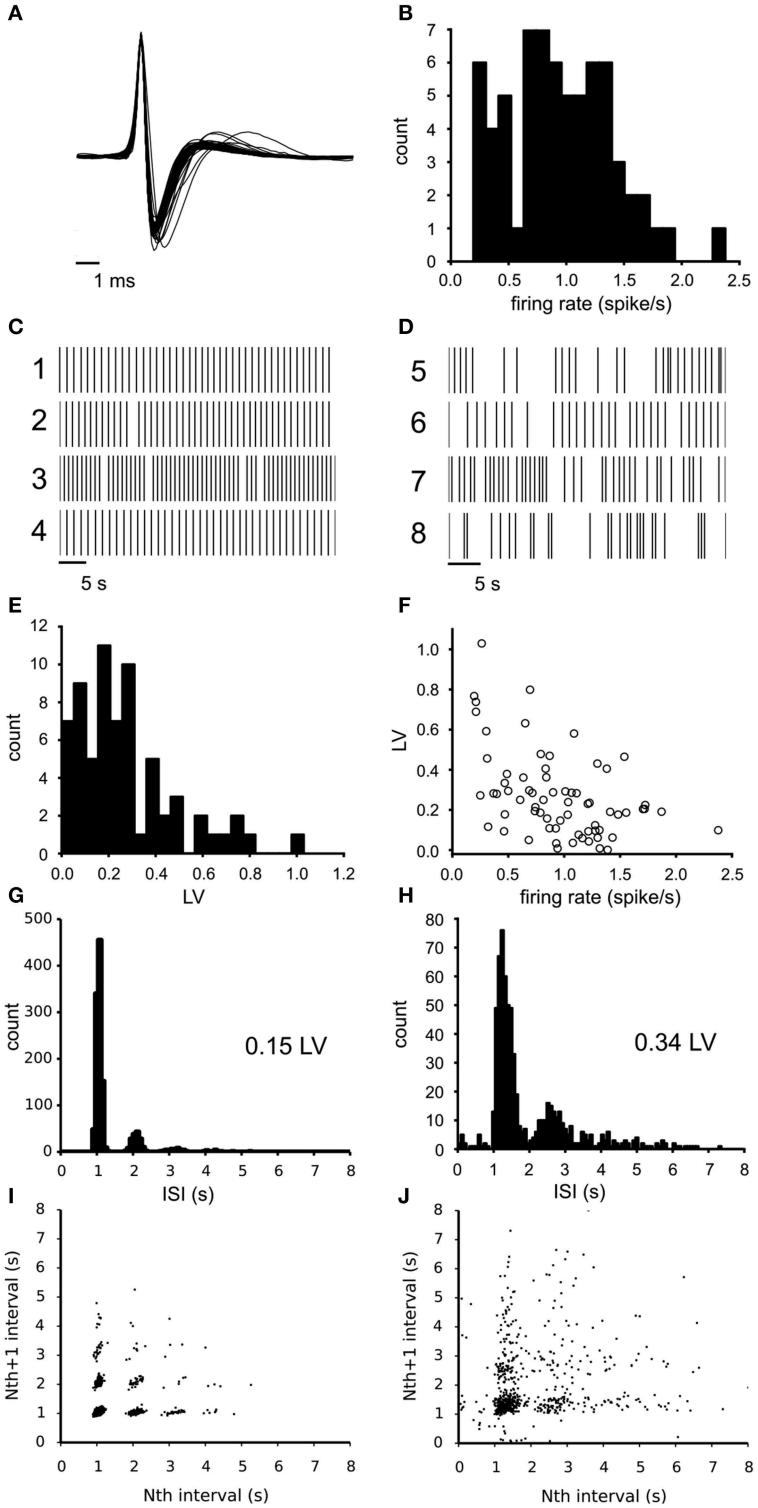

FIGURE 1 | Electrophysiological properties of the recorded VTA dopamine neurons during baseline conditions. **(A)** Superimposed traces of the mean spike waveform per neuron (normalized to the maximum amplitude and centered at the peak). All neurons exhibited the tri-phasic waveform typical of dopamine neurons. **(B)** Distribution of the mean firing rate of all VTA neurons recorded during baseline ($n = 68$). Mean firing rate is computed over a 25–30 min time period. **(C)** Typical firing pattern of four highly regular firing neurons (rasterplot of point processes). **(D)** Typical firing patterns of four irregularly firing neurons (rasterplot of point processes). **(E)** Distribution of LV (firing irregularity) for the same neurons as in **(A)**, determined over the same period. The distribution was not different from unimodal

(Continued)

FIGURE 1 | (Hartigan's diptest). **(F)** The relation between firing irregularity (LV) and the mean firing rate showed a negative correlation, indicating that neurons with a lower firing rate fired more irregularly. **(G)** Inter-Spike-Interval (ISI) distribution of a highly regular firing neuron, showing one main peak. **(H)** ISI distribution of a less regular firing neuron (broader main peak), which shows cycle skipping and also includes longer periods of silence. **(I)** return map of a highly regular firing neuron, which relates adjacent spike intervals. A relatively low number of spike cycle skipping (multiples of the preferred interval) was seen in combination with highly discrete point clouds. **(J)** Return map of an irregular firing neuron, which relates adjacent spike intervals. The more diffuse point clouds indicate a less stable oscillation frequency.

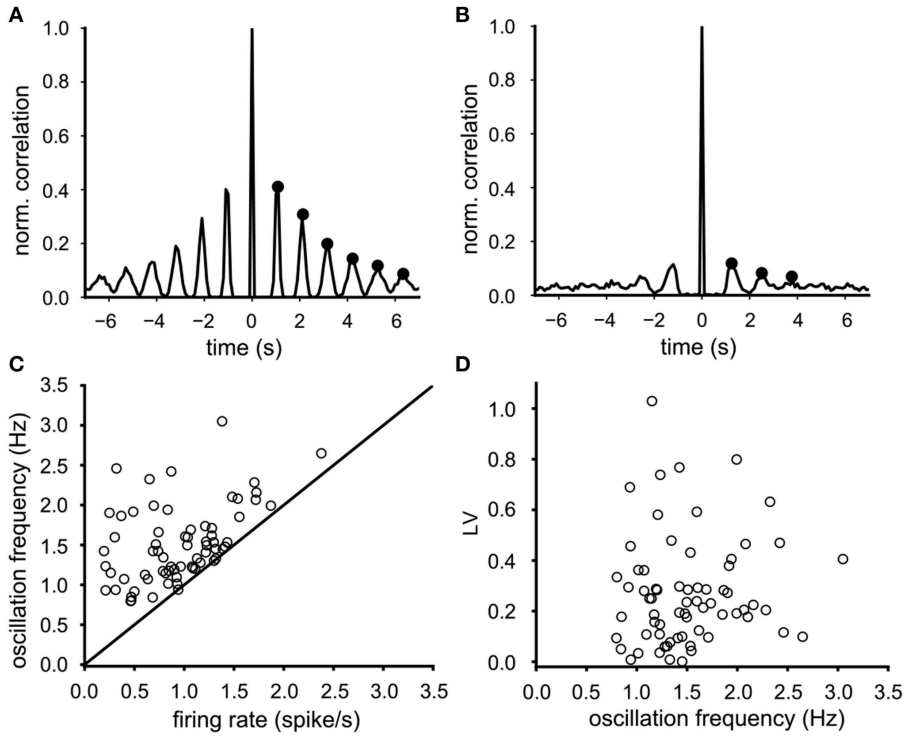

FIGURE 2 | Determination and properties of the oscillation frequency. **(A)** Typical auto-correlation function of a highly regular firing neuron ($LV = 0.008$), where the oscillation frequency (0.9 Hz) was computed from the six side-lobes (indicated by black dots). **(B)** Typical auto-correlation function of a less regular firing neuron ($LV = 0.34$), where the oscillation frequency (0.75 Hz) was computed from three side-lobes (indicated by black dots). **(C)** The oscillation frequency in relation to the mean firing rate for all 68 neurons (baseline condition). The oscillation frequency was either higher or equal to the mean firing rate. This difference was larger for neurons with lower firing rates. **(D)** the firing irregularity (LV) did not correlate with the oscillation frequency.

3.4. Baseline Functional Connectivity

In the MEA data the potential interaction between the VTA dopamine neurons was further quantified by the functional connectivity between all neuron pairs. The strength of this functional connectivity was measured by their phase coupling, which is based on the relative timing of the spikes in their spike trains. The Paired Phase Consistency (Vinck et al., 2010), was calculated in respect to the oscillation frequencies of the neurons in the pair and over a time period of 1,200-1,500 s. The PPC could be computed for 368 unique pairs of neurons in the recorded population. **Figure 4A** shows the distribution of the PPC values for the neuron pairs in the baseline condition (**Figure 4A**, black line). The mean value was larger than zero (2.3 (SE 0.2) $*10^{-3}$; t-test against 0, $p = 7.2 * 10^{-23}$). The same statistics were also performed on the shuffled controls (see methods). The PPC values of the shuffled pairs were tightly distributed around zero (green line, **Figure 4A**), indicating that the measured phase coupling originated from the combined temporal structure in the spike trains.

The statistics of the PPC of an individual pair could also be assessed through bootstrapping: 42 of the 368 (11%) of neuron pairs (involving 32 of the 68 neurons, distributed over all experiments) were significantly coupled (**Figure 4A**, filled green bars), when directly tested against their shuffled control trains. **Figure 4B** shows the PPC spectrum for an example neuron pair exhibiting significant coupling. A discrete peak was seen at the oscillation frequency of one of the two neurons (oscillation frequencies indicated by black dots, **Figure 4B**), indicating that the coupling was selective for the oscillation frequencies of the neurons.

3.5. Network Activity Modulation

The measured PPC can be considered an emerging property of the VTA network. We investigated the relation between the PPC and increased mean neuronal firing rate, using two different forms of excitability modulation. First, raising $[K^+]_o$ from a baseline level of 3.5 to 5 mM, increased the firing rate (4 experiments, 57 neurons under baseline condition). Second, bath

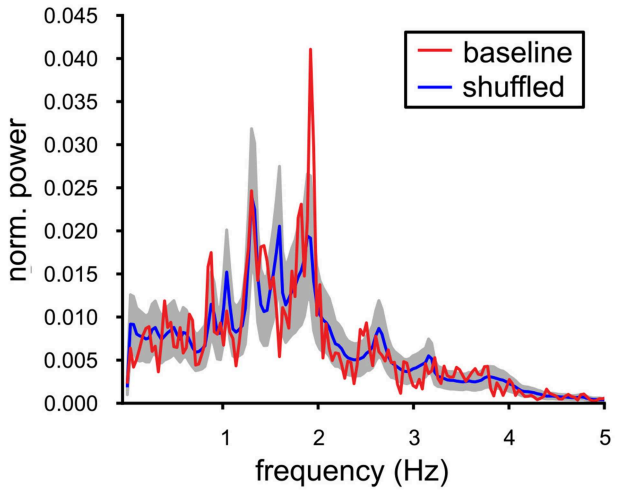

FIGURE 3 | Oscillations in the baseline VTA population output. Power spectrum of the population output of a recorded baseline VTA population (red line, single experiment: 14 neurons). It was contrasted with the power spectrum averaged over 1000 temporally shuffled populations (mean ± SD, resp. blue line and gray filling). The peak in the spectrum of the Population Output at 1.9 Hz was significantly different from the shuffled control and close to the mean oscillation frequency of the VTA neuron population [1.70(SE0.11)] Hz.

application of 300 μM glutamate increases firing rate to about the same level (5 experiments, 55 neurons under baseline condition). The enhanced excitability induced by both procedures was reversible, although some of the newly recruited neurons did not silence at wash-out. The time course of the mean firing rate of a VTA dopamine neuron population can be observed during a high potassium (**Figure 5A**) and a glutamate application (**Figure 5B**). Enhancing the potassium concentration to 5 mM, increased the mean firing rate of the VTA neurons by about 45% (**Figure 6A**) and the oscillation frequency by about 33% (**Figure 6C**) (statistics: ANOVA, $n = 57$, firing rate: $p = 3.4{*}10^{-6}$ oscillation frequency: $p = 1 * 10^{-6}$). Application of 300 μM glutamate induced similar changes in firing rate (**Figure 6B**) and oscillation frequency (**Figure 6D**) (statistics: ANOVA, $n = 55$, firing rate: $p = 5.8 * 10^{-5}$, oscillation frequency: $p = 0.001$). In addition to increasing the neuronal activity, both treatments recruited 22% (high potassium) and 16% (glutamate) additional neurons, shown in **Figures 6A,B** as neurons with a baseline firing rate set to zero. The neurons recruited by glutamate had an oscillation frequency higher than baseline [**Figure 6D**; ANOVA, $p = 0.027$, n = (55 baseline, 9 recruited)], whereas those recruited by high potassium had an oscillation frequency similar to baseline [**Figure 6C**, n = (57 baseline, 13 recruited)]. The increase in firing rate due to high potassium correlated with a decrease in LV (Spearman's $\rho = -0.35, p = 0.0074, n = 57$, **Figure 6E**), but this relationship was not present for the increase in firing rate induced by glutamate (Spearman's $\rho = -0.05, p = 0.71, n = 55$, **Figure 6F**). The two correlation coefficients were different ($\alpha = 0.05$, Fisher transform). These results indicated that, although the effect on the firing rate was quite similar for high potassium and glutamate, the manipulations had a differential effect on firing irregularity.

3.6. Population Output Modulation

The population output of the VTA neuronal population was compared under the two excitatory conditions. The analysis was performed on the slices with the largest number of neurons (19 neurons for high $[K^+]_o$ and 25 neurons for the 300 μM glutamate condition, active during the preceding baseline condition). The normalized power spectra were computed from the Population Output over a 20 min time period. **Figure 7** illustrates that the dominant peak in the spectrum shifts to a higher frequency for both the high potassium (**Figure 7A**) and glutamate condition (**Figure 7B**). These shifts confirmed the changes in mean oscillation frequency of the individual neurons (**Figures 6C,D**). The normalized Population Output spectra allowed to compare the shape of the dominant peak, but in order to investigate the consequences of increased excitability for the functional connectivity, we determined the pairwise functional connectivity.

3.7. Connectivity Modulation

The modulation of functional connectivity in the VTA network by high potassium or glutamate was analyzed by computing the PPC values of all pairs of neurons within the VTA populations for the conditions of increased excitability and their preceding baseline. The PPC was calculated over 1,200–1,500 s windows for the baseline as well as the excitatory condition. The cumulative PPC distribution for all neuron pairs in the high potassium condition compared to baseline, shows the shift to lower PPC values during high potassium (baseline PPC: 3.0 (SE 0.2) $*10^{-3}$, 442 pairs, 5 mM $[K^+]_o$ PPC: 2.3 (SE 0.2) $*10^{-3}$, 652 pairs, Kolmogorov-Smirnov, $p = 5.3 * 10^{-4}$, **Figure 8A**). Comparing the PPC distribution before and during glutamate application confirmed the absence of an effect on the PPC (baseline PPC: 2.4 (SE 0.2) $*10^{-3}$, 367 pairs, with glutamate PPC: 2.3(SE0.2) $* 10^{-3}$, 486 pairs, Kolmogorov-Smirnov, $p = 0.24$, **Figure 8B**). To assess the differential effect of glutamate and potassium on the PPC we performed multi-level analyses on neuron pairs active before and after the treatment (Field et al., 2012). First, we performed the analysis on the datasets of potassium and glutamate separately. PPC values of the neuron pairs, before and after treatment represented a repeated measure. The test included a fixed effect for the treatment and a random effect for the baseline PPC (across slices, potassium: 4 slices, glutamate: 5 slices). The latter controls for potential differences in baseline PPC values between slices. High potassium reduced the PPC of neuron pairs (δPPC: $-1.2(SE0.3) * 10^{-3}, n = 415, p < 0.01$). In contrast, glutamate treatment did not significantly affect the PPC within neuron pairs (δPPC: $-0.26(SE0.3) * 10^{-3}, n = 349, p = 0.28$). Adding a fixed effect for the substance type (potassium and glutamate) allowed us to determine the interaction between the type of substance and treatment (repeated measure). The multi-level analysis demonstrated a significant interaction ($p = 0.037$) between the substance type and treatment (repeated measure), substantiating the difference between potassium and glutamate. Similar results

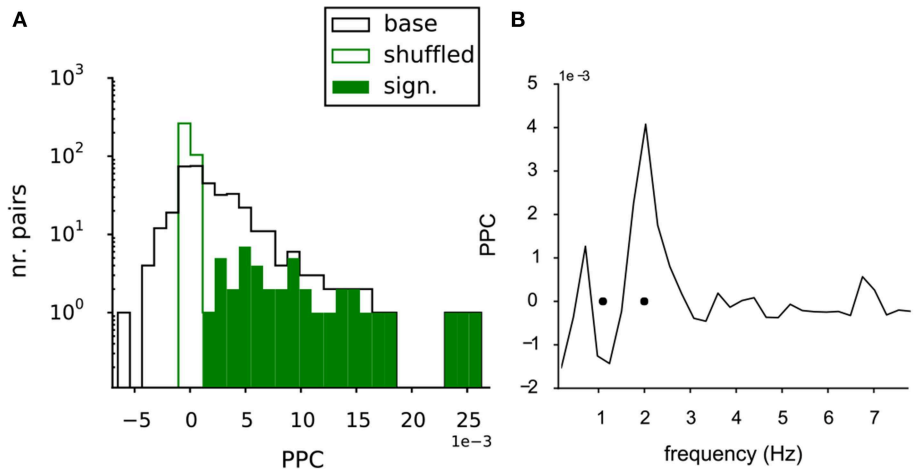

FIGURE 4 | Functional connectivity of the baseline VTA network. **(A)** The PPC values for all pairs within the baseline VTA networks (black line, 368 pairs from 6 experiments) had a mean PPC higher than zero, indicating significant functional coupling. The PPC values of the shuffled controls (green line) were centered at zero, showing the dependency of the phase coupling on the temporal structure in the neuronal spike trains. The significantly bootstrapped pairs (filled green bars, 42 pairs, 11% of total pairs containing 32 out of 68 neurons). **(B)** PPC spectrum for a significantly coupled neuron pair, showing selective coupling (peak in the spectrum) at one of the two oscillation frequencies of the neurons in the pair (black dots).

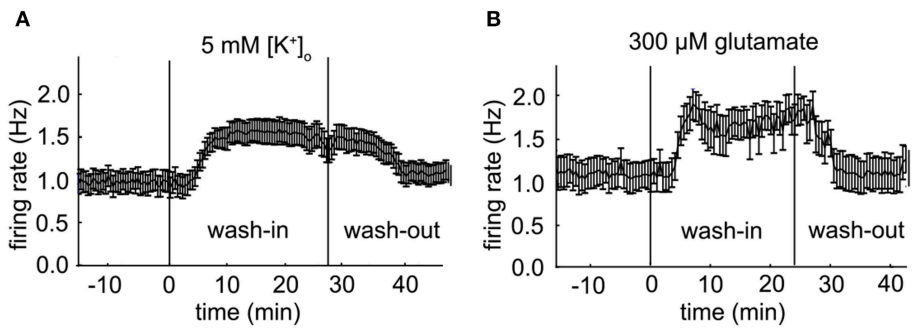

FIGURE 5 | Population firing rate of two VTA dopamine neuron populations. **(A)** Mean firing rate of a VTA dopamine neuron population during an administration of 5 mM extra-cellular potassium $[K^+]_o$ at 30 s time points (error bars indicate SE, baseline: n = 20; high potassium: n = 22). **(B)** Mean firing rate of a VTA dopamine neuron population during an administration of 300 µM glutamate at 30 s time points (error bars indicate SE, baseline: n = 10; glutamate: n = 12).

were obtained when we repeated the analysis controlling for random effects of PPC baseline values of slices and pairs. Both conditions recruited neurons due to their excitatory effect (high $[K^+]_o$: 22% and glutamate: 16%). These recruited neurons were engaged in new functional connections with each other and with the neurons already active under baseline conditions. **Figures 8C,D** contain bar plots of the PPC divided up for the pairs of neurons that were present in both the baseline and experimental conditions and the new pairs made with the recruited neurons. This division showed that the lower PPC value during high potassium was due to a reduction in PPC of the baseline network and not due to newly recruited neurons (ANOVA, $p = 2.3 * 10^{-5}, n = 415$, **Figure 8C**). The neurons recruited by glutamate generated higher PPC values compared to baseline (ANOVA, $p = 0.01, n = 349$, baseline, $n = 137$, **Figure 8D**). High potassium decreased the LV of the neurons active under baseline conditions (**Figure 6E**). Additionally, it reduced the PPC between these neurons (**Figure 8C**). We investigated the possible relation between the firing irregularity and PPC through mutual information analysis. The mutual information measured the strength of the association between the LV values of the neurons in a pair and their PPC value. It was computed for all neuron pairs during the baseline condition (n = 368). The LV values predicted 2.8 out of the 8.4 bits of total entropy in the PPC distribution after mutual information shuffle correction (Panzeri et al., 2007).

3.8. GABA-A Modulation

Around 20–30% of the neurons in the VTA are known to be GABAergic interneurons (Nair-Roberts et al., 2008). They could play a role in explaining the difference between the potassium and glutamate experiments. To assess their contribution to the observed effects on the PPC we used the GABA-A antagonist bicuculline. Bicuculline administration (20 µM) did not affect

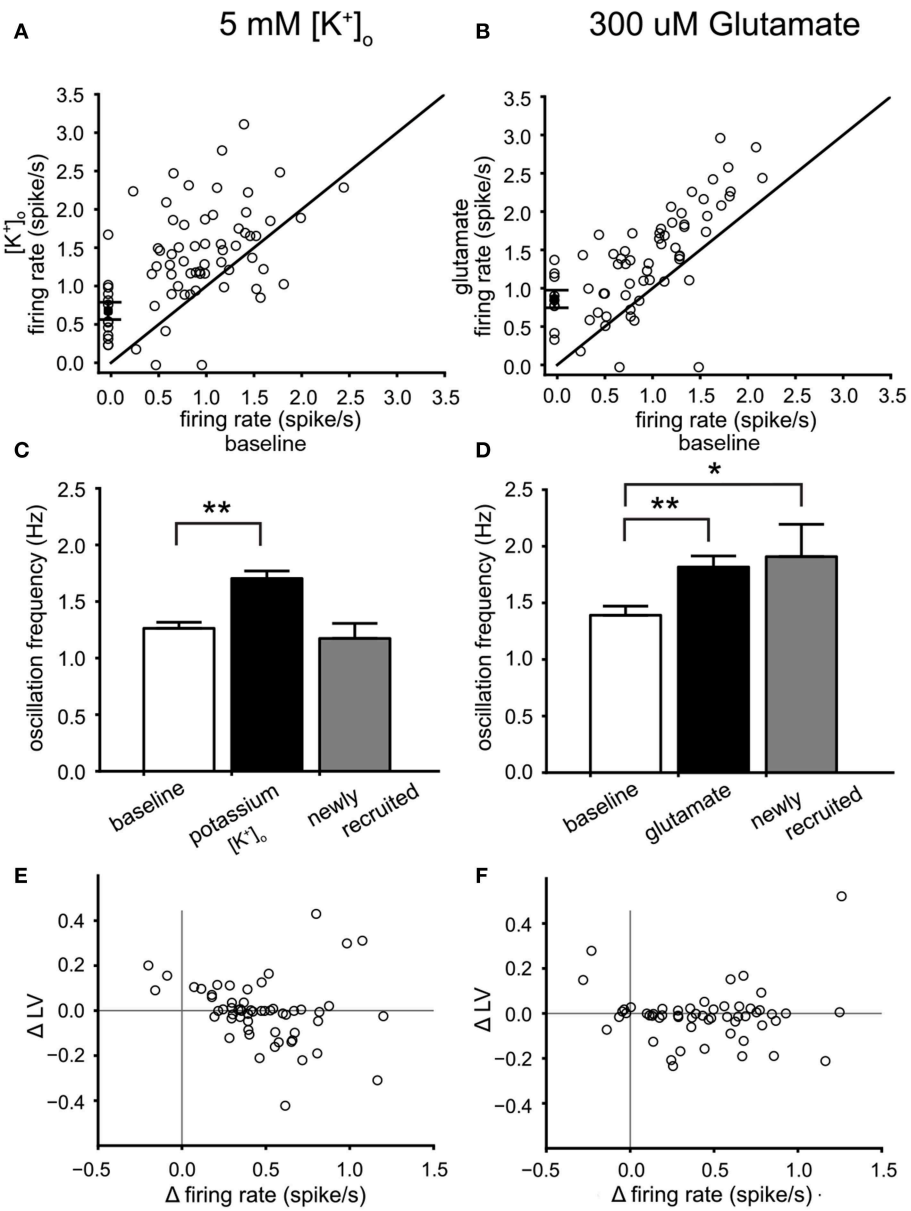

FIGURE 6 | Dopamine neuron activity modulation by high potassium and glutamate. **(A)** Relation between the mean firing rate during 5 mM $[K^+]_o$ and the mean firing rate at baseline. Newly recruited neurons are depicted with a zero mean firing rate at baseline. High potassium induced a 45% increase of the mean firing rate. **(B)** relation between the mean firing rate during 300 μM glutamate and the mean firing rate at baseline. Newly recruited neurons are depicted with a zero mean firing rate at baseline. Glutamate induced a 43% increase of the mean firing rate. **(C)** Comparison of the mean oscillation frequency of VTA neurons at baseline (white bar), with that of the same neurons during 5 mM $[K^+]_o$ (black bar) and the newly recruited neurons by 5 mM $[K^+]_o$ (22% recruited, gray bar). The mean oscillation frequency was increased during 5 mM $[K^+]_o$ for the neurons active during baseline. **(D)** Comparison of the mean oscillation frequency of VTA neurons at baseline (white bar), with that of the same neurons during 300 μM glutamate (black bar) and the newly recruited neurons by 300 μM glutamate (16% recruited, gray bar). Glutamate increased the mean oscillation frequency of the neurons present during baseline. In addition the recruited neurons also had an oscillation frequency higher than baseline. **(E)** The change in firing irregularity (LV) as a function of the change in mean firing rate induced by 5 mM $[K^+]_o$. An increase in mean firing rate correlated with lower firing irregularity. **(F)** The difference (300 μM glutamate minus baseline) in firing irregularity (LV) as a function of the difference in mean firing rate. A correlation was not found. $^*p < 0.05$, $^{**}p < 0.01$.

the mean firing rates of the dopamine neurons, averaged across four experiments (**Figure 9A**). The scatter plot confirms the absence of systematic changes of the neuronal firing rate (**Figure 9B**). Bicuculline (20 μM) also did not affect the PPC, when averaged over four experiments (Kolmogorov-Smirnov, $p = 0.72$, **Figure 9C**).

4. DISCUSSION

In this study we quantify the PPC between pairs of neurons and interpret it as a measure of functional connectivity; it appears as an emerging property during excitability modulation of dopaminergic neurons in acute brain slices of the VTA.

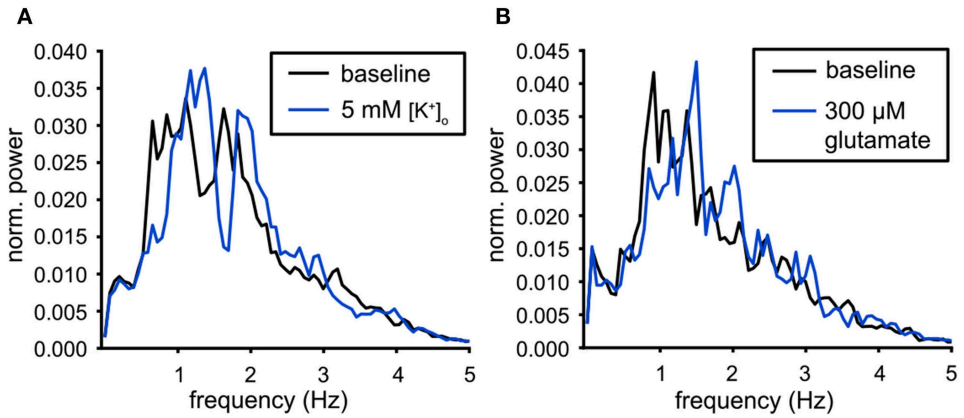

FIGURE 7 | Modulation of the population output. **(A)** Power spectra of the population output for baseline conditions (black) and for high potassium (5 mM $[K^+]_O$, blue). The dominant peak was shifted to a higher frequency during the administration of 5 mM $[K^+]_O$ (baseline: $n = 19$, 5 mM $[K^+]_O$: $n = 25$). **(B)** Power spectra of the population output for baseline conditions (black) and for 300 µM glutamate (blue). The dominant peak was shifted to a higher frequency during the administration of glutamate (baseline: $n = 22$, glutamate: $n = 26$).

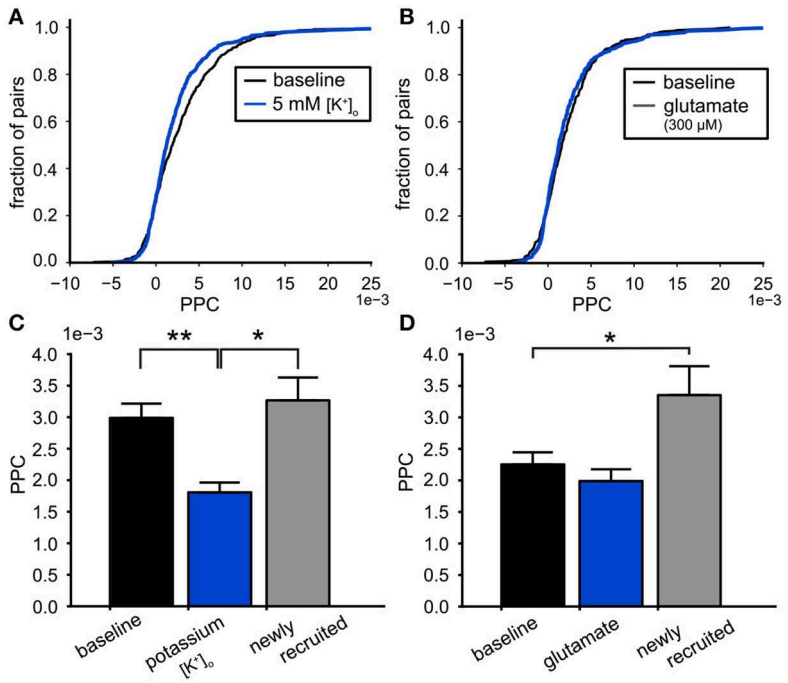

FIGURE 8 | Effects on functional connectivity induced by high potassium and glutamate. **(A)** The cumulative PPC for the high potassium (5 mM $[K^+]_O$) and baseline conditions. The leftward shift of the distribution shows that the pairwise connectivity was weakened during the high potassium condition (baseline: 442 pairs, high potassium: 652 pairs). **(B)** The cumulative PPC histograms for the baseline and 300 µM glutamate conditions. The two distributions were identical, indicating that the network strength was similar for both conditions (baseline: 367 pairs, glutamate: 486 pairs). **(C)** Comparison of the mean PPC for neuron pairs during baseline conditions (black bar, $n = 415$), with the same pairs during 5 mM $[K^+]_O$ (blue bar, $n = 415$) and the connections made by neurons newly recruited by 5 mM $[K^+]_O$ (gray bar, $n = 237$). The weakening of the functional connectivity by high potassium originated from the functional connections already present during baseline. **(D)** Comparison of the mean PPC for neuron pairs during baseline conditions (black bar, $n = 349$), with the same pairs during 300 µM glutamate administration (blue bar, $n = 349$) and the connections made by neurons newly recruited by 300 µM glutamate (gray bar, $n = 137$). The functional connections already present during baseline were not affected by glutamate (black), however the connections made by the recruited neurons were stronger than baseline (gray). $^*p < 0.05$, $^{**}p < 0.01$.

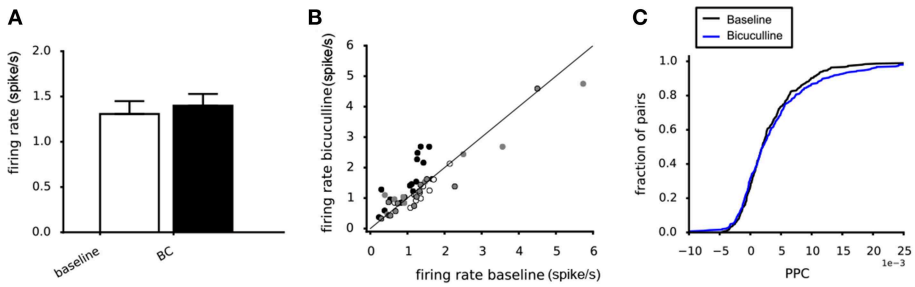

FIGURE 9 | Effect of Bicuculline on the dopamine neuron population activity and connectivity. **(A)** The mean firing rate was not affected by 20 μM Bicuculline, when averaging across four experiments (white vs. black bar). **(B)** Relation between the mean firing rate during 20 μM Bicuculline and the mean firing rate at baseline for four experiments (filled and non-filled dots). A systematic response to Bicuculline was not found. **(C)** The cumulative PPC histograms for the baseline and 20 μM bicuculline conditions. The two distributions were not different, indicating that bicuculline had no effect on the network strength when averaged over four experiments (baseline: 305 pairs, bicuculline: 305 pairs).

Two manipulations that induced a similar increase of the mean neuronal firing rate, had a differential effect on the PPC. This could make the organization of network synchrony as captured in the PPC a unique element of the population message that is send to the projection areas.

VTA dopamine neurons possess a strong intrinsic rhythm generator, illustrated by the spontaneous oscillating firing patterns in the isolated slice. These rhythms even persist in dissociated single neurons (Koyama et al., 2005). It has been hypothesized that the VTA circuit could generate a low frequency rhythm important for memory, which entrains the prefrontal cortex and the hippocampus (Fujisawa and Buzsáki, 2011). The functional connectivity local to the VTA that we find can be an important factor in generating such a rhythm. The mean firing rate of all dopaminergic neurons recorded in our preparation (∼1 Hz) is lower than the one reported *in vivo*: 5 (SD 6) Hz (Fujisawa and Buzsáki, 2011). The same holds for the mean oscillation frequency (∼2 Hz in our experiments) vs. the oscillations in the field potential reported *in vivo*: 4 Hz. We carefully restricted our recordings to the "classical" lateral part of the VTA that contains an uniform class of mesolimbic projection neurons (Lammel et al., 2008). All recorded neurons were identified as principal neurons based on their electrophysiological properties: characteristic broad spike waveform and regular low frequency firing pattern. In addition we could unambiguously link these properties to a positive response to the D_2-receptor agonist quinpirole in all tested neurons. Although the accuracy of these markers has been debated, our findings correspond to the abundance (∼60%) of these neurons in the VTA (Margolis et al., 2006). The VTA also contains glutamatergic and GABAergic neurons that can make up 20 to 30% of the population (Nair-Roberts et al., 2008). They are typically faster spiking (> 10 Hz) neurons and we never picked up more than one such unit per slice; too few to incorporate them in a statistical sound way into this study. The acute brain slice is very suitable for work with the Micro Electrode Array, but has limitations. In the slicing process we may have lost quite a few connections and the preparation is devoid of its normal background input. The first aspect will lead to an underestimation of the functional connectivity, while the second one brings a careful analysis within reach. The MEA has only two dimensional electrodes with a diameter of 30 μM and spacing of 100 μm, which inevitably leads to sparse sampling of the population; nevertheless we can simultaneously record ensembles of 10–30 spiking dopamine neurons.

Most of our VTA dopaminergic neurons had a distinct firing rate and LV, but these values varied considerably over the population, even within the same slice. Our distribution of firing rate and LV was, however, unimodal in contrast to the observations made in acutely dissociated neurons, where two distinct populations have been described (Koyama et al., 2005). The assumption of an underlying oscillatory process that triggers the action potentials, predicts that the auto-correlation of the spike train is a better estimator of the oscillation frequency than the firing rate, which we confirmed. The LV is determined by local irregularity of consecutive spikes and by cycle skipping. This is in accordance with the finding that a lower baseline firing rate correlated with a higher firing irregularity, but that this correlation did not exist between the oscillation frequency and the irregularity. Rhythmic action potential firing in VTA neurons originates from a balance between an inward sodium current and a calcium-based slow oscillatory potential (SOP) (Drion et al., 2011). There is evidence that sodium currents form the dominant mechanism and generate quite regular spike firing in the VTA (Khaliq and Bean, 2010; Drion et al., 2011), while in dopamine neurons from the substantia nigra spikes are generated predominantly by the less regular SOP (Putzier et al., 2009; Drion et al., 2011).

We measured the strength of the synchrony between all possible pairs of VTA neurons in the same slice with the Paired Phase Consistency (PPC) (Lachaux et al., 1999; Vinck et al., 2010). The mean value of the PPC of the baseline VTA network was different from zero which indicates the presence of significant functional coupling in the network. When tested individually, 42 out of 382 all possible pairs demonstrated significant coupling at their oscillation frequency (∼2 Hz). Of the 68 neurons sampled in all experiments,

32 were part of a pair with significant coupling, which shows that the VTA contains a collection of functionally connected dopamine neurons that synchronizes at the time scale of their preferred oscillation frequency. The details of the underlying mechanism of this coupling cannot be deduced from the current measurements alone. Direct synaptic connectivity as demonstrated for dopaminergic neurons in the Substantia Nigra (Vandecasteele et al., 2008) could be involved as could be electrical coupling through gap junctions. However in a previous study (van der Velden et al., 2017) we concluded that probably the best explation of the apparent synchrony in the VTA network is through a mechanism that is called dopamine volume transmission (Zoli et al., 1998). In volume transmission the collective and synchronized release of dopamine in the extracellular space creates an oscillating dopamine concentration, where the precize role of each neuron depends on its cellular sensitivity for dopamine. In the current paper we refer to our form of coupling as functional connectivity in contrast to the more straigtforward anatomical/synaptic connectivity.

Most pharmacological studies of the VTA have focused on describing modulations of individual neuronal firing rates (Hand et al., 1987; Wang and French, 1993; Werkman et al., 2001). The MEA recordings allow us to incorporate functional connectivity into this modulation. Here we used two well understood manipulations of neuronal activity: increasing extracellular potassium from 3.5 mM to 5 mM and bath-application of 300 μM glutamate, to study the relation between neuronal activity and functional connectivity. The manipulations were chosen so that their effect on mean firing rate and mean oscillation frequency was similar. Interestingly, they had a differential effect on firing irregularity (LV) as well as on functional connectivity as measured with the PPC. The increase in firing rate induced by high potassium correlated with a lower LV, while such correlation was absent in the case of glutamate application. The increased activity of individual neurons propagates through the population level, as high potassium and glutamate both increased the oscillation frequency of the population output. Calculating the connectivity strength using the PPC demonstrated that the functional connectivity was weakened by high potassium, while it was unaffected by glutamate. The weakening of the functional connectivity originated from neurons already present during the baseline condition and not from connections made by neurons recruited by high potassium. Depolarization of the membrane potential by high potassium can strengthen the role of the 'persistent' sodium current in the generation of spikes (Khaliq and Bean, 2010) and this can explain the reduction in firing irregularity (Drion et al., 2011). Glutamate depolarizes the neuron through the activation of post-synaptic AMPA and NMDA receptors (Wang and French, 1993). The strongly increased synaptic activity enhances firing irregularity (Drion et al., 2011) and could therefore explain the difference between the effect of glutamate- and potassium-induced depolarization. The strengthening of the intrinsic rhythm by larger sodium currents could make the dopamine neurons less sensitive to synchronizing inputs and thus show up as weaker neuronal interactions, based on resonance principles (Hunter et al., 1998; Coombes and Bressloff, 1999). Neurons recruited by glutamate had a higher than baseline oscillation frequency and functional connectivity, suggesting that recruitment through increased synaptic input (glutamate) leads to more network participation than recruitment through a direct increase of the membrane potential (high potassium).

GABAergic transmission contributed on average very little to the functional connectivity. Our activity modulation demonstrates that physiological relevant stimuli (high potassium and glutamate) can alter the functional connectivity of the local VTA network. This effect seems independent from the modulation of the neuronal firing rate. The difference between the effects on the functional connectivity by high potassium and glutamate, indicates that the specific mechanisms with which these substances excite the individual dopamine neurons (electro-chemical vs. synaptic) is of importance for the response at the network level. Functional connectivity shows up as an aspect of spatial synchronization that is present in the output of the VTA and can thus form a unique element of the message that is sent to the projection areas. Functional connectivity between VTA dopamine neurons is involved in reward processing (Kim et al., 2012; Moghaddam et al., 2017) and we analyze this connectivity in detail. Functional connectivity adds a dimension to pharmacological manipulation of the VTA micro circuit and could lead to a better understanding of pharmacological (side) effects of e.g., anti-psychotic drugs on the mesolimbic and mesocortical projection areas.

DATA AVAILABILITY

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

ETHICS STATEMENT

All experiments and methods were approved by the ethical committee for animal experimentation of the University of Amsterdam.

AUTHOR CONTRIBUTIONS

LvdV, MV, TW, and WW: designed the study. LvdV, MV, and WW: contributed to the data analysis and writing of the manuscript. LvdV: performed the experiments and the data analysis.

FUNDING

This work was supported by grant 114000091 to WW from the Platform alternatief voor dierproeven of the Dutch organization for fundamental research.

REFERENCES

Ardid, S., Vinck, M., Kaping, D., Marquez, S., Everling, S., and Womelsdorf, T. (2015). Mapping of functionally characterized cell classes onto canonical circuit operations in primate prefrontal cortex. *J. Neurosci.* 35, 2975–2991. doi: 10.1523/JNEUROSCI.2700-14.2015

Bayer, H. M., Lau, B., and Glimcher, P. W. (2007). Statistics of midbrain dopamine neuron spike trains in the awake primate. *J. Neurophysiol.* 98, 1428–1439. doi: 10.1152/jn.01140.2006

Bayer, V. E., and Pickel, V. M. (1990). Ultrastructural localization of tyrosine hydroxylase in the rat ventral tegmental area: relationship between immunolabeling density and neuronal associations. *J. Neurosci.* 10, 2996–3013. doi: 10.1523/JNEUROSCI.10-09-02996.1990

Björklund, A., and Dunnett, S. B. (2007). Dopamine neuron systems in the brain: an update. *Trends Neurosci.* 30, 194–202. doi: 10.1016/j.tins.2007.03.006

Bowery, B., Rothwell, L. a., and Seabrook, G. R. (1994). Comparison between the pharmacology of dopamine receptors mediating the inhibition of cell firing in rat brain slices through the *Substantia nigra* pars compacta and ventral tegmental area. *Brit. J. Pharmacol.* 112, 873–880. doi: 10.1111/j.1476-5381.1994.tb13161.x

Coombes, S., and Bressloff, P. C. (1999). Mode locking and Arnold tongues in integrate-and-fire neural oscillators. *Phys. Rev. E* 60(2 Pt B):2086–2096. doi: 10.1103/PhysRevE.60.2086

Drion, G., Massotte, L., Sepulchre, R., and Seutin, V. (2011). How modeling can reconcile apparently discrepant experimental results: the case of pacemaking in dopaminergic neurons. *PLoS Comput. Biol.* 7:e1002050. doi: 10.1371/journal.pcbi.1002050

Field, A., Miles, J., and Field, Z. (2012). *Discovering Statistics Using R*. Thousand Oaks, CA: SAGE Publications.

Fields, H. L., Hjelmstad, G. O., Margolis, E. B., and Nicola, S. M. (2007). Ventral tegmental area neurons in learned appetitive behavior and positive reinforcement. *Annu. Rev. Neurosci.* 30, 289–316. doi: 10.1146/annurev.neuro.30.051606.094341

Fujisawa, S., and Buzsáki, G. (2011). A 4 Hz oscillation adaptively synchronizes prefrontal, VTA, and hippocampal activities. *Neuron* 72, 153–165. doi: 10.1016/j.neuron.2011.08.018

Gonon, F. (1988). Nonlinear relationship between impulse flow and dopamine released by rat midbrain dopaminergic neurons as studied by *in vivo* electrochemistry. *Neuroscience* 24, 19–28. doi: 10.1016/0306-4522(88)90307-7

Grace, A. A., and Onn, S. p. (1989). Morphology and electrophysiological properties of lmmunocytochemically identified rat dopamine neurons recorded in wifro. *J. Neurosci.* 9, 3463–3481. doi: 10.1523/JNEUROSCI.09-10-03463.1989

Hand, T. H., Hu, X. T., and Wang, R. Y. (1987). Differential effects of acute clozapine and haloperidol on the activity of ventral tegmental (A10) and nigrostriatal (A9) dopamine neurons. *Brain Res.* 415, 257–269. doi: 10.1016/0006-8993(87)90207-1

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Comput. Sci. Eng.* 9, 90–95. doi: 10.1109/MCSE.2007.55

Hunter, J. D., Milton, J. G., Thomas, P. J., and Cowan, J. D. (1998). Resonance effect for neural spike time reliability. *J. Neurophysiol.* 80, 1427–1438. doi: 10.1152/jn.1998.80.3.1427

Ince, R. A. A., Petersen, R. S., Swan, D. C., and Panzeri, S. (2009). Python for information theoretic analysis of neural data. *Front. Neuroinformatics* 3:4. doi: 10.3389/neuro.11.004.2009

Khaliq, Z. M., and Bean, B. P. (2010). Pacemaking in dopaminergic ventral tegmental area neurons: depolarizing drive from background and voltage-dependent sodium conductances. *J. Neurosci.* 30, 7401–7413. doi: 10.1523/JNEUROSCI.0143-10.2010

Kim, Y., Wood, J., and Moghaddam, B. (2012). Coordinated activity of ventral tegmental neurons adapts to appetitive and aversive learning. *PloS ONE* 7:e29766. doi: 10.1371/journal.pone.0029766

Koyama, S., Kanemitsu, Y., and Weight, F. F. (2005). Spontaneous activity and properties of two types of principal neurons from the ventral tegmental area of rat. *J. Neurophysiol.* 93, 3282–3293. doi: 10.1152/jn.00776.2004

Lachaux, J. P., Rodriguez, E., Martinerie, J., and Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* 8, 194–208.

Lammel, S., Hetzel, A., Häckel, O., Jones, I., Liss, B., and Roeper, J. (2008). Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron* 57, 760–773. doi: 10.1016/j.neuron.2008.01.022

Lisman, J. E., and Grace, A. A. (2005). The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* 46, 703–713. doi: 10.1016/j.neuron.2005.05.002

MacKay, D. J. C. (2003). *Information Theory, Inference and Learning Algorithms*. Cambridge, UK; New York, NY: Cambridge University Press.

Margolis, E. B., Lock, H., Hjelmstad, G. O., and Fields, H. L. (2006). The ventral tegmental area revisited: is there an electrophysiological marker for dopaminergic neurons? *J. Physiol.* 577(Pt 3):907–924. doi: 10.1113/jphysiol.2006.117069

Moghaddam, B., Koerner, F. S., Kass, R. E., Wood, J., and Simon, N. W. (2017). Networks of VTA neurons encode real-time information about uncertain numbers of actions executed to earn a reward. *Front. Behav. Neurosci.* 11, 1–16. doi: 10.3389/fnbeh.2017.00140

Nair-Roberts, R. G., Chatelain-Badie, S. D., Benson, E., White-Cooper, H., Bolam, J. P., Ungless, M. A., et al. (2008). Stereological estimates of dopaminergic, GABAergic and glutamatergic neurons in the ventral tegmental area, substantia nigra and retrorubral field in the rat. *Neuroscience* 152, 1024–1031. doi: 10.1016/j.neuroscience.2008.01.046

Olivier, M., Fejtl, M., Raggenbass, M., Bertrand, D., Renaud, P., Heuschkel, M. O., et al. (2002). A three-dimensional multi-electrode array for multi-site stimulation and recording in acute brain slices. *J. Neurosci. Methods* 114, 135–148. doi: 10.1016/S0165-0270(01)00514-3

Omelchenko, N., and Sesack, S. R. (2009). Ultrastructural analysis of local collaterals of rat ventral tegmental area neurons: GABA phenotype and synapses onto dopamine and GABA cells. *Synapse* 63, 895–906. doi: 10.1002/syn.20668

Paladini, C. A., and Roeper, J. (2014). Generating bursts (and pauses) in the dopamine midbrain neurons. *Neuroscience* 282C, 109–121. doi: 10.1016/j.neuroscience.2014.07.032

Panzeri, S., Senatore, R., Montemurro, M. A., and Petersen, R. S. (2007). Correcting for the sampling bias problem in spike train information measures. *J. Neurophysiol.* 98, 1064–1072. doi: 10.1152/jn.00559.2007

Pucak, M. L., and Grace, A. A. (1994). Evidence that systemically administered dopamine antagonists activate dopamine neuron firing primarily blockade of somatodendritic. *J. Pharmacol. Exp. Therapeut.* 271, 1181–1192.

Pucak, M. L., and Grace, A. a. (1996). Effects of haloperidol on the activity and membrane physiology of substantia nigra dopamine neurons recorded *in vitro*. *Brain Res.* 713, 44–52. doi: 10.1016/0006-8993(95)01460-8

Putzier, I., Kullmann, P. H. M., Horn, J. P., and Levitan, E. S. (2009). Cav1.3 channel voltage dependence, not Ca2+ selectivity, drives pacemaker activity and amplifies bursts in nigral dopamine neurons. *J. Neurosci.* 29, 15414–15419. doi: 10.1523/JNEUROSCI.4742-09.2009

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

Shinomoto, S., Miura, K., and Koyama, S. (2005). A measure of local variation of inter-spike intervals. *Bio Systems* 79, 67–72. doi: 10.1016/j.biosystems.2004.09.023

Taketani, M., and Baudry, M. (2010). *Advances in Network Electrophysiology*. New York, NY: Springer.

Traub, R. D., Miles, R., and Wong, R. K. (1989). Model of the origin of rhythmic population oscillations in the hippocampal slice. *Science* 243, 1319–1325. doi: 10.1126/science.2646715

van der Velden, L., Vinck, M., Werkman, T. R., and Wadman, W. J. (2017). Tuning of neuronal interactions in the lateral ventral tegmental area by dopamine sensitivity. *Neuroscience* 366, 62–69. doi: 10.1016/j.neuroscience.2017.10.009

Vandecasteele, M., Glowinski, J., Deniau, J. M., and Venance, L. (2008). Chemical transmission between dopaminergic neuron pairs. *Proc. Natl. Acad. Sci. U.S.A.* 105, 4904–4909. doi: 10.1073/pnas.0703121105

Vinck, M., van Wingerden, M., Womelsdorf, T., Fries, P., and Pennartz, C. M. A. (2010). The pairwise phase consistency: a bias-free measure

of rhythmic neuronal synchronization. *NeuroImage* 51, 112–122. doi: 10.1016/j.neuroimage.2010.01.073

Wang, T., and French, E. D. (1993). L-glutamate excitation of A10 dopamine neurons is preferentially mediated by activation of NMDA receptors: extra- and intracellular electrophysiological studies in brain slices. *Brain Res.* 627, 299–306. doi: 10.1016/0006-8993(93)90334-J

Welch, P. D. (1967). The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* 15, 70–73. doi: 10.1109/TAU.1967.1161901

Werkman, T. R., Kruse, C. G., Nievelstein, H., Long, S. K., and Wadman, W. J. (2001). *In vitro* modulation of the firing rate of dopamine neurons in the rat substantia nigra pars compacta and the ventral tegmental area by antipsychotic drugs. *Neuropharmacology* 40, 927–936. doi: 10.1016/S0028-3908(01)00015-6

Zoli, M., Torri, C., Ferrari, R., Jansson, A., Zini, I., Fuxe, K., et al. (1998). The emergence of the volume transmission concept. *Brain Res. Rev.* 26, 136–147. doi: 10.1016/S0165-0173(97)00048-9

Conflict of Interest Statement: 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.

Copyright © 2019 van der Velden, Vinck, Werkman and Wadman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 11 July 2019
doi: 10.3389/fncom.2019.00047

The Orientation Selectivity of Spike-LFP Synchronization in Macaque V1 and V4

Zhaohui Li[1,2]*, *Mengyu Gao*[1,2] *and Yongtian Wang*[1,2]

[1] School of Information Science and Engineering, Yanshan University, Qinhuangdao, China, [2] Hebei Key Laboratory of Information Transmission and Signal Processing, Yanshan University, Qinhuangdao, China

Orientation selectivity is a fundamental property of visual cortical neurons and plays a crucial role in pattern perception. Although many studies have dedicated to explain how the orientation selectivity emerged, the mechanism underlying orientation selectivity is still not clear. In this work, we investigated the synchronization between spikes and local field potentials (LFP) in gamma band, with the aim of providing a new avenue to analyze the orientation selectivity. The experimental data were recorded by utilizing two chronically implanted multi-electrode arrays, where each array consisted of 48 electrodes and was placed over V1 and V4, respectively, in two macaques performing a selective visual attention task. An unbiased and robust measure for quantifying the synchronization between spikes and LFP was employed in the analysis process, which is termed as spike-triggered correlation matrix synchronization (SCMS) and performs reliably for limited samples of data. We observed the spike-LFP synchronization in three cases, i.e., spikes and LFP in V1, spikes and LFP in V4, spikes in V4 and LFP in V1. From the orientation tuning curves based on the spike-LFP synchronization, it is found that there is a strong correlation between the synchronization and grating orientation. The neurons in both V1 and V4 exhibit orientation selectivity, but V1 is stronger. In addition, the spike-LFP synchronization strength between V1 and V4 also shows orientation selectivity to drifting gratings. It means that the synchronization not only reflects the basic features of visual stimulation, but also describes the orientation tuning characteristics of neurons in different regions. Our results suggest that the spike-LFP synchronization can be used as an alternative and effective method to study the mechanism for generating orientation selectivity of visual neurons.

Keywords: orientation selectivity, multi-electrode array, spike, local field potential, synchronization

OPEN ACCESS

Edited by:
Andreas Bahmer, Universitätsklinikum Würzburg, Germany

Reviewed by:
Paul H. E. Tiesinga, Radboud University Nijmegen, Netherlands
Conrado Arturo Bosman, University of Amsterdam, Netherlands

***Correspondence:**
Zhaohui Li
lizhaohui@ysu.edu.cn

Received: *15 January 2019*
Accepted: *26 June 2019*
Published: *11 July 2019*

Citation:
Li Z, Gao M and Wang Y (2019) The Orientation Selectivity of Spike-LFP Synchronization in Macaque V1 and V4. Front. Comput. Neurosci. 13:47. doi: 10.3389/fncom.2019.00047

INTRODUCTION

The orientation is a basic feature of natural images. The orientation selective response of visual cortical neurons to the object boundary plays a key role in the shape perception and other perception processes (Hubel and Wiesel, 1962; Mansfield, 1974; Girshick et al., 2011; Durant et al., 2017). In the past decades, many studies have been devoted to explore the mechanism for generating orientation selectivity. Generally, two important types of signal were employed in these analyses, i.e., spikes (action potentials) and local field potentials (LFP), which were simultaneously

recorded from the visual cortex by multiple electrode arrays (Zhang and Li, 2013; Bharmauria et al., 2016). On the one hand, the spikes are identified by high-pass filtering, detection and sorting, indicating the firing activities of individual neurons. Its firing rate has been widely used since the orientation selectivity is discovered, e.g., some researchers found that the spike firing rate has different response values under different orientation and contrast stimuli (Hubel and Wiesel, 1962; Anderson et al., 2000; McLaughlin et al., 2000; Manyakov and Van Hulle, 2010). It indicates that the neuronal discharge activity is able to encode the orientation information of visual images. On the other hand, the LFP is obtained by low-pass filtering the original wideband signal, representing the synaptic activities of local populations of cortical neurons (Buzsaki et al., 2012; Gaucher et al., 2012). Because the spike firing rate cannot reflect the synaptic activities of multiple neurons in a local region, the LFP frequency or energy is adopted in many studies. For example, the high frequency oscillations of LFP in the striate cortex of awake monkeys showed stronger orientation selectivity than low frequency oscillations (Frien et al., 2000), and the energy variation of the LFP in gamma band was able to effectively encode the stimuli (such as time, frequency, orientation, etc.) in images (Siegel and Konig, 2003; Henrie and Shapley, 2005; Ince et al., 2012).

In recent years, a large number of neurophysiological studies have shown that there is a close relationship between spike and LFP gamma band (Ray and Maunsell, 2011a,b; Li et al., 2014). Combining these two signals to decode the behaviors can provide more information than using one signal separately (Mehring et al., 2003; Mollazadeh et al., 2009), which means that, it is able to provide a comprehensive description about the neural mechanism of signal processing. Moreover, it has been shown that the spike and LFP both participate in the coding of visual information (Quian Quiroga and Panzeri, 2009; Perge et al., 2014). Thus, we think relating the spike-LFP correlation and orientation is an effective tool to investigate the orientation selectivity. Therefore, we used the SCMS method (Li et al., 2016) to estimate the spike-LFP synchronization of the data which was obtained by simultaneously implanting two multi-electrode arrays in V1 and V4 of visual cortex, respectively.

MATERIALS AND METHODS

Experiment Procedure

The experimental data was recorded from two male rhesus monkeys. All procedures were conducted in compliance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals, and were approved by the Institutional Animal Care and Use Committee of Beijing Normal University. Under general anesthesia induced with ketamine (10 mg/kg) and maintained with isoflurane (1.5–2.0%), a titanium post was attached to the skull with bone screws for immobilizing the animal's head during behavioral training. After the monkeys had been trained in a simple fixation task, two 6 × 8 multi-electrode arrays (with interelectrode spacing of 0.4 mm, electrode length of 0.5–0.6 mm, and typical electrode impedances of a few hundred kiloohms; Blackrock Microsystems) were implanted intoV1 and V4, respectively. LFP and spike data were recorded at 10 kHz using a 128-channel Cerebus neural electrophysiological signal recording system (Blackrock Microsystems).

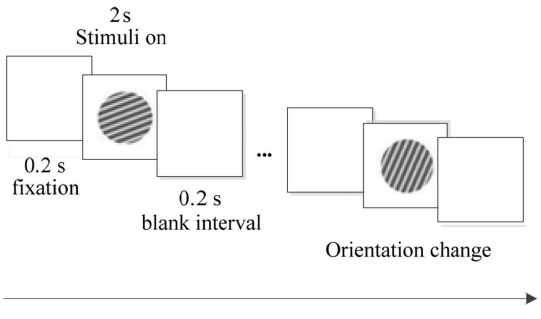

FIGURE 1 | The experimental procedure. Each trial was performed for 2.4 s, including three phases: fixation in the first 0.2 s, stimulus in the next 2 s, and a blank interval in the final 0.2 s.

Visual Stimulation

The visual stimulation in the experiment were generated by a stimulus generator system ViSaGe and displayed on a 22-inch CRT monitor at a viewing distance of 100 cm. The stimulus patterns were drifting sinusoidal gratings of different orientations which were displayed within a circular patch of 4° visual angle in diameter, covering the visual field locations of all recording sites. The orientation of the sinusoidal grating used in the experiment was uniformly distributed between 0° and 360° in steps of 22.5°. Other stimulus settings were identical in the whole experiment, including the contrast of 99%, the spatial frequency of 2 cycle/degree and the temporal frequency of 4 Hz. The experimental procedure is shown in **Figure 1**.

On each trial, the grating with different orientations was appeared on the screen in a pseudorandom order. Every stimulus was presented for 2 s and repeated 30 times. A trial started when a lever was pulled by the animal. Then, a fixation point (FP) of 0.1° was displayed in the CRT center. Within 600 ms after FP presentation, the animal was required to fixate within an invisible circular window of 0.6° in radius around the FP. Before the stimulus was displayed, the animal maintained its fixation on the screen for 200 ms. And after the stimulus, there was a blank interval of 200 ms. The FP was then slightly dimmed, and the animal had to release the lever within 600 ms for a drop of juice as reward.

Signal Preprocessing

With increasing popularity of LFP analysis, oscillations in LFP gamma band have been used to study orientation selectivity (Berens et al., 2008; Xing et al., 2012). In order to obtain LFP gamma band signal and preserve the phase relationship between LFP and spikes, we used a two-way least-squares FIR filter in the EEGLAB toolbox (Delorme and Makeig, 2004) to perform zero-phase shift band-pass filtering of 30–80 Hz on the original signals recorded in the experiment.

To identify spikes fired by neurons, the recorded signals were first filtered with a band-pass filter of 300–3,000 Hz. Then, the threshold detection method was used to determine the

spiking time and extract the spike waveform. Finally, spikes were classified by utilizing an unsupervised detection and sorting method based on wavelets and superparamagnetic clustering (Quiroga et al., 2004). An example of the raw data and the corresponding procedure of signal preprocessing is illustrated in **Figure 2**. During the implementation of spike sorting, the cluster with most spikes was taken as the firing activity of a neuron recorded by one electrode and the other spikes were discarded. Thus, there were 48 neurons in V1 and V4, respectively. The inter-spike interval (ISI) histogram metric was used to evaluate the spike sorting accuracy. An example of the ISI is shown in bottom of **Figure 2**. Considering the effect of refractory period, it is an acceptable result of spike sorting for this electrode. The ISI also exhibits similar distribution for other electrodes. The single-unit activity and LFP recorded from the same electrodes were used to calculate the spike-LFP synchronization in this paper.

Synchronization Analysis

The SCMS method was used to analyze the data recorded in the macaques' visual cortex V1 and V4. The main idea of this method is to take the LFP segments centered on each spike as multi-channel signals and measure the synchronization between these LFP signal segments using the phase locking value. The global synchronization is calculated by constructing a correlation matrix to quantify the coupling strength of the spike and LFP. In the data analysis, the influence of window length on the algorithm is very small. However, it is possible that there are other spikes immediately before or after a specific spike which may alter the frequency and phase of the LFP (Zanos et al., 2011). On the other hand, the algorithm uses the similarity of variation in LFP phase as the mechanism for the calculation of spike-LFP synchronization. Considering the impact of these two aspects, we used a window of 20 ms in this study. More details are as follows and the calculation procedure is shown in **Figure 3**.

First, the instantaneous phase of the whole filtered LFP signal is calculated by Hilbert transform. Then, construct the correlation matrix **C** by calculating the phase locking value between pairs of LFP segments, i.e.,

$$c_{mn} = \left| \frac{1}{M} \sum_{k=1}^{M} e^{i(\phi_m(t_k) - \phi_n(t_k))} \right| \tag{1}$$

where $\phi_m(t_k)$ and $\phi_n(t_k)$ denotes the phase of mth and nth LFP segments, respectively, t_k is the sampling time and M denotes the number of samples in the time window. All elements of matrix **C** range from 0 to 1: when c_{mn}= 1, there is a perfect phase synchronization between the mth and nth LFP segments; and when c_{mn}= 0, there is no synchronization. Thus, **C** is a real symmetric matrix of order N and all diagonal elements are equal to 1, where N denotes the number of LFP segments. Moreover, the eigenvalues of matrix **C** ($\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$) are real numbers and the sum of them is N. If all of the LFP segments are totally non-synchronized with each other, **C** will become an identity matrix and all of the eigenvalues will be equal to 1. Once all of the LFP segments are perfectly synchronized, the maximum eigenvalue of **C** will be equal to N and other eigenvalues zero. Above all, eigenvalues can provide information about the synchronization between LFP segments.

Surrogate Data

Finally, in order to obtain a normalized value of spike-LFP synchronization which is independent of the number of spikes, this paper used the Rank-Shuffled Surrogate (RSS) method to generate surrogate data (Junfeng et al., 2012). Assume that {g(n)} denotes a Gaussian random sequence, and R[g(k)] denotes the order in which g(k) is ranked in the time series {g(n)}. For example, if g(k) is the 5th smallest sample point in {g(n)}, then R[g(k)] = 5. Then, use $\tilde{s}$(n) to represent the rank-shuffled surrogate data of the original signals {s(n)}, where $\tilde{s}$(n) = s [k(n)], and k(n) = R[g(n)]. That is to say, the surrogate data is generated by randomizing the order of the original signals, destroying the time structure, but retaining the amplitude distribution, mean and variance.

By using such a method, all spike-triggered LFP segments are randomized to calculate a surrogate correlation matrix **R**. That is, the surrogate data is generated by randomizing the order of the original signals. Similarly, the ordered eigenvalues of surrogate correlation matrix **R** can be obtained, which are denoted as $\lambda_1{}^s \geq \lambda_2{}^s \geq \cdots \geq \lambda_N{}^s$. This randomization process is repeated and calculated 100 times, the mean and standard deviation of the maximum eigenvalues are denoted as $\overline{\lambda_1{}^s}$and σ_1, respectively. Then, the normalized spike-LFP synchronization value is calculated by the following equation:

$$\eta = \begin{cases} \left(\frac{\lambda_1 - \overline{\lambda_1{}^s}}{N - \overline{\lambda_1{}^s}} \right) & if \ \lambda_1 > \left(\overline{\lambda_1{}^s} + K \times \sigma_1 \right) \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where K is a constant that determines the threshold, and K=3 is selected for 99% confidence intervals (Li et al., 2007).

Circle Variance

The circle variance (CV) (Ringach et al., 2002) is an orientation selectivity index obtained by the vector sum of neuron's responses to each orientation of the stimulus divided by the scalar sum of the responses, which can effectively describe the degree of orientation selectivity. Its definition is:

$$\mathrm{CV} = 1 - \frac{\sum_k r_k \exp(\mathrm{i}2\theta_k)}{\sum_k r_k} \tag{3}$$

RESULTS

In order to acquire more accurate and significant results, the trials with very few spikes (<10) and distorted recordings with very small amplitude are rejected, and then the mean spike-LFP synchronization of the remaining trials is calculated. The experimental data was analyzed using Matlab.

Orientation Selectivity of Neurons in V1 and V4

First, to examine the neuronal response in the two brain regions V1 and V4 under the stimuli of sinusoidal grating with different

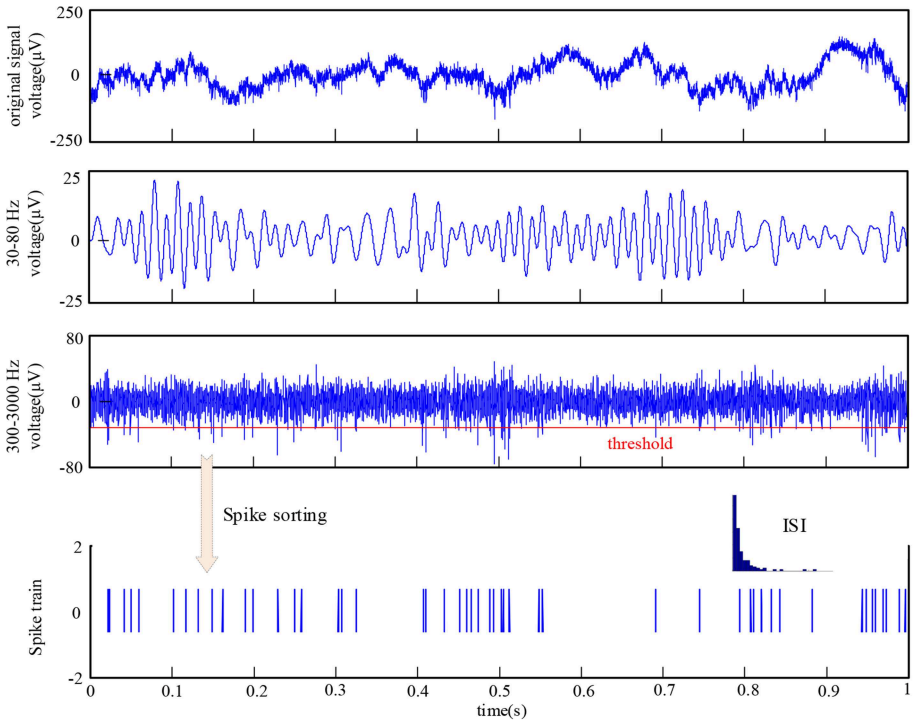

FIGURE 2 | The raw data recorded by one electrode and the procedure of signal preprocessing.

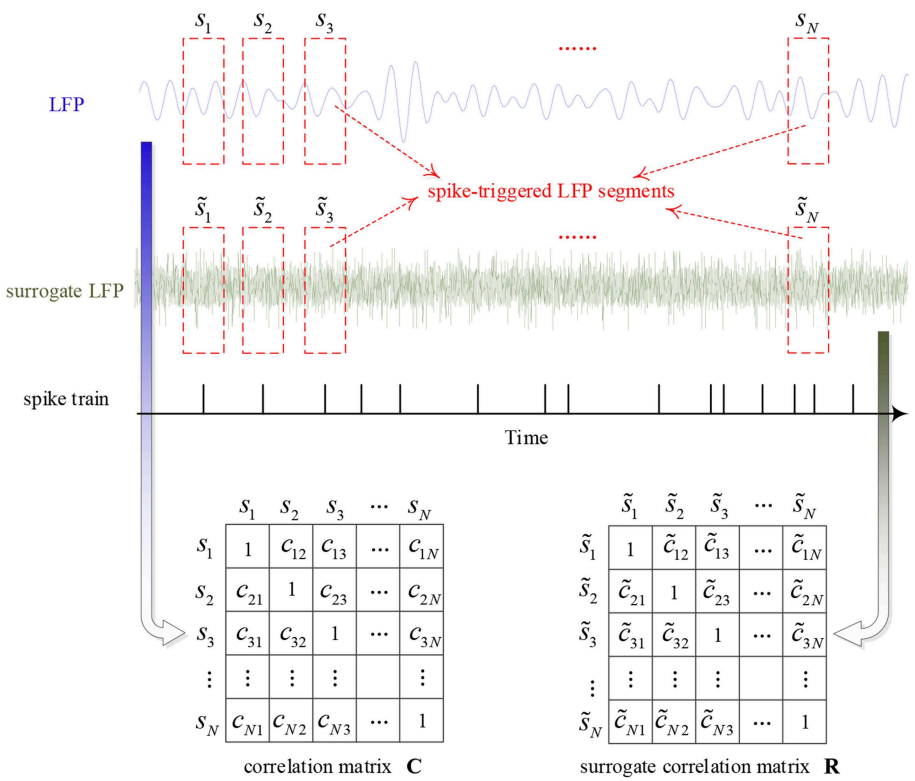

FIGURE 3 | The calculation procedure of the algorithm for characterizing the strength of spike-LFP synchronization.

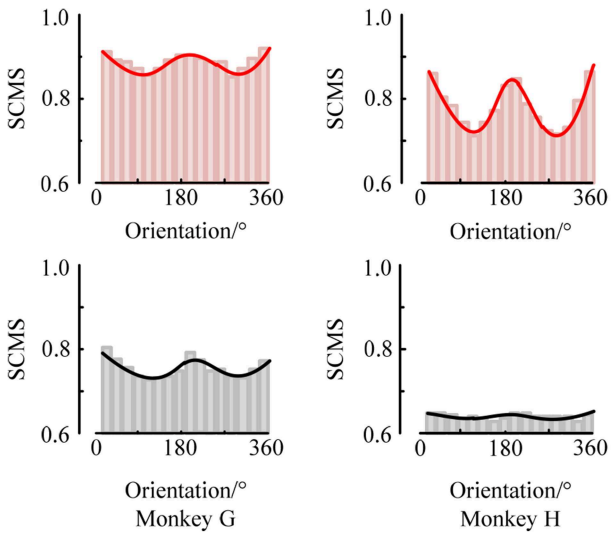

FIGURE 4 | The mean spike-LFP synchronization values in V1 and V4 of the two monkeys. Response of the V1 neurons is indicated by a red line, and response of the V4 neurons is indicated by a black line.

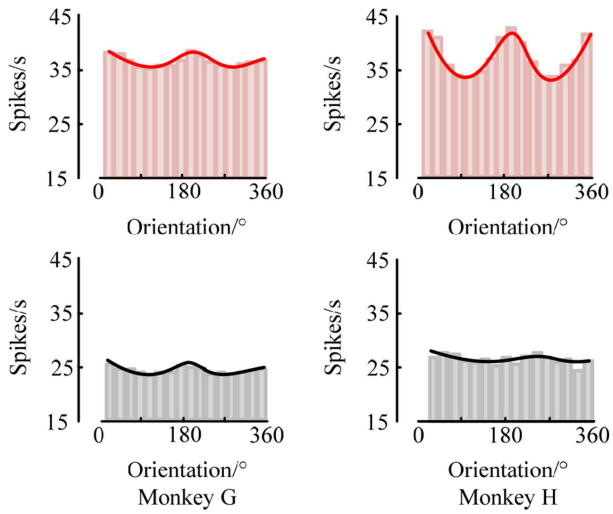

FIGURE 5 | The orientation tuning of the spiking response. The values represent the mean firing rates of neurons to sinusoidal grating with different orientations.

orientations, we used the SCMS method to estimate the spike-LFP synchronization of the experimental data recorded by each electrode respectively. The mean spike-LFP synchronization values of the 48 electrodes in V1 and V4 are plotted in the curve of **Figure 4**.

The red curves show the response of neurons in V1 for the two monkeys. It can be seen that the mean spike-LFP synchronization values exhibit obvious orientation selectivity. More concretely, the neurons respond more strongly to the gratings around 22.5° or 202.5° than the other orientations, which means that the synchronization values to the preference-oriented stimulus and the non-preferential orientation stimulus are markedly different. And the values distribute symmetrically. On the other hand, the black curves show the response of neurons in V4. Clearly, although the distribution of the mean spike-LFP synchronization values is almost symmetric, the difference between them is not as obvious as V1. Especially for monkey H, the neurons did not respond well to the stimulus. There are two reasons may lead to this phenomenon. One is that there are individual differences in the two monkeys and they responded not consistently to the drift gratings. Another, and more important, is that the V4 cortex itself has small patches that encode shape and orientation (Roe et al., 2012). In the experiment, the electrodes in V4 of Monkey H were more likely located close to the color-coded region.

Moreover, the synchronization value of the neurons in V1 is higher than the synchronization value of neurons in V4, indicating that the neurons in V1 are more active and more sensitive to the orientation of the drift grating. Considering that the local field potential is the sum of the excitatory and inhibitory postsynaptic potentials in the vicinity of the recording electrode, it is the superposition of the neuron cluster firing activity in a local area. Then, the synchronization relationship between the spike train fired by a neuron and the local field potential can be understood as the connection between a single neuron and multiple neurons around it. Therefore, it can be considered that the activity of a single neuron in V1 is more affected by the network composed of peripheral neurons, while the neurons in V4 are relatively less affected by peripheral neurons.

At the same time, we also analyzed the orientation tuning of the spiking response under different orientations toward grating stimulation. It can be easily seen from **Figures 4**, **5** that in the two brain regions V1 and V4 under the grating stimulation, peaks appear in the orientation tuning curves based on both the firing rate and the spike-LFP synchronization. This indicates that both the spiking response and the spike-LFP synchronization of neurons have a clear orientation selectivity. Meanwhile, it can be found that the firing rate of neurons in the two brain regions is different. The mean firing rate of neurons in V1 is higher than that in V4, indicating that the neurons in V1 are active and have more firing activity under grating stimulation. Similar results were observed for the mean intensity of the spike-LFP synchronization.

Second, we used the CV to measure the orientation selectivity of neurons and compared the differences between different brain regions. The CV values were calculated separately for the orientation tuning curves obtained by the two methods of the firing rate and the spike-LFP synchronization, and then the neurons in the two brain regions V1 and V4 were statistically analyzed. The result is shown in **Figure 6**.

It can be seen that the CV values obtained by the spike-LFP synchronization are smaller than the CV values obtained by the firing rate. All neurons have significant orientation selectivity and most of the CV values range from 0.6 to 0.8. In addition, it can also be found that the CV values

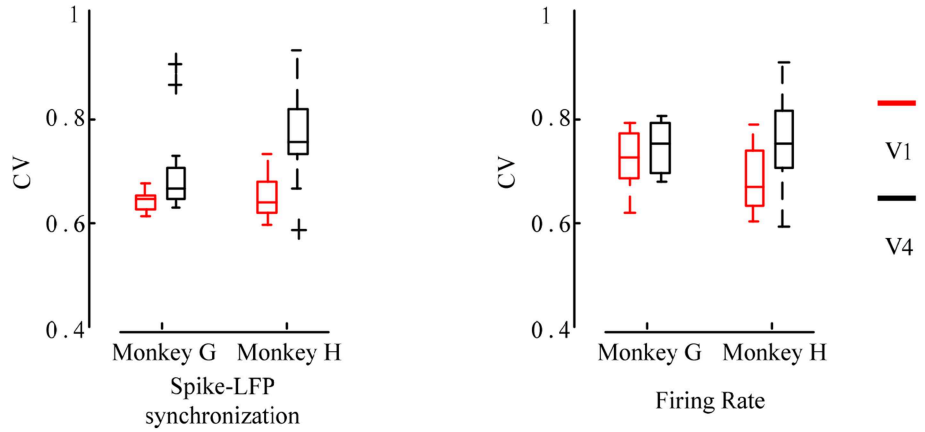

FIGURE 6 | Box plot of the CV values for the firing rate and the spike-LFP synchronization of neurons in different regions. Red indicates the V1 and black indicates the V4. The smaller the CV values, the more significant the orientation selectivity of the neurons.

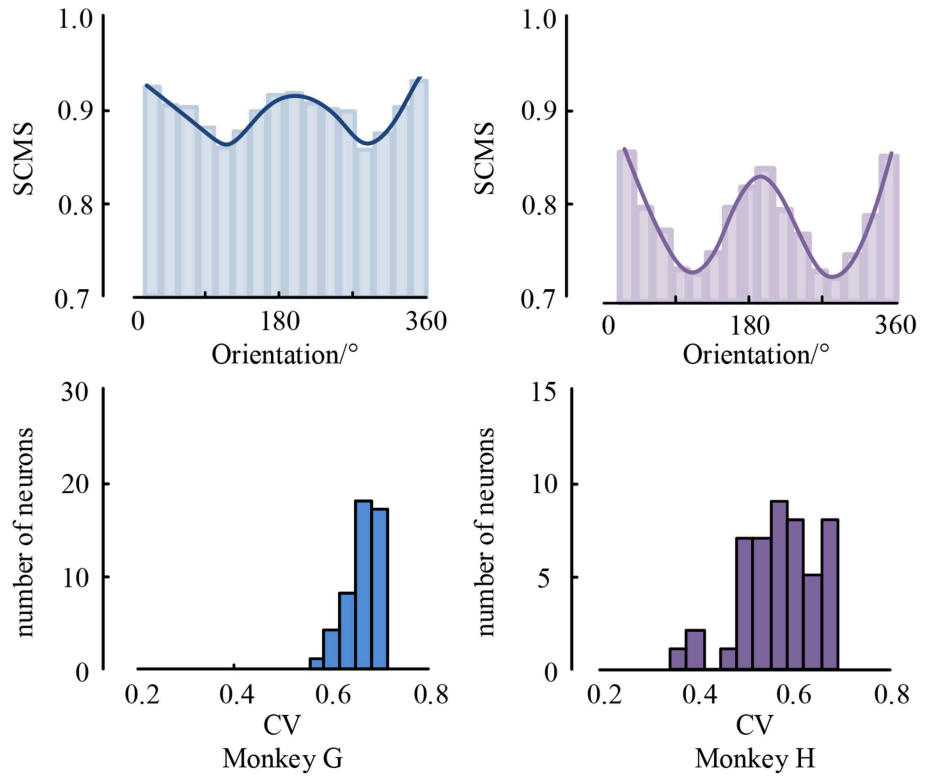

FIGURE 7 | The mean spike-LFP synchronization values between V1 and V4 and the histogram of CV values.

of the neurons in V1 is lower than that in V4. The F test revealed a significant difference between these two regions ($p < 0.05$). It further indicates that the neurons in V1 exhibit stronger orientation selectivity than that in V4. This shows that for sinusoidal grating with different orientations, the firing activity of neurons will show a certain orientation selectivity. However, this orientation selectivity is stronger when considering the spike-LFP synchronization, that is, the synchronization relationship is more sensitive to grating stimulation with different orientations. Therefore, it can be considered that studying the spike-LFP synchronization relationship provides a more effective method for exploring the formation mechanism of visual neurons toward orientation selectivity. It is able to effectively describe the orientation tuning characteristics of neurons and the difference of orientation selectivity in different regions.

Co-modulation Effect on Orientation in V1 and V4

As is known, the processing of visual information requires mutual communication and cooperation among multiple brain areas (Jensen and Mazaheri, 2010; Akam and Kullmann, 2014; Fries, 2015). However, it is still elusive whether and how the distant cortical areas cooperate in visual tasks (Tiesinga and Buia, 2009; Ter Wal and Tiesinga, 2017). For example, whether LFPs coordinate spiking output timely between distant cortical areas that have been traditionally associated with the sensory encoding of visual information, and is the precision of coordination between these areas related to changes in visual information? To understand these questions, we used local field potential and spike recorded in monkeys performing a visual task to study neural interactions between visual area V1 and V4. The result is shown in **Figure 7**.

As can be seen, the spike-LFP synchronization between V1 and V4 is modulated by visual information content, and its intensity also shows orientation selectivity to drifting gratings during the stimulation. Moreover, all neurons have significant orientation selectivity with CV values less than 0.8. In addition, we also found that the spike-LFP synchronization value between regions is higher than in a single region. This suggests that the spike-LFP synchronization coordinates potential communication between V1 and V4. Specifically, the spike-LFP synchronization is enhanced during visual tasks in both V1 and V4, and increased synchronization is accompanied by the phase coding of visual stimulus.

Moreover, spiking activity in V4 was more strongly locked to LFP in V1 and vice versa, i.e., V4 spiking seems to be more sensitive to V1 gamma than V1 spiking to V4 gamma. The asymmetry of spike-LFP synchronization between the regions implies a possible directedness in the interaction and communication pattern between the regions, the details of which remain to be explored.

DISCUSSION

We combined spike and LFP signals to investigate the orientation tuning characteristics of neurons in macaques' V1 and V4 under drifting sinusoidal gratings by calculating the synchronization between spike and LFP gamma band. The results are as follows:

First, we found a strong correlation between the spike-LFP synchronization and the stimulus orientation, which is modulated by the orientation and reflects the basic feature information of the visual stimulation. Second, we also investigated the modulation of orientation selectivity through the spike-LFP synchronization of V1 and V4 neurons. The results show that the spike-LFP synchronization not only can effectively encode the stimulus information for different orientations, but also can distinctly distinguish the orientation tuning characteristics of neurons in different regions. Finally, it was observed that there is a clear mutual modulation of orientation between V1 and V4, suggesting that the neural interaction based on the spike-LFP synchronization between these two long-range cortical regions is related to the coding of visual information.

Our findings are consistent with previous studies. For instance, Frien et al. found that gamma-band LFP displays sharper orientation tuning than slower components of the same recordings in striate cortex of the awake monkey (Frien et al., 2000). Lashgari et al. systematically compared the stimulus selectivity of LFP and neighboring single-unit activity recorded in V1 of awake rhesus monkeys. They demonstrated that LFP and single-unit activity have similar stimulus preferences for orientation, direction of motion, contrast and other features (Lashgari et al., 2012). Womelsdorf et al. determined for each spike its phase relative to the gamma cycle and used the pairwise phase consistency to quantify the concentration of phases around the mean gamma phase. They observed that orientation selectivity is modulated by gamma phase and the spike firing rate that occurred close to a neuron's mean gamma phase is most orientation selective (Womelsdorf et al., 2012). Although these results are closely related to ours in this paper, there are two clear distinctions between them. One is that we used an unbiased and robust measure to quantify the spike-LFP synchronization, which provided a reliable comparison between trials with different spike numbers. Then, it is feasible to investigate the tuning characteristics of spike-LFP synchronization under stimulus with different orientations. Another is that we analyzed the spike-LFP synchronization not limited to V1, but expanded it to V4. And we also studied the mutual modulation between V1 and V4.

Taken together, our results illustrate that the connection between spike-LFP synchronization and orientation not only exists in an individual region (V1 or V4), but also between distant cortical regions (V1 and V4). That is, the neural interaction based on spike-LFP synchronization may be related to the maintenance and communication of information during visual information processing. We suggest that this method provides a new direction to study the formation mechanism of orientation selectivity.

ETHICS STATEMENT

All procedures were conducted in compliance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals, and were approved by the Institutional Animal Care and Use Committee of Beijing Normal University.

AUTHOR CONTRIBUTIONS

ZL contributed to simulation design and manuscript preparation. MG contributed to performing simulations and analysis. YW contributed to manuscript preparation.

ACKNOWLEDGMENTS

This research was supported in part by the National Natural Science Foundation of China (61603327, 61673333) and the Natural Science Foundation of Hebei China (F2017203010). We thank Wu Li of the State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University for providing the Rhesus monkey data and for useful discussions. We also thank Ying Yan for his assistance in preparing the manuscript.

REFERENCES

Akam, T., and Kullmann, D. M. (2014). Oscillatory multiplexing of population codes for selective communication in the mammalian brain. *Nat. Rev. Neurosci.* 15, 111–122. doi: 10.1038/nrn3668

Anderson, J. S., Lampl, I., Gillespie, D. C., and Ferster, D. (2000). The contribution of noise to contrast invariance of orientation tuning in cat visual cortex. *Science* 290, 1968–1972. doi: 10.1126/science.290.5498.1968

Berens, P., Keliris, G. A., Ecker, A. S., Logothetis, N. K., and Tolias, A. S. (2008). Comparing the feature selectivity of the gamma-band of the local field potential and the underlying spiking activity in primate visual cortex. *Front. Syst. Neurosci.* 2:2. doi: 10.3389/neuro.06.002.2008

Bharmauria, V., Bachatene, L., Ouelhazi, A., Cattan, S., Chanauria, N., Etindele-Sosso, F. A., et al. (2016). Interplay of orientation selectivity and the power of low- and high-gamma bands in the cat primary visual cortex. *Neurosci. Lett.* 620, 14–19. doi: 10.1016/j.neulet.2016.03.033

Buzsaki, G., Anastassiou, C. A., and Koch, C. (2012). The origin of extracellular fields and currents–EEG, ECoG, LFP and spikes. *Nat. Rev. Neurosci.* 13, 407–420. doi: 10.1038/nrn3241

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Durant, S., Sulykos, I., and Czigler, I. (2017). Automatic detection of orientation variance. *Neurosci. Lett.* 658, 43–47. doi: 10.1016/j.neulet.2017.08.027

Frien, A., Eckhorn, R., Bauer, R., Woelbern, T., and Gabriel, A. (2000). Fast oscillations display sharper orientation tuning than slower components of the same recordings in striate cortex of the awake monkey. *Eur. J. Neurosci.* 12, 1453–1465. doi: 10.1046/j.1460-9568.2000.00025.x

Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235. doi: 10.1016/j.neuron.2015.09.034

Gaucher, Q., Edeline, J. M., and Gourevitch, B. (2012). How different are the local field potentials and spiking activities? Insights from multi-electrodes arrays. *J. Physiol. Paris* 106, 93–103. doi: 10.1016/j.jphysparis.2011.09.006

Girshick, A. R., Landy, M. S., and Simoncelli, E. P. (2011). Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nat. Neurosci.* 14, 926–932. doi: 10.1038/nn.2831

Henrie, J. A., and Shapley, R. (2005). LFP power spectra in V1 cortex: the graded effect of stimulus contrast. *J. Neurophysiol.* 94, 479–490. doi: 10.1152/jn.00919.2004

Hubel, D. H., and Wiesel, T. N. (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

Ince, R. A., Mazzoni, A., Bartels, A., Logothetis, N. K., and Panzeri, S. (2012). A novel test to determine the significance of neural selectivity to single and multiple potentially correlated stimulus features. *J. Neurosci. Methods* 210, 49–65. doi: 10.1016/j.jneumeth.2011.11.013

Jensen, O., and Mazaheri, A. (2010). Shaping functional architecture by oscillatory alpha activity: gating by inhibition. *Front. Hum. Neurosci.* 4:186. doi: 10.3389/fnhum.2010.00186

Junfeng, S., Xiangfei, H., and Shanbao, T. (2012). Phase synchronization analysis of EEG signals: an evaluation based on surrogate tests. *IEEE Trans. Biomed. Eng.* 59, 2254–2263. doi: 10.1109/TBME.2012.2199490

Lashgari, R., Li, X., Chen, Y., Kremkow, J., Bereshpolova, Y., Swadlow, H. A., et al. (2012). Response properties of local field potentials and neighboring single neurons in awake primary visual cortex. *J. Neurosci.* 32, 11396–11413. doi: 10.1523/JNEUROSCI.0429-12.2012

Li, S., Ouyang, M., Liu, T., Bai, W., Yi, H., and Tian, X. (2014). Increase of spike-LFP coordination in rat prefrontal cortex during working memory. *Behav. Brain Res.* 261, 297–304. doi: 10.1016/j.bbr.2013.12.030

Li, X., Cui, D., Jiruska, P., Fox, J. E., Yao, X., and Jefferys, J. G. (2007). Synchronization measurement of multiple neuronal populations. *J. Neurophysiol.* 98, 3341–3348. doi: 10.1152/jn.00977.2007

Li, Z., Cui, D., and Li, X. (2016). Unbiased and robust quantification of synchronization between spikes and local field potential. *J. Neurosci. Methods* 269, 33–38. doi: 10.1016/j.jneumeth.2016.05.004

Mansfield, R. J. (1974). Neural basis of orientation perception in primate vision. *Science* 186, 1133–1135. doi: 10.1126/science.186.4169.1133

Manyakov, N. V., and Van Hulle, M. M. (2010). Decoding grating orientation from microelectrode array recordings in monkey cortical area V4. *Int. J. Neural Syst.* 20, 95–108. doi: 10.1142/S0129065710002280

McLaughlin, D., Shapley, R., Shelley, M., and Wielaard, D. J. (2000). A neuronal network model of macaque primary visual cortex (V1): orientation selectivity and dynamics in the input layer 4Calpha. *Proc. Natl. Acad. Sci. U.S.A.* 97, 8087–8092. doi: 10.1073/pnas.110135097

Mehring, C., Rickert, J., Vaadia, E., Cardosa de Oliveira, S., Aertsen, A., and Rotter, S. (2003). Inference of hand movements from local field potentials in monkey motor cortex. *Nat. Neurosci.* 6, 1253–1254. doi: 10.1038/nn1158

Mollazadeh, M., Aggarwal, V., Thakor, N. V., Law, A. J., Davidson, A., and Schieber, M. H. (2009). "Coherency between spike and LFP activity in M1 during hand movements," in *4th International IEEE/EMBS Conference on Neural Engineering* (Antalya), 506–509. doi: 10.1109/NER.2009.5109344

Perge, J. A., Zhang, S., Malik, W. Q., Homer, M. L., Cash, S., Friehs, G., et al. (2014). Reliability of directional information in unsorted spikes and local field potentials recorded in human motor cortex. *J. Neural Eng.* 11:046007. doi: 10.1088/1741-2560/11/4/046007

Quian Quiroga, R., and Panzeri, S. (2009). Extracting information from neuronal populations: information theory and decoding approaches. *Nat. Rev. Neurosci.* 10, 173–185. doi: 10.1038/nrn2578

Quiroga, R. Q., Nadasdy, Z., and Ben-Shaul, Y. (2004). Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering. *Neural Comput.* 16, 1661–1687. doi: 10.1162/089976604774201631

Ray, S., and Maunsell, J. H. (2011a). Different origins of gamma rhythm and high-gamma activity in macaque visual cortex. *PLoS Biol.* 9:e1000610. doi: 10.1371/journal.pbio.1000610

Ray, S., and Maunsell, J. H. (2011b). Network rhythms influence the relationship between spike-triggered local field potential and functional connectivity. *J. Neurosci.* 31, 12674–12682. doi: 10.1523/JNEUROSCI.1856-11.2011

Ringach, D. L., Shapley, R. M., and Hawken, M. J. (2002). Orientation selectivity in macaque V1: diversity and laminar dependence. *J. Neurosci.* 22, 5639–5651. doi: 10.1523/JNEUROSCI.22-13-05639.2002

Roe, A. W., Chelazzi, L., Connor, C. E., Conway, B. R., Fujita, I., Gallant, J. L., et al. (2012). Toward a unified theory of visual area V4. *Neuron* 74, 12–29. doi: 10.1016/j.neuron.2012.03.011

Siegel, M., and Konig, P. (2003). A functional gamma-band defined by stimulus-dependent synchronization in area 18 of awake behaving cats. *J. Neurosci.* 23, 4251–4260. doi: 10.1523/JNEUROSCI.23-10-04251.2003

Ter Wal, M., and Tiesinga, P. H. (2017). Phase difference between model cortical areas determines level of information transfer. *Front. Comput. Neurosci.* 11:6. doi: 10.3389/fncom.2017.00006

Tiesinga, P. H., and Buia, C. I. (2009). Spatial attention in area V4 is mediated by circuits in primary visual cortex. *Neural Netw.* 22, 1039–1054. doi: 10.1016/j.neunet.2009.07.010

Womelsdorf, T., Lima, B., Vinck, M., Oostenveld, R., Singer, W., Neuenschwander, S., et al. (2012). Orientation selectivity and noise correlation in awake monkey area V1 are modulated by the gamma cycle. *Proc. Natl. Acad. Sci. U.S.A.* 109, 4302–4307. doi: 10.1073/pnas.1114223109

Xing, D., Shen, Y., Burns, S., Yeh, C. I., Shapley, R., and Li, W. (2012). Stochastic generation of gamma-band activity in primary visual cortex of awake and anesthetized monkeys. *J. Neurosci.* 32, 13873–13880a. doi: 10.1523/JNEUROSCI.5644-11.2012

Zanos, T. P., Mineault, P. J., and Pack, C. C. (2011). Removal of spurious correlations between spikes and local field potentials. *J. Neurophysiol.* 105, 474–486. doi: 10.1152/jn.00642.2010

Zhang, L., and Li, B. (2013). Surround modulation characteristics of local field potential and spiking activity in primary visual cortex of cat. *PLoS ONE* 8:e64492. doi: 10.1371/journal.pone.0064492

Conflict of Interest Statement: 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.

Copyright © 2019 Li, Gao and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Integrative Neuroscience

ORIGINAL RESEARCH
published: 17 July 2019
doi: 10.3389/fnint.2019.00028

The Lateralization of Speech-Brain Coupling Is Differentially Modulated by Intrinsic Auditory and Top-Down Mechanisms

M. F. Assaneo[1], J. M. Rimmele[2], J. Orpella[1], P. Ripollés[1], R. de Diego-Balaguer[3,4,5,6] and D. Poeppel[1,2]*

[1] Department of Psychology, New York University, New York, NY, United States, [2] Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany, [3] Departament de Cognició, Desenvolupament i Psicologia de l'Educació, University of Barcelona, Barcelona, Spain, [4] Catalan Institute for Research and Advance Studies, Barcelona, Spain, [5] Cognition and Brain Plasticity Unit, IDIBELL, L'Hospitalet de Llobregat, Spain, [6] Institute of Neuroscience, University of Barcelona, Barcelona, Spain

The lateralization of neuronal processing underpinning hearing, speech, language, and music is widely studied, vigorously debated, and still not understood in a satisfactory manner. One set of hypotheses focuses on the temporal structure of perceptual experience and links auditory cortex asymmetries to underlying differences in neural populations with differential temporal sensitivity (e.g., ideas advanced by Zatorre et al. (2002) and Poeppel (2003). The Asymmetric Sampling in Time theory (AST) (Poeppel, 2003), builds on cytoarchitectonic differences between auditory cortices and predicts that modulation frequencies within the range of, roughly, the syllable rate, are more accurately tracked by the right hemisphere. To date, this conjecture is reasonably well supported, since – while there is some heterogeneity in the reported findings – the predicted asymmetrical entrainment has been observed in various experimental protocols. Here, we show that under specific processing demands, the rightward dominance disappears. We propose an enriched and modified version of the asymmetric sampling hypothesis in the context of speech. Recent work (Rimmele et al., 2018b) proposes two different mechanisms to underlie the auditory tracking of the speech envelope: one derived from the intrinsic oscillatory properties of auditory regions; the other induced by top-down signals coming from other non-auditory regions of the brain. We propose that under non-speech listening conditions, the intrinsic auditory mechanism dominates and thus, in line with AST, entrainment is rightward lateralized, as is widely observed. However, (i) depending on individual brain structural/functional differences, and/or (ii) in the context of specific *speech listening* conditions, the relative weight of the top-down mechanism can increase. In this scenario, the typically observed auditory sampling asymmetry (and its rightward dominance) diminishes or vanishes.

Keywords: asymmetrical sampling, brain to stimulus synchronization, MEG (magnetoencephalography), speech perception, speech envelope tracking

OPEN ACCESS

Edited by:
Andreas Bahmer, Universitätsklinikum Würzburg, Germany

Reviewed by:
Kenneth Hugdahl, University of Bergen, Norway
Keiichi Kitajo, National Institute for Physiological Sciences (NIPS), Japan

****Correspondence:***
M. F. Assaneo
fassaneo@gmail.com

Received: *30 March 2019*
Accepted: *02 July 2019*
Published: *17 July 2019*

Citation:
Assaneo MF, Rimmele JM, Orpella J, Ripollés P, de Diego-Balaguer R and Poeppel D (2019) The Lateralization of Speech-Brain Coupling Is Differentially Modulated by Intrinsic Auditory and Top-Down Mechanisms. Front. Integr. Neurosci. 13:28. doi: 10.3389/fnint.2019.00028

INTRODUCTION

Considerable advances in our understanding of the neural basis of speech processing have been made in the last decade. There is emerging consensus about a functional dissociation of the neuronal substrate underlying speech processing into ventral and dorsal pathways (Hickok and Poeppel, 2007; Saur et al., 2008; Rauschecker and Scott, 2009; Friederici, 2012) increasing evidence suggests an important role of both hemispheres (albeit contributions differ depending on the processing levels) (Binder et al., 2000; Cogan et al., 2014; Sammler et al., 2015) and the crucial role played by the sensorimotor circuitry during verbal learning and speech comprehension processes (Lopez-Barroso et al., 2011, 2013) is well-established. There are, to be sure, open questions and unsolved puzzles. Here we address controversial findings regarding hemispheric lateralization in the auditory cortex during the processing of speech. We propose that the differential contribution of both hemispheres to the processing of the speech acoustics reflects intrinsic attributes of the neural populations in the auditory cortex as well as modulation by top-down influence from non-auditory cortical areas. We provide new neurophysiological data supporting these claims.

Based on early foundational work (Giraud et al., 2000; Ahissar et al., 2001; Rimmele et al., 2018a) followed by a number of recent experiments (Luo and Poeppel, 2007; Kubanek et al., 2013; Ding et al., 2014; Crosse et al., 2015; Rimmele et al., 2015), it is now established that during speech comprehension low-frequency neural activity is *entrained* by connected speech, and in particular by attributes of the speech envelope. Neuronal entrainment (or speech tracking) denotes the alignment of the neuronal excitability phase of slow oscillations in auditory cortex with slow energy fluctuations in the speech acoustics. Crucially, entrainment to speech has been argued to have causal force (Doelling et al., 2014; Ghitza, 2014) (rather than being epiphenomenal), and, accordingly, the associated neurophysiological mechanisms have received much attention. However, there are controversial findings in this growing literature that challenge existing explanations.

One hypothesized mechanism to account for the neuronal entrainment to speech and its hemispheric lateralization is the Asymmetric Sampling in Time model (AST) (Poeppel, 2003). AST postulates that there are two different temporal integration constants in non-primary auditory cortex that result from the *intrinsic* properties of local neuronal ensembles. An *asymmetric sampling in time* results from the right hemispheric auditory cortical structures having a larger population of neural ensembles with longer temporal integration windows [∼100 to 300 ms, i.e., roughly corresponding to the syllabic rate (Ding et al., 2017)] compared to the left. These temporal windows, or specifically their neural instantiation, is reflected in neuronal oscillatory activity (longer window: theta; shorter window: low gamma) that aligns with basic units of speech, viz. syllables (theta) and phonetic or segmental information (gamma). In accordance with this hypothesis, there exists a growing body of evidence supporting a rightward preference for the processing of more slowly modulated acoustic information (Boemio et al., 2005; Giraud et al., 2007; Abrams et al., 2008; Telkemeyer et al., 2009; Morillon et al., 2012) in this regard, the AST conjecture accords well with related hypotheses about hemispheric asymmetries in processing spectral versus temporal sound characteristics (Zatorre and Belin, 2001). As the proposed temporal integration constants relate closely to the intrinsic properties of the auditory cortex in each hemisphere (Zatorre et al., 2002; Poeppel, 2003), we refer to the neuronal oscillatory activity in auditory cortex as an *intrinsic mechanism*.

In spite of this evidence, a closer inspection of previous findings reveals that most of the studies that report a rightward lateralization of the processing of slow acoustic modulations rely on tasks with low language processing demands, typically using auditory stimuli such as non-speech signals (Zatorre and Belin, 2001; Boemio et al., 2005; Telkemeyer et al., 2009; Vanvooren et al., 2014) (e.g., modulated noise or pure-tone patterns), unattended speech (Abrams et al., 2008, 2009), streams of monosyllables (Doelling et al., 2014), or a small number of sentences repeated many times (Luo and Poeppel, 2007).

Furthermore, it has been shown that the strength of speech entrainment in the left (Ahissar et al., 2001; Zoefel et al., 2018), but not in the right (Peelle et al., 2013), auditory cortex covaries with speech intelligibility. Accordingly, another influence on auditory cortex entrainment to speech has been recently described. A set of experiments showed that top-down signals, coming from frontal areas, increase the synchronization between the auditory cortex and the stimulus envelope, particularly in the left hemisphere (Park et al., 2015, 2018; Morillon and Baillet, 2017). In light of these findings, Rimmele et al. (2018b) postulated that frontal areas modulate the intrinsic oscillatory activity of the auditory cortex on the basis of predictive cues in the speech signal, such as rate fluctuations, syntactic or semantic information, or motor production-related predictions, permitting a more flexible tracking of speech than that attained with oscillatory entrainment alone.

A natural question derives from this elegant proposal: how does the integration of these two mechanisms – i.e., *intrinsic auditory* and *externally driven* – modulate the canonical *rightward lateralization* of the slow frequency neuronal entrainment in auditory cortex? To answer this question, we built on the following observations: (i) most of the research that shows rightward lateralization relies on tasks with low language processing demands (Zatorre and Belin, 2001; Boemio et al., 2005; Abrams et al., 2008, 2009; Telkemeyer et al., 2009; Doelling et al., 2014; Vanvooren et al., 2014) (ii) the strength of speech entrainment in the left (Ahissar et al., 2001; Zoefel et al., 2018), but not in the right (Peelle et al., 2013) auditory cortex covaries *particularly* with speech intelligibility; (iii) frontal top-down signals can enhance the entrainment of the left auditory cortex to the speech envelope (Park et al., 2015, 2018; Federmeier, 2007); and (iv) a recent study demonstrates that speech tracking is affected by neurophysiological and neuroanatomical individual differences and that for a subset of the population – characterized by strong audio-motor interaction – the auditory tracking is balanced between hemispheres (i.e., the expected rightward asymmetry disappears) (Assaneo et al., 2019). Connecting these empirical observations, we propose that, while listening to non-speech stimuli, auditory entrainment principally reflects

the *intrinsic auditory* mechanism, thus exposing the rightward hemispheric asymmetry. However, under specific speech-listening conditions, or due to neuronal functional and structural individual differences, the *externally driven* mechanism can affect the neuronal activity, mostly in the left hemisphere, equalizing the strength of entrainment across hemispheres. We apply new analyses to three published magnetoencephalography (MEG) datasets (Assaneo and Poeppel, 2018; Assaneo et al., 2019; Rimmele et al., 2019) to present new evidence to support these claims.

MATERIALS AND METHODS

The datasets used in this manuscript belong to three previously published experiments. Materials and methods of each experiment are briefly described below. For more detail see Assaneo and Poeppel (2018) for Experiment A, Assaneo et al. (2019) for Experiment B, and Rimmele et al. (2019) for Experiment C.

Participants

All participants self-reported normal hearing and no neurological deficits, and all had normal structural MRI scans. Participants were paid for taking part in the different studies and provided written informed consent.

Experiments A and B

The protocol was approved by the local Institutional Review Board (New York University's Committee on Activities Involving Human Subjects).

Experiment C

The protocol was approved by the local ethics committee of the University Hospital Frankfurt (Germany).

Experiment A

A cohort of 19 individuals participated in the study and two were removed – one was not able to perform the task, for the other one the MEG signal was too noisy. The analyzed sample consisted of 17 participants (9 males; mean age 28, range 20–40; 15 native speakers of American English and 2 native speakers of Spanish).

Experiment B

A group of 40 participants completed the experiment, the data from three was not analyze, since the acquired MEG signal was too noisy. The final database included 37 right handed participants (18 males; mean age, 30; age range, 21 to 55).

Experiment C

Twenty-one individuals participated in this study. Two participants were removed, because of outlier behavioral performance (accuracy $<$ mean $-$ 2 $\times$ SD) and because of technical issues (i.e., audio problems). The final sample comprised 19 right-handed German native speakers (n = 19) with no previous knowledge of Turkish (male: 9; mean age: 24.46 years; SD: 3.73 years).

Task

Experiments A and B

In both experiments, participants passively listened to a set of syllable streams while their neural activity was recorded. At the end of each trial, participants indicated, by pressing a button, whether a given syllable had been presented. In Experiment A participants also completed a motor and an auditory localizer task.

Experiment C

During the MEG recording, participants were asked to listen attentively to sequences of di-syllabic German words (*Semantic Condition*) or Turkish pseudo-words (*Non-Semantic Condition*). Overall, 15 blocks were presented, each consisting of 210 trials (105 per condition). In total, each German and Turkish word (note that the syllables of the Turkish words were randomized) was repeated 15 times. Each block contained 29% trials with a target stimulus (i.e., a syllable repetition) equally distributed across conditions. After each trial, participants indicated the presence of a target stimulus with a button press.

Stimuli

Experiment A

English syllables */ba/*, */wa/*, */ma/*, and */va/* were synthesized using an online text-to-speech software www.fromtexttospeech.com/. The stimulus intensity was normalized based on the amplitude root mean square and the signal was compressed to 120-ms duration using Praat software (Boersma, 2001). Trials contained 3 s of silence (baseline) followed by 6 s of syllables. Two syllables were randomly selected from out of the four syllables for each trial. The syllables were sequentially presented with an occurrence frequency of 0.7 for one and 0.3 for the other. Varying the syllable rate generated six different conditions of trials: 2.5, 3.5, 4.5, 5.5, and 6.5 syllables per second.

Experiment B

Five sets of syllable streams were generated using the MBROLA text-to-speech synthesizer (Bozkurt et al., 1996). All phonemes were equal in pitch (200 Hz), pitch rise and fall (with the maximum amplitude at 50% of the phoneme) and duration, which was set to 111 ms to get a presentation rate of 4.5 syllables per second. Each set lasted 2 min and consisted of 12 distinct syllables (unique consonant-vowel combinations).

Experiment C

In total, 134 German disyllabic words were selected from the CELEX lexical database (Baayen et al., 1995) and 134 Turkish disyllabic words from the TELL database[1] [for details see Rimmele et al. (2019)]. German and Turkish syllables produced by a female German/Turkish bilingual speaker were recorded. The recordings were high-pass filtered at 60 Hz, compressed in duration (250 ms), and normalized in peak-amplitude and pitch contour (at 250 Hz). The two syllables of each word were concatenated to generate the German word stimuli. Di-syllabic pseudo-words were created by concatenating two syllables that

[1] http://linguistics.berkeley.edu/TELL/

were quasi-randomly selected from all Turkish syllable stimuli with equal probability of first/second syllable position. For each sequence, randomly selected disyllabic stimuli were concatenated (19 disyllabic stimuli per sequence). Overall, three different sets of sequences were created.

Data Acquisition and Processing

Experiments A and B

Neural activity was recorded with a 157-channel whole-head axial gradiometer system (KIT, Kanazawa Institute of Technology, Japan) emplaced in a magnetically shielded room. The recordings were acquired at 1000 Hz. An online bandpass filter between 1 and 200 Hz and a notch filter at 60 Hz were applied.

In order to monitor the subject's head position, five electromagnetic coils were attached and localized to the MEG sensors at the beginning of the experiment. The position of the coils with respect to three anatomical landmarks: nasion, and left and right tragus were determined using a 3D digitizer software (Source Signal Imaging, Inc.) and digitizing hardware (Polhemus, Inc.). This measurement was used to coregister the MEG data with the subjects' anatomical magnetic resonance image (MRI).

Data processing and analyses were conducted using custom MATLAB code and the FieldTrip toolbox (Oostenveld et al., 2011). Noisy channels were visually rejected for each participant's dataset. The continuous MEG recordings were submitted to two different procedures. First, a least squares projection was fitted to the data from the 2 min of *empty room* recorded at the end of each session. The corresponding component was removed from the recordings (Adachi et al., 2001). Second, the environmental magnetic field was measured with three reference sensors located away from the participant's head, and was regressed out from the MEG signals using time-shifted PCA (de Cheveigné and Simon, 2007). The MEG signals were detrended and artifacts related to eyeblinks and heartbeats were removed using independent component analysis.

Experiment C

A 269-channel whole-head system (Omega 2000, CTF Systems Inc.) situated in a magnetically shielded room was used for the MEG recordings. A sampling rate of 1200 Hz, an online low pass filter (cut-off: 300 Hz), and online denoising (higher-order gradiometer balancing) were applied. The head position relative to the MEG sensors was continuously tracked and head displacement was corrected in the breaks using the fieldtrip toolbox (Stolk et al., 2013). The data were band-pass filtered off-line (1–160 Hz, Butterworth filter; filter order 4) and line-noise was removed using bandstop filters (49.5–50.5; filter order 4). Muscle, jump and slow artifacts were removed in a semi-automatic artifact detection procedure. Trials that contained head movements that exceeded a threshold (5 mm) were rejected. Sensors with high variance were rejected. Eye-blink, eye-movement and heartbeat-related artifacts were removed, using independent component analysis [infomax algorithm (Makeig et al., 1996)]. The data was first reduced to 64 components using principal component analysis. Trials with correct responses were selected and the trial number was matched between the conditions by randomly selecting trials of the condition with less trials (overall trial number, mean = 68.68, SD = 10.27).

Structural MRI

Experiments A and B

High-resolution T1-weighted 3D volume MR data were acquired using a Siemens Allegra 3T and a Siemens Prisma 3T scanner for Experiment A and B, respectively. Each participant's MRI data were preprocessed following the FieldTrip pipeline. Cortical reconstruction and volumetric segmentation were performed with the FreeSurfer image analysis suite.

Experiment C

Individual T1-weighted MRI scans were acquired (for all participants except for three). The standard Montreal Neurological Institute (MNI) template brain was used, in case an individual MRI was missing. MRIs were recorded on a 3 Tesla scanner (Siemens Magnetom Trio, Siemens, Erlangen, Germany) and anatomical landmarks (nasion, left and right pre-auricular points) were marked via Vitamin-E capsules. From the individual MRIs of all participants, probabilistic tissue maps (including cerebrospinal fluid white and gray matter) were retrieved using the FieldtTrip toolbox.

Source Reconstruction

Different approaches were used to reconstruct the brain activity across experiments: cortically constrained MNE (Dale et al., 2000) in Experiment A, linearly constrained minimum variance beamforming (Nolte, 2003) in Experiment B, and Dynamic Imaging of Coherent Sources (Gross et al., 2001) (DICS) in Experiment C.

Brain-to-Stimulus Synchronization

Experiments A and B

Synchronization was estimated by computing the phase locking value (PLV) between the brain activity and the cochlear envelope (Ding et al., 2017) of the perceived stream of syllables. Specifically, the PLV was computed using the following formula:$PLV = \frac{1}{T}\left|\sum_{t=1}^{T} e^{i(\theta_1(t)-\theta_2(t))}\right|$, where t is the discretized time, T is the total number of time points, and θ_1 and θ_2 represent the phase of the brain activity and the cochlear envelope, respectively.

For Experiment A the PLV was computed within a frequency band of ± 0.5 Hz around the syllable rate using windows of 2-s length and 1-s overlap. The percentage of change from baseline was estimated as the difference between the PLV computed for the stimulation window and the PLV computed for the baseline window divided by the latter. For Experiment B the PLV was computed within a frequency band from 3.5 to 5.5 Hz using windows of 1-s length and 0.5-s overlap. In both cases the results for all time windows were averaged separately for each condition obtaining one PLV per voxel and per subject.

Auditory entrainment was estimated by averaging the PLVs of all voxels within the auditory cortex. The method used to define this region of interest (ROI) varied along experiments. In Experiment A, it was functionally localized. In Experiment B,

areas were anatomically defined as BA 41/41, TE 1.0 and TE 1.2 using the Brainnetome Atlas (Fan et al., 2016).

Experiment C

The speech envelope was computed for each acoustic trial by using the following procedure (Smith et al., 2002): the waveforms were filtered into 8 frequency bands, the Hilbert transform was applied for each band, and the absolute magnitude of the 8 analytic signals was averaged. The obtained speech envelope was downsampled to 500 Hz. The spectral complex coefficients at 4 Hz were computed trial-wise for the speech envelope and the MEG data with a 0.1111 Hz resolution, and coherence was computed between all sensors and the speech envelope. The data were projected to source space using a common filter (DICS; λ = 100%; 0.8 cm grid), and Fischer *z*-transformation was applied. Voxels of the left and right Heschl's Gyrus were selected based on the *automated anatomical labeling* atlas (Tzourio-Mazoyer et al., 2002) (AAL).

Connectivity Analysis

Experiment B

The connectivity between the left primary auditory cortex and the 34 regions within the left frontal lobe was estimated using the weighted phase lag index (wPLI). Regions were anatomically defined using the Brainnetome Atlas (Fan et al., 2016), and activity was averaged for all sources within the same region. Primary auditory cortex was defined as BA 41/41, TE 1.0 and TE 1.2. In accordance with the Brainnetome Atlas, the frontal lobe comprised 34 regions: medial BA 8, dorsolateral BA 8, lateral BA 9, medial BA 9, medial BA 6, dorsolateral BA 6, medial BA 10, dorsal BA 9/46, Inferior Frontal Junction, BA 46, ventral BA 9/46, ventrolateral BA8, ventrolateral BA 6, lateral BA 10, dorsal BA 44, Inferior Frontal Sulcus, caudal BA 45, rostral BA 45, opercular BA 44, ventral BA 44, medial BA 14, orbital BA 12/47, lateral BA 11, medial BA 11, BA 13, lateral 12/47, BA 4 head and face region, BA 4 upper limb region, caudal dorsolateral BA 6, BA 4 trunk region, BA 4 tongue and larynx region, caudal ventrolateral BA 6, BA 1/2/3 lower limb region, BA 4 lower limb region.

The wPLI was computed between the left primary auditory cortex activity and the signal originated in each region of the left frontal lobe. First, the cross-spectrum between signals was computed as $X = Z_iFrontalZ_audLeft^*$, where Z represents the Morlet wavelet transform of the signal – centered at 4.5 Hz and with the number of cycles of the wavelet set at 9 (Lachaux et al., 2002). Next, the wPLI square estimator was computed as (Vinck et al., 2011):

$$Y = \frac{Im(X)}{|X|}$$

$$wPLI(f) = \frac{|\sum_{t=1}^{T} Y(f,t)|^2 - \sum_{t=1}^{T} Y(f,t)}{(\sum_{t=1}^{T} |Y(f,t)|)^2 - \sum_{t=1}^{T} Y(f,t)^2}$$

where f is the frequency, t is the discretized time, and T is the total number of time points.

Experiment C

Source space connectivity was computed by multiplying the spectral complex coefficients of each trial (single taper frequency transformation; 0.1111 Hz resolution) with a common filter (DICS; across 2 and 4 Hz), computed across conditions separately for each trial. The debiased weighted phase lag index (Vinck et al., 2011) (dwPLI) was computed between all voxels. Fischer *z*-transformation was applied to normalize the data prior to further analysis. The connectivity between the STG and IFG was computed by averaging the dwPLI values within each ROI. The ROIs were selected based on the AAL (Tzourio-Mazoyer et al., 2002) (*Temporal_Sup_L* and *Frontal_Inf_Tri_L*). The connectivity of the ROI with itself was set to zero.

Data Exclusion Criteria

In all the analyses, data points exceeding two standard deviations were removed. In Experiment C only correct responses were analyzed.

RESULTS

Experiment A: Rightward Dominance Disappears for Speech Rates Deviating From Most Natural During a Syllable Perception Task

Previous studies showing the rightward dominance for speech envelope tracking focused on stimuli with a temporal modulation close to the natural syllable rate (Ding et al., 2017) (i.e., 4.5 syllables per second (Ding et al., 2017)). Here, we explored how the asymmetry is modified when the perceived syllable rate departs from the natural range by testing speech tracking at the typical rate and at the borders of the natural range. The auditory trials consisted of streams of syllables at different rates: 2.5, 3.5, 4.5, 5.5, and 6.5 syllables per second. We estimated, by means of the PLV, the synchronization between the activity in auditory cortex and the envelope of the perceived speech (see Materials and Methods). The results show that synchronization in the right auditory cortex, but not in the left, is modulated by the syllable rate (**Figure 1A**). Furthermore, we found that the auditory coupling asymmetry – defined as $2\left(PLV_{right} - PLV_{left}\right) / \left(PLV_{right} + PLV_{left}\right)$, positive values indicating a rightward asymmetry – is significantly different from zero only for the 4.5 and 5.5 syllables per second conditions (see **Figure 1B**).

Experiment B: The Degree of Asymmetry Correlates With the Strength of Auditory-Frontal Connectivity During a Syllable Perception Task

Assaneo et al. (2019) showed that, while participants listened to a stream of syllables, the rightward dominance of the envelope tracking is strongly diminished in a subset of the subject population. When the data from the whole participant cohort were pooled, the well-known asymmetry is evident (**Figure 2A**). However, when the cohort was segregated into two groups, subjects with high versus low performance on an audio-motor speech synchronization task, the envelope tracking is no longer

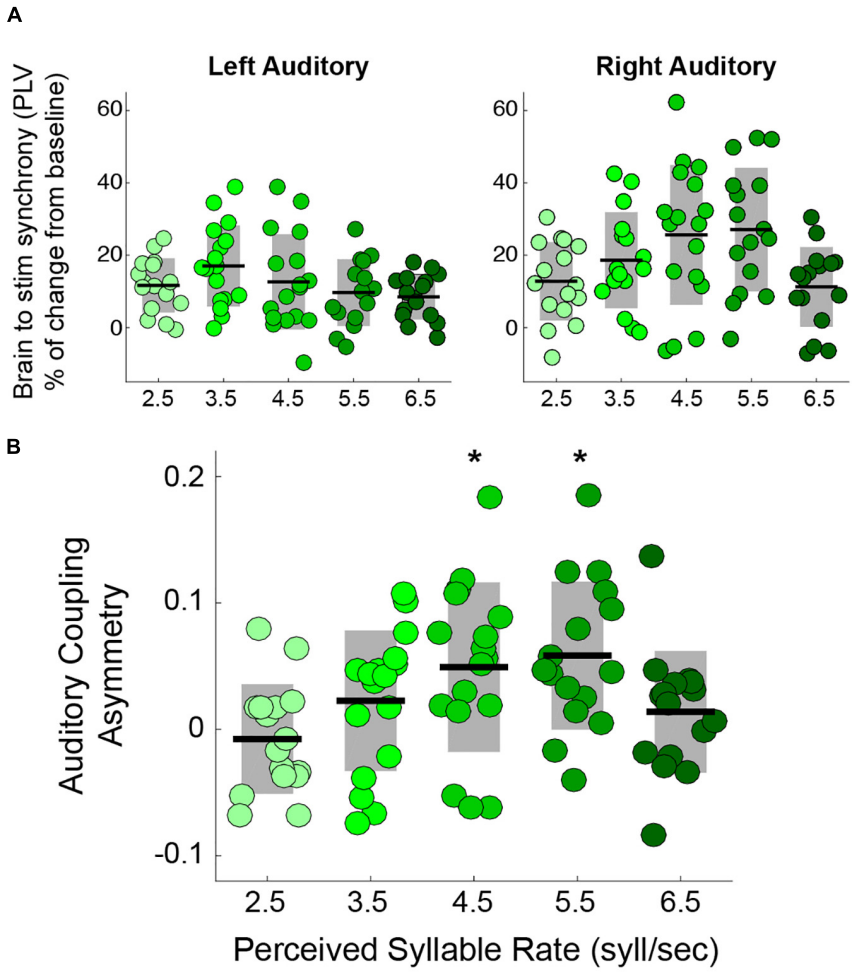

FIGURE 1 | Rightward dominance is affected by speech rate during a syllable perception task. **(A)** PLV between auditory cortices and speech envelope, increment from resting state. Mean PLV around the syllable rate of each condition (syllable rate ± 0.5 Hz). Left auditory synchronization shows no change between conditions (Kruskal-Wallis test: $\chi^2(4) = 5.6$, two-sided $p = 0.23$). However, the right auditory cortex does (Kruskal-Wallis test: $\chi^2(4) = 12.45$, two-sided $p = 0.014$) (Adapted from Assaneo and Poeppel, 2018). **(B)** Auditory coupling asymmetry for the different syllable rate conditions: the degree of asymmetry is modulated by the syllable rate (Kruskal-Wallis test: $\chi^2(4) = 13.63$, two-sided $p = 0.008$). The asymmetry is significantly above zero only for 4.5 and 5.5 syllables per second. * Stands for two-sided $p < 0.05$ (Wilcoxon Signed-Rank test, FDR corrected). Dots: individual participants, the scattering in the *X*-axis is for visualization purposes. Black lines: mean across participants. Shaded region: SD. $N = 17$.

significantly different across hemispheres for *high* performance participants (i.e., it is symmetrical; **Figure 2B**). Moreover, the asymmetry index was different between groups (see **Figure 2C**). We hypothesize that the asymmetry decreased due to stronger auditory-frontal interactions, that is, due to top-down influences from cortical regions external to the auditory cortex, especially to the left auditory cortex. Thus, by means of the wPLI analysis, we estimated the connectivity between the left auditory cortex and frontal regions at 4.5 Hz (see Materials and Methods). We found that left auditory to Brodmann area 45 (BA 45) and left auditory to inferior frontal sulcus (IFS) wPLI significantly correlated with the asymmetry index (**Figures 2D–F**). This result demonstrates that a stronger functional connectivity between left BA 45/IFS and left auditory cortex correlates with a more balanced (i.e., symmetrical) envelope tracking across hemispheres.

Experiment C: Word-Level Linguistic Processing Reverses Rightward Dominance for the Envelope Tracking

Previous research showed that speech intelligibility increases envelope tracking, particularly in the left auditory cortex (Peelle et al., 2013). Here we test how the asymmetry is modified by the presence of semantic information in the auditory stimulus. A cohort of German speakers performed an auditory syllable repetition detection task under two different conditions, defined by the type of stimuli: (i) *Non-Semantic*: Turkish disyllabic pseudo-word streams are presented (no lexical processing); and (ii) *Semantic*: German di-syllabic words streams are presented. This protocol allows us to explore how lexical-semantic processing modifies the asymmetry, independent of the task being performed – note that the task remained the same

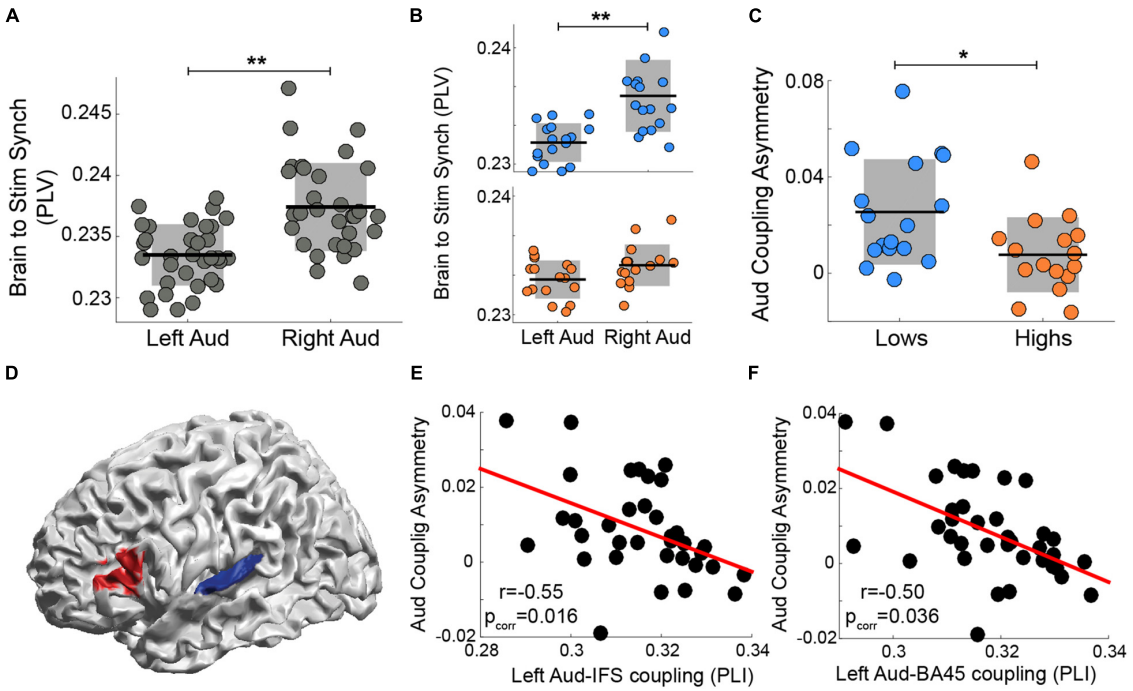

FIGURE 2 | The degree of asymmetry correlates with the strength of auditory-frontal connectivity during a syllable perception task. Synchronization (PLV) between auditory cortex activity and the perceived speech envelope in left and right hemispheres: **(A)** all subjects pooled and **(B)** low synchronizers in the upper (blue) panel versus high synchronizers in the lower panel. **(C)** Auditory coupling asymmetry: comparison between groups. [**B,C** reanalyzed and replotted from Assaneo et al. (2019)] **(D)** Connectivity at 4.5 Hz (wPLI) between early auditory cortex (BA 41/41, TE 1.0 and TE 1.2 in blue) and frontal regions (caudal BA 45 and IFS in red) correlates with the degree of auditory-to-stimulus coupling asymmetry (Spearman correlation, two-sided $p < 0.05$, FDR-corrected). Scatter plots of the correlation between auditory-to-stimulus asymmetry and the wPLI between left auditory cortex and areas highlighted in red: **(E)** inferior frontal sulcus (IFS; $N = 37$) and **(F)** caudal BA 45 ($N = 36$). Orange/blue corresponds to high/low synchronizers, respectively. ** Stands for two-sided $p < 0.005$ (Wilcoxon signed-rank test), * for two-sided $p < 0.05$ (Mann-Whitney-Wilcoxon test). Dots: individual participants; in panels **A–C** the scattering in the X-axis is for visualization purposes. Black lines: mean across participants. Shaded region in panels **A–C**: SD. Red line: linear regression.

across conditions. First, we quantified the coherence between primary auditory cortex activity and the envelope of the perceived auditory stimulus, in a frequency band around the syllable rate (4 Hz). The results show that there was a significant difference in the asymmetry index for the two conditions of interest (**Figure 3A**). For the *Non-Semantic* condition, the asymmetry exposed a trend for the classical rightward dominance (right panel); for the *Semantic* condition the asymmetry was reversed (**Figure 3B**). Finally, we examined if there was a correlation between the auditory tracking and left fronto-temporal connectivity. For the *Non-Semantic* condition, the data showed the same trend as observed in Experiment B – the stronger the connectivity with the left auditory cortex, the weaker the rightward dominance (i.e., the more symmetrical the tracking was; see **Figure 3C**, right panel). However, the fronto-temporal connectivity did not seem to be directly related to the reversed asymmetry in the *Semantic* condition (see **Figure 3C**, left panel).

DISCUSSION

The neural architecture that forms the basis of speech processing is structurally and functionally complex, comprising a suite of computations that perform necessary subroutines on input and output processes. The tracking of the speech envelope by the auditory cortex has been proposed to be one of the basic mechanisms for the subsequent decoding of the signal. A common assumption is that envelope tracking is stronger in the right hemisphere than in the left – an effect that has been theoretically proposed (Poeppel, 2003) and experimentally demonstrated (Luo and Poeppel, 2007; Abrams et al., 2008; Doelling et al., 2014). However, here we show three examples in which the canonical rightward dominance is broken.

We hypothesize that in addition to the oscillatory activity intrinsic to the auditory cortex, speech tracking depends on input from other non-auditory brain regions, and we propose different functional roles for these mechanisms – i.e., *intrinsic auditory* and top down *externally driven* mechanisms. Lateralization depends on the extent to what these mechanisms are engaged to support a specific task.

Right Hemisphere Tracking Preference Disappears for Non-natural Speech Rates During Syllable Perception

The AST theory (Poeppel, 2003) – as well as similar theories (Zatorre et al., 2002; Flinker et al., 2019) – builds on

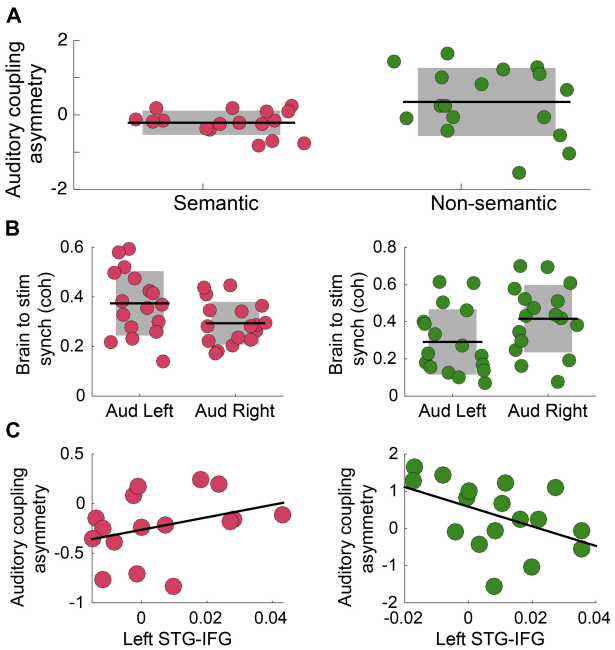

FIGURE 3 | Semantic access reverses the classical rightward dominance for the envelope tracking. **(A)** Auditory coupling asymmetry index: comparison between conditions (N = 17; two-sided p = 0.035, paired Wilcoxon Signed-Rank test). **(B)** Coherence between *Heschl's* gyrus (right and left hemispheres) activity and the auditory stimulus envelope, in a frequency band around 4 Hz. Left panel: *Semantic Condition* (N = 17; two-sided p = 0.010, paired Wilcoxon Signed Rank test). Right panel: *Non-Semantic Condition* (N = 17; two-sided p = 0.15, paired Wilcoxon Signed-Rank test). **(C)** Scatterplot of the auditory coupling asymmetry as a function of the connectivity between left STG- IFG, in a frequency band around 4 Hz. Left panel: *Semantic Condition* (N = 16; Spearman correlation coefficient r = 0.4, two-sided p = 0.12). Right panel: *Non-Semantic Condition* (N = 17; Spearman correlation coefficient r = –0.46, two-sided p = 0.063). In all panels: Pink/green correspond to *Semantic/Non-Semantic* (German words/Turkish pseudo-words) respectively, and dots: individual participants. In panels **A,B**: the shaded region represents SD, the black line the mean across participants and the dots scattering in the X-axis is for visualization purposes. In panel **C:** the black line represents the linear regression.

cytoarchitectonic differences between the (primary and non-primary) auditory cortices of the left and right hemispheres (Hutsler and Galuske, 2003). It proposes that those differences grant specific oscillatory properties to each hemisphere. Specifically, due to the biophysical properties of the neural populations, neuronal activity within the right auditory cortex shows characteristics of a neural oscillator with a natural frequency between 3 and 6 Hz. The basic features defining an oscillator are (Pikovsky et al., 2003): (i) it generates rhythmic activity at its own natural frequency, which is defined by the internal properties of the system; and (ii) it entrains to other oscillators or, as relevant in the speech case, synchronizes to external rhythmic stimuli, within a restricted range of frequencies close enough to its natural one.

Previous work found slow rhythmic neuronal activity during resting state within right auditory cortices (Giraud et al., 2007; Morillon et al., 2012), suggesting a neuronal population behaving as a low-frequency oscillator – consistent with criterion (i). Here, in line with criterion (ii), we found a tuning curve for the synchronization of the right auditory cortex (**Figure 1A**), with enhanced values for stimulus rates between 4 and 6 Hz. This was not found for the left auditory cortex, as presumably its natural frequency range differs. We propose that the brain-to-envelope coupling in the right hemisphere is driven by the oscillatory features of the auditory cortex, which are tuned to maximally resonate (in phase space) in response to frequencies close to the natural syllable rate (Ding et al., 2017). Thus, when the stimulation rate departs from the *natural* frequency of this area, the right cortex is less responsive and the tracking asymmetry disappears (**Figure 1B**). Furthermore, we hypothesize that the tuning curve obtained here is not inflexible; we believe that the range of entrainment could be extended under different task demands. According to previous proposals, the function of the right hemispheric speech tracking is related to the decoding of phonetic/spectral features of the audio signal (Zatorre and Belin, 2001; Poeppel, 2003). From a mathematical point of view, the entrainment range of an oscillator can be extended (Pikovsky et al., 2003) by increasing the strength of the coupling between the oscillator and the external driving force (in this case the auditory stimulus). Bringing those points together, we speculate that by modifying the goal of the task (e.g., a pitch perception or voice identity task, instead of the syllable perception task with a working memory component performed here) the right auditory-to-envelope coupling could be strengthened and the asymmetry could be recovered even for the less optimal modulation frequencies.

During Syllable Perception Hemispheric Asymmetry Correlates With Auditory-Frontal Connectivity

Different functional roles have been attributed to speech envelope tracking: (i) segmentation of the input stream into temporal units of the appropriate granularity for subsequent decoding (Ghitza, 2014); (ii) extracting paralinguistic information; and (iii) integration of smaller phoneme-like units into larger syllable-like units for the subsequent phonological decoding. While (ii) is preferentially conducted by the right hemisphere, (iii) is more represented in the left (Giraud and Poeppel, 2012). We propose that the right hemisphere envelope tracking mostly reflects *intrinsic auditory* (bottom-up) oscillatory activity, while the left tracking (at the syllabic rate) is preferentially driven by cortical areas outside of auditory cortex (*externally driven*, top-down mechanism). **Table 1** summarizes these conjectures.

In Experiment B, on this view, the observed tracking reflects both *intrinsic auditory* and *externally driven* influences. On the one hand, since the temporal properties of the acoustic signal (i.e., the syllable rate of 4.5 Hz) match the natural frequency of the right auditory cortex, the right-lateralized intrinsic oscillatory mechanism is activated. On the other hand, the phonological processing required to complete the syllable perception task activates the *externally driven* mechanism. Thus, the envelope tracking lateralization is determined by the interplay between the recruited mechanisms. Our findings

TABLE 1 | Different origins for the observed auditory-to-envelope synchronization.

	Function	Dominant Hemisphere	Nature
(i)	Paralinguistic/Phonetic	Right	Intrinsic auditory
(ii)	Phonological/Semantic	Left	Externally driven

Speech envelope tracking is generated by distinct mechanisms. Each mechanism supports different functional roles and is asymmetrically represented across hemispheres. We define as externally driven a process relying on an interplay between the auditory cortex and other brain areas, while intrinsic refers to activity that reflects the internal oscillatory features of the auditory cortices.

provide insight into how these influences interact and suggest that individual differences also play a role in the contribution of both mechanisms. Interestingly, while a part of the population shows the classic rightward dominance for the speech tracking (**Figure 2A**; note that when pooling together data from all participants this effect is observed), a subgroup of the population – with enhanced microstructural properties in the white matter pathways connecting the left auditory cortex with frontal regions (Assaneo et al., 2019)- displays no asymmetry in tracking (**Figure 2B**). We suggest that for this group, due to functional and structural differences, the influence of the *externally driven* mechanism is enhanced, equilibrating the tracking across hemispheres (**Figure 2C**). Moreover, the correlation between speech envelope tracking asymmetry and fronto-auditory connectivity (**Figure 2D**) supports the claim that *externally driven* top-down influences from the left frontal cortex to the left auditory one reverse the classical right hemispheric dominance. The same trend is found in Experiment C for the condition wherein a phonological task is performed on random streams of syllables (Turkish pseudo-words condition; note that this condition resembles the paradigm of Experiment B, **Figure 3C** right panel).

Semantic Processing Reverses the Right Hemisphere Dominance

In Experiment C, as in the previously discussed study, speech tracking in both conditions presumably reflects both *intrinsic auditory* and *externally driven* contributions: on the one hand, the syllables are presented at a rate of 4 Hz – close to the natural frequency of the right auditory cortex – and on the other hand, the task requires phonological processing of the signal. During the condition in which German words were presented, additional lexical-semantic computations are necessarily performed (Rimmele et al., 2019).

Here, we show that, even though the task remains the same across conditions, and although the acoustic properties of the stimuli are similar, the asymmetry of the auditory tracking is reversed when semantic information is present (see **Figures 3A,B** left panel). We propose that semantic processing further enhances the envelope tracking performed by the left hemisphere, and thus reduces the right hemispheric dominance. This proposal aligns well with previous studies showing that the auditory to speech synchronization increases with intelligibility, specifically in the left hemisphere (Ahissar et al., 2001; Peelle et al., 2013; Rimmele et al., 2015).

Note that in Experiment C (in the German words condition), in spite of a reduced hemispheric asymmetry we found no correlation between fronto-temporal connectivity and the asymmetry index (as we do in Experiment B). Different reasons can underpin the dissimilarity between results. On one side, semantic access is a complex process – as compared to syllable perception – relying on large-scale brain networks (Hickok and Poeppel, 2000; Scott et al., 2002; Binder et al., 2009; Rodd et al., 2015) then, the asymmetry reduction can derive from the connectivity between temporal areas and a different region of the brain. On the other side, the task in Experiment B contains a working memory component, while the task in Experiment C does not load high on working memory. Further research is required to clarify the complex connectivity patterns between auditory cortex and other brain regions underpinning the hemispheric asymmetry and to investigate whether the correlation between the fronto-temporal connectivity and the asymmetry index might be related to working memory mechanisms.

It is worth noting that we employed distinct methods in the current work – different experimental designs, phase synchrony measurements and source reconstruction techniques. The results presented here derived from three already published studies. Thus, we chose to adopt for each analysis the approach applied in the original work. We believe that the fact that different experimental designs and methodological approaches show converging results further strengthens the reliability of our hypothesis.

To summarize, speech tracking (measured as auditory-cortex-to-speech envelope synchronization) is a complex process determined by an interplay between the intrinsic properties of the auditory cortices (Zatorre et al., 2002; Giraud et al., 2007) and top down influences from other non-auditory cortical areas related to different factors such as speech intelligibility, attention and/or acoustic properties of the perceived signal (Zion Golumbic et al., 2013; Ding et al., 2014; Zoefel and VanRullen, 2015; Bidelman and Howell, 2016). Moreover, individual differences in neural function and structure can also strongly affect the symmetry of the speech tracking between the hemispheres. Crucially, the *intrinsic auditory* and *externally driven* influences differently affect the hemispheric lateralization patterns of the speech tracking in the auditory cortex. Our findings illustrate the interaction between the different influences on speech tracking and suggest that the observed hemispheric lateralization patterns depend in subtle ways on task demands and the properties of the auditory signal. However, understanding the distinct origins of the assessed synchrony requires further research.

DATA AVAILABILITY

The datasets for this manuscript are not publicly available because data available upon request. Requests to access the datasets should be directed to fassaneo@gmail.com.

AUTHOR CONTRIBUTIONS

All authors designed the research. MA, JR, JO, and PR acquired the data. MA and JR analyzed the data and wrote the manuscript.

FUNDING

This work was supported by NIH grant 2R01DC05660 (DP) and FP7 Ideas: European Research Council grant ERC-StG-313841 (RdD-B).

REFERENCES

Abrams, D. A., Nicol, T., Zecker, S., and Kraus, N. (2008). Right-hemisphere auditory cortex is dominant for coding syllable patterns in speech. *J. Neurosci.* 28, 3958–3965. doi: 10.1523/JNEUROSCI.0187-08.2008

Abrams, D. A., Nicol, T., Zecker, S., and Kraus, N. (2009). Abnormal cortical processing of the syllable rate of speech in poor readers. *J. Neurosci.* 29, 7686–7693. doi: 10.1523/JNEUROSCI.5242-08.2009

Adachi, Y., Shimogawara, M., Higuchi, M., Haruta, Y., and Ochiai, M. (2001). Reduction of non-periodic environmental magnetic noise in MEG measurement by continuously adjusted least squares method. *IEEE Tran. Appl. Supercond.* 11, 669–672. doi: 10.1109/77.919433

Ahissar, E., Nagarajan, S., Ahissar, M., Protopapas, A., Mahncke, H., Merzenich, M. M., et al. (2001). Speech comprehension is correlated with temporal response patterns recorded from auditory cortex. *Proc. Natl. Acad. Sci. U. S. A* 98, 13367–13372. doi: 10.1073/pnas.201400998

Assaneo, M. F., and Poeppel, D. (2018). The coupling between auditory and motor cortices is rate-restricted: evidence for an intrinsic speech-motor rhythm. *Sci. Adv.* 4:eaao3842. doi: 10.1126/sciadv.aao3842

Assaneo, M. F., Ripollés, P., Orpella, J., Lin, W. M., de Diego-Balaguer, R., and Poeppel, D. (2019). Spontaneous synchronization to speech reveals neural mechanisms facilitating language learning. *Nat. Neurosci.* 22, 627–632. doi: 10.1038/s41593-019-0353-z

Baayen, R. H., Piepenbrock, R., and Gulikers, L. (1995). *The CELEX Lexical Database (CD-ROM)*. Philadelphia, PA: Philadelphia Linguistics Data Consortium University of Pennsylvania.

Bidelman, G. M., and Howell, M. (2016). Functional changes in inter- and intra-hemispheric cortical processing underlying degraded speech perception. *Neuroimage* 124((Pt A), 581–590. doi: 10.1016/j.neuroimage.2015.09.09.020

Binder, J. R., Desai, R. H., Graves, W. W., and Conant, L. L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cereb. Cortex* 19, 2767–2796. doi: 10.1093/cercor/bhp055

Binder, J. R., Frost, J. A., Hammeke, T. A., Bellgowan, P. S., Springer, J. A., Kaufman, J. N., et al. (2000). Human temporal lobe activation by speech and nonspeech sounds. *Cereb. Cortex* 10, 512–528. doi: 10.1093/cercor/10.5.512

Boemio, A., Fromm, S., Braun, A., and Poeppel, D. (2005). Hierarchical and asymmetric temporal sensitivity in human auditory cortices. *Nat. Neurosci.* 8, 389–395. doi: 10.1038/nn1409

Boersma, P. (2001). PRAAT, a system for doing phonetics by computer. *Glot Int.* 5, 341–345.

Bozkurt, B., Dutoit, T., and Pagel, V. (1996). "Le projet MBROLA : vers un ensemble de synthétiseurs vocaux disponibles gratuitement pour utilisation non-commerciale," in *Proceedings of the* Actes des JEP *XXIèmes Journées d'Etude sur la Parole*, (Avignon), 441–444.

Cogan, G. B., Thesen, T., Carlson, C., Doyle, W., Devinsky, O., Pesaran, B., et al. (2014). Sensory-motor transformations for speech occur bilaterally. *Nature* 507, 94–98. doi: 10.1038/nature12935

Crosse, M. J., Butler, J. S., and Lalor, E. C. (2015). Congruent visual speech enhances cortical entrainment to continuous auditory speech in noise-free conditions. *J. Neurosci.* 35, 14195–14204. doi: 10.1523/JNEUROSCI.1829-15.2015

Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., et al. (2000). Dynamic statistical parametric mapping: combining fmri and meg for high-resolution imaging of cortical activity. *Neuron* 26, 55–67.

de Cheveigné, A., and Simon, J. Z. (2007). Denoising based on time-shift PCA. *J. Neurosci. Methods* 165, 297–305. doi: 10.1016/j.jneumeth.2007.06.003

Ding, N., Chatterjee, M., and Simon, J. Z. (2014).). Robust cortical entrainment to the speech envelope relies on the spectro-temporal fine structure. *Neuroimage* 88, 41–46. doi: 10.1016/j.neuroimage.2013.10.054

Ding, N., Patel, A., Chen, L., Butler, H., Luo, C., and Poeppel, D. (2017). Temporal modulations in speech and music. *Neurosci. Biobehav. Rev.* 81(Pt B), 181–187. doi: 10.1016/j.neubiorev.2017.02.011

Doelling, K. B., Arnal, L. H., Ghitza, O., and Poeppel, D. (2014). Acoustic landmarks drive delta-theta oscillations to enable speech comprehension by facilitating perceptual parsing. *Neuroimage* 85, 761–768. doi: 10.1016/j.neuroimage.2013.06.035

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., et al. (2016). The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cereb. Cortex* 26, 3508–3526. doi: 10.1093/cercor/bhw157

Federmeier, K. D. (2007). Thinking ahead: the role and roots of prediction in language comprehension. *Psychophysiology* 44, 491–505. doi: 10.1111/j.1469-8986.2007.00531.x

Flinker, A., Doyle, W. K., Mehta, A. D., Devinsky, O., and Poeppel, D. (2019). Spectrotemporal modulation provides a unifying framework for auditory cortical asymmetries. *Nat. Hum. Behav.* 3, 393–405. doi: 10.1038/s41562-019-0548-z

Friederici, A. D. (2012). The cortical language circuit: from auditory perception to sentence comprehension. *Trends Cog. Sci.* 16, 262–268. doi: 10.1016/j.tics.2012.04.001

Ghitza, O. (2014). Behavioral evidence for the role of cortical θ oscillations in determining auditory channel capacity for speech. *Front. Psychol.* 5:652. doi: 10.3389/fpsyg.2014.00652

Giraud, A. L., Kleinschmidt, A., Poeppel, D., Lund, T. E., Frackowiak, R. S., Laufs, H., et al. (2007). Endogenous cortical rhythms determine cerebral specialization for speech perception and production. *Neuron* 56, 1127–1134. doi: 10.1016/j.neuron.2007.09.038

Giraud, A. L., Lorenzi, C., Ashburner, J., Wable, J., Johnsrude, I., Frackowiak, R., et al. (2000). Representation of the temporal envelope of sounds in the human brain. *J. Neurophysiol.* 84, 1588–1598. doi: 10.1152/jn.2000.84.3.1588

Giraud, A. L., and Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nature. Neurosci.* 15, 511–517. doi: 10.1038/nn.3063

Gross, J., Kujala, J., Hämäläinen, M., Timmermann, L., Schnitzler, A., and Salmel, R. (2001). Dynamic imaging of coherent sources: studying neural interactions in the human brain. *Proc. Natl. Acad. Sci. U.S.A.* 98, 694–699. doi: 10.1073/pnas.98.2.694

Hickok, G., and Poeppel, D. (2000). Towards a functional neuroanatomy of speech perception. *Trends Cog. Sci.* 4, 131–138. doi: 10.1016/S1364-6613(00)01463-7

Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nature Rev. Neurosci.* 8, 393–402. doi: 10.1038/nrn2113

Hutsler, J., and Galuske, R. A. (2003). Hemispheric asymmetries in cerebral cortical networks. *Trends Neurosci.* 26, 429–435. doi: 10.1016/S0166-2236(03)00198-X

Kubanek, J., Brunner, P., Gunduz, A., Poeppel, D., and Schalk, G. (2013). The tracking of speech envelope in the human cortex. *PLoS One* 8:e53398. doi: 10.1371/journal.pone.0053398

Lachaux, J.-P., Eugenio, R., Le van Quyen, M., Antoine, L., Martinerie, J., and Varela, Francisco J. (2002). Studying single-trial of phase synchronous activity in the brain. *Int. J. Bifurc. Chaos* 10, 2429–2439. doi: 10.1142/s0218127400001560

Lopez-Barroso, D., Catani, M., Ripollés, P., Dell'Acqua, F., Rodríguez-Fornells, A., de Diego-Balaguer, R., et al. (2013). Word learning is mediated by the left arcuate fasciculus. *Proc. Natl. Acad. Sci. U.S.A.* 110, 13168–13173. doi: 10.1073/pnas.1301696110

Lopez-Barroso, D., de Diego-Balaguer, R., Cunillera, T., Camara, E., Münte, T. F., Rodriguez-Fornells, A., et al. (2011). Language learning under working memory constraints correlates with microstructural differences in the ventral language pathway. *Cereb. Cortex* 21, 2742–2750. doi: 10.1093/cercor/bhr064

Luo, H., and Poeppel, D. (2007). Phase patterns of neuronal responses reliably discriminate speech in human auditory cortex. *Neuron* 54, 1001–1010. doi: 10.1016/j.neuron.2007.06.004

Makeig, S., Bell, A. J., Jung, T.-P., and Sejnowski, T. J. (1996). "Independent Component Analysis of Electroencephalographic Data," in *Proceedings of the 8th International Conference on Neural Information Processing Systems*, (Cambridge, MA: MIT Press), doi: 10.1109/ICOSP.2002.1180091

Morillon, B., and Baillet, S. (2017). Motor origin of temporal predictions in auditory attention. *Proc. Natl. Acad. Sci. U.S.A.* 114, E8913–E8921. doi: 10.1073/pnas.1705373114

Morillon, B., Liégeois-Chauvel, C., Arnal, L. H., Bénar, C. G., and Giraud, A. L. (2012). Asymmetric function of theta and gamma activity in syllable processing: an intra-cortical study. *Front. Psychol.* 3:248. doi: 10.3389/fpsyg.2012.00248

Nolte, G. (2003). The magnetic lead field theorem in the quasi-static approximation and its use for magnetoenchephalography forward calculation in realistic volume conductors. *Phys. Med. Biol.* 48, 3637–3652. doi: 10.1088/0031-9155/48/22/002

Oostenveld, R., Fries, P., Maris, E., and Schoffelen, J. M. (2011). FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput. Intell. Neurosci.* 2011:156869. doi: 10.1155/2011/156869

Park, H., Ince, R. A., Schyns, P. G., Thut, G., and Gross, J. (2015). Frontal top-down signals increase coupling of auditory low-frequency oscillations to continuous speech in human listeners. *Curr. Biol.* 25, 1649–1653. doi: 10.1016/j.cub.2015.04.049

Park, H., Thut, G., and Gross, J. (2018). Predictive entrainment of natural speech through two fronto-motor top-down channels. *Lang. Cogn. Neurosci.* 1–13. doi: 10.1080/23273798.2018.1506589

Peelle, J. E., Gross, J., and Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cereb. Cortex* 23, 1378–1387. doi: 10.1093/cercor/bhs118

Pikovsky, A., Rosenblum, M., and Kurths, J. (2003). *Synchronization: a universal concept in nonlinear sciences*. Cambridge: University Press, 432. doi: 10.1063/1.1554136

Poeppel, D. (2003). The analysis of speech in different temporal integration windows: cerebral lateralization as 'asymmetric sampling in time'. *Speech Commun.* 41, 245–255. doi: 10.1016/s0167-6393(02)00107-3

Rauschecker, J. P., and Scott, S. K. (2009). Maps and streams in the auditory cortex: nonhuman primates illuminate human speech processing. *Nature Neurosci.* 12, 718–724. doi: 10.1038/nn.2331

Rimmele, J. M., Gross, J., Molholm, S., and Keitel, A. (2018a). Editorial: brain oscillations in human communication. *Front. Hum. Neurosci.* 12:39. doi: 10.3389/fnhum.2018.00039

Rimmele, J. M., Morillon, B., Poeppel, D., and Arnal, L. H. (2018b). Proactive sensing of periodic and aperiodic auditory patterns. *Trends Cog. Sci.* 22, 870–882. doi: 10.1016/j.tics.2018.08.003

Rimmele, J. M., Sun, Y., Michalareas, G., Ghitza, O., and Poeppel, D. (2019). Dynamics of functional networks for syllable and word-level processing. *BioRxiv* 584375.

Rimmele, J. M., Zion Golumbic, E., Schröger, E., and Poeppel, D. (2015). The effects of selective attention and speech acoustics on neural speech-tracking in a multi-talker scene. *Cortex.* 68, 144–154. doi: 10.1016/j.cortex.2014.12.014

Rodd, J. M., Vitello, S., Woollams, A. M., and Adank, P. (2015). Localising semantic and syntactic processing in spoken and written language comprehension: an activation likelihood estimation meta-analysis. *Brain Lang.* 141, 89–102. doi: 10.1016/j.bandl.2014.11.012

Sammler, D., Grosbras, M. H., Anwander, A., Bestelmeyer, P. E., and Belin, P. (2015). Dorsal and ventral pathways for prosody. *Curr. Biol.* 25, 3079–3085. doi: 10.1016/j.cub.2015.10.009

Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M. S., et al. (2008). Ventral and dorsal pathways for language. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18035–18040. doi: 10.1073/pnas.0805234105

Scott, S. K., Blank, C. C., Rosen, S., and Wise, R. J. (2002). Identification of a pathway for intelligible speech in the left temporal lobe. *Brain* 123(Pt 12), 2400–2406. doi: 10.1093/brain/123.12.2400

Smith, Z. M., Delgutte, B., and Oxenham, A. J. (2002). Chimaeric sounds reveal dichotomies in auditory perception. *Nature* 416, 87–90. doi: 10.1038/416087a

Stolk, A., Todorovic, A., Schoffelen, J.-M., and Oostenveld, R. (2013). Online and offline tools for head movement compensation in MEG. *Neuroimage* 68, 39–48. doi: 10.1016/j.neuroimage.2012.11.047

Telkemeyer, S., Rossi, S., Koch, S. P., Nierhaus, T., Steinbrink, J., Poeppel, D., et al. (2009). Sensitivity of newborn auditory cortex to the temporal structure of sounds. *J. Neurosci.* 29, 14726–14733. doi: 10.1523/JNEUROSCI.1246-09.2009

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978

Vanvooren, S., Poelmans, H., Hofmann, M., Ghesquière, P., and Wouters, J. (2014). Hemispheric asymmetry in auditory processing of speech envelope modulations in prereading children. *J. Neurosci.* 34, 1523–1529. doi: 10.1523/JNEUROSCI.3209-13.2014

Vinck, M., Oostenveld, R., van Wingerden, M., Battaglia, F., and Pennartz, C. M. (2011). An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias. *Neuroimage* 55, 1548–1565. doi: 10.1016/j.neuroimage.2011.01.055

Zatorre, R. J., and Belin, P. (2001). Spectral and temporal processing in human auditory cortex. *Cereb. Cortex* 11, 946–953. doi: 10.1093/cercor/11.10.946

Zatorre, R. J., Belin, P., and Penhune, V. B. (2002). Structure and function of auditory cortex: music and speech. *Trends Cog. Sci.* 6, 37–46. doi: 10.1016/S1364-6613(00)01816-7

Zion Golumbic EM, Ding, N., Bickel, S., Lakatos, P., Schevon, C. A., McKhann, G. M., et al. (2013). Mechanisms underlying selective neuronal tracking of attended speech at a 'cocktail party'. *Neuron* 77, 980–991. doi: 10.1016/j.neuron.2012.12.037

Zoefel, B., Archer-Boyd, A., and Davis, M. H. (2018). Phase Entrainment of Brain Oscillations Causally Modulates Neural Responses to Intelligible Speech. *Curr. Biol* 28, 401.e5–e408. doi: 10.1016/j.cub.2017.11.071

Zoefel, B., and VanRullen, R. (2015). The Role of High-Level Processes for Oscillatory Phase Entrainment to Speech Sound. *Front. Hum. Neurosci* 9:651. doi: 10.3389/fnhum.2015.00651

Conflict of Interest Statement: 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.

Copyright © 2019 Assaneo, Rimmele, Orpella, Ripollés, de Diego-Balaguer and Poeppel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Integrative Neuroscience

CORRECTION
published: 04 February 2020
doi: 10.3389/fnint.2020.00002

Corrigendum: The Lateralization of Speech-Brain Coupling Is Differentially Modulated by Intrinsic Auditory and Top-Down Mechanisms

M. F. Assaneo[1]*, *J. M. Rimmele*[2], *J. Orpella*[1], *P. Ripollés*[1], *R. de Diego-Balaguer*[3,4,5,6] *and D. Poeppel*[1,2]

[1] Department of Psychology, New York University, New York, NY, United States, [2] Department of Neuroscience, Max Planck Institute for Empirical Aesthetics, Frankfurt, Germany, [3] Departament de Cognició, Desenvolupament i Psicologia de l'Educació, University of Barcelona, Barcelona, Spain, [4] Catalan Institute for Research and Advance Studies, Barcelona, Spain, [5] Cognition and Brain Plasticity Unit, IDIBELL, L'Hospitalet de Llobregat, Spain, [6] Institute of Neuroscience, University of Barcelona, Barcelona, Spain

Keywords: asymmetrical sampling, brain to stimulus synchronization, MEG (magnetoencephalography), speech perception, speech envelope tracking

OPEN ACCESS

Approved by:
Frontiers Editorial Office, Frontiers Media SA, Switzerland

***Correspondence:**
M. F. Assaneo
fassaneo@gmail.com

Received: *07 January 2020*
Accepted: *09 January 2020*
Published: *04 February 2020*

Citation:
Assaneo MF, Rimmele JM, Orpella J, Ripollés P, de Diego-Balaguer R and Poeppel D (2020) Corrigendum: The Lateralization of Speech-Brain Coupling Is Differentially Modulated by Intrinsic Auditory and Top-Down Mechanisms. Front. Integr. Neurosci. 14:2. doi: 10.3389/fnint.2020.00002

A Corrigendum on

The Lateralization of Speech-Brain Coupling Is Differentially Modulated by Intrinsic Auditory and Top-Down Mechanisms

by Assaneo, M. F., Rimmele, J. M., Orpella, J., Ripollés, P., de Diego-Balaguer, R., and Poeppel, D. (2019). Front. Integr. Neurosci. 13:28. doi: 10.3389/fnint.2019.00028

In the published article, there was an error regarding the affiliations for R. de Diego-Balaguer. As well as having affiliation 3, she should also have:

[4] Catalan Institute for Research and Advance Studies, Barcelona, Spain
[5] Cognition and Brain Plasticity Unit, IDIBELL, L'Hospitalet de Llobregat, Spain
[6] Institute of Neuroscience, University of Barcelona, Barcelona, Spain

The authors apologize for this error and state that this does not change the scientific conclusions of the article in any way. The original article has been updated.

Copyright © 2020 Assaneo, Rimmele, Orpella, Ripollés, de Diego-Balaguer and Poeppel. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 26 September 2019
doi: 10.3389/fpsyg.2019.02198

The Inter-Regional Connectivity Within the Default Mode Network During the Attentional Processes of Internal Focus and External Focus: An fMRI Study of Continuous Finger Force Feedback

Zhi-Wei Zhou[1,2,3†], Xia-Qing Lan[1,2,3†], Yan-Tong Fang[2,3], Yun Gong[2,3], Yu-Feng Zang[1,2,3], Hong Luo[2,3] and Hang Zhang[1,2,3*]*

[1] *Institute of Psychological Sciences, College of Education, Hangzhou Normal University, Hangzhou, China,* [2] *Center for Cognition and Brain Disorders and the Affiliated Hospital, Hangzhou Normal University, Hangzhou, China,* [3] *Zhejiang Key Laboratory for Research in Assessment of Cognitive Impairments, Hangzhou, China*

OPEN ACCESS

Edited by:
Andreas Bahmer, Universitätsklinikum Würzburg, Germany

Reviewed by:
Kewei Chen, Banner Alzheimer's Institute, United States
Judson Brewer, Brown University, United States

***Correspondence:**
Hong Luo
luohong@vip.126.com
Hang Zhang
kevinhangbnu@foxmail.com

† These authors have contributed equally to this work

Specialty section:
This article was submitted to Cognitive Science, a section of the journal Frontiers in Psychology

Received: *03 June 2019*
Accepted: *12 September 2019*
Published: *26 September 2019*

Citation:
Zhou Z-W, Lan X-Q, Fang Y-T, Gong Y, Zang Y-F, Luo H and Zhang H (2019) The Inter-Regional Connectivity Within the Default Mode Network During the Attentional Processes of Internal Focus and External Focus: An fMRI Study of Continuous Finger Force Feedback. Front. Psychol. 10:2198. doi: 10.3389/fpsyg.2019.02198

Sustained attention involves two distinct processes, i.e., external focus and internal focus. Some recent neuroimaging studies employed the instruction of experimenters or the self-report from participants to generate the two attentional processes, and observed that the default mode network (DMN) was also responding to the external focus. These observations challenged the general view that the DMN accounts for the internally directed cognition, e.g., unfocused mind wandering, task independent-thoughts and internally focused events. Notably, the instruction or self-report may not effectively ensure the participants engage in the external focus/internal focus, and thus, the functional significance of the DMN for the externally focused process remains to be verified. In the present study, a new task paradigm, i.e., real/sham continuous feedback of finger force, was employed to generate the attentional process of external focus/internal focus, and the functional connectivity among the node regions of the DMN was further investigated in the two processes respectively. We found that two regions of the DMN, posterior cingulate cortex and left inferior parietal cortex/angular gyrus showed stronger inter-regional connectivity in the externally focused process than it in the internally focused process. Intriguingly, this functional connectivity was closely related to the behavioral performance in the process of external focus. These findings implicated that the functional significance of the DMN in sustained attention was more than responding to the internally directed cognition, and the task paradigm of continuous finger force feedback could benefit for the future studies on the externally focused/internally focused process of sustained attention.

Keywords: sustained attention, external focus/internal focus, fMRI, continuous finger force feedback, functional connectivity

INTRODUCTION

Almost every task in the daily life requires human to maintain the attentional focus. Sustaining attention over prolonged periods of time is of great interest to cognitive psychologists. This is recognized by the fact that the sustained attention is critical for successful cognitive processing (Coull et al., 1996), and a number of psychiatric disorders, e.g., attention deficit hyperactivity disorder (Liu et al., 2015), schizophrenia (Burton et al., 2018), depression (Ye et al., 2018), etc., always show symptoms of deficits in sustained attention. The increasing number of evidences indicate that sustained attention could be dissociated into two attentional processes based on whether the attentional focus is external or internal (Chun et al., 2011; Fortenbaugh et al., 2017). The attentional process of external focus refers to external information of tasks, e.g., locations in space, the shape of targets, etc., while the process of internal focus refers to the internal experience of tasks, such as rules, decisions, responses, etc (Chun et al., 2011). Behavioral investigations suggested that the externally focused process could promote the performance in many sport items, e.g., volleyball, dart throwing and so forth (Wulf et al., 2002; Shafizadeh et al., 2013), and the internally focused process benefits the performance when the context was related to specific movement form rather than the performance outcomes (Poolton et al., 2006; Schücker et al., 2014). Of note, these intriguing findings came from investigations on quite different paradigms, such as probe-detection paradigm (Mansell et al., 2003), thought-probe paradigm (Scheibner et al., 2017) and so on. In these paradigms, the attentional processes of external focus and internal focus were mostly generated through the instruction of experimenters or self-reports from participants (Wulf et al., 2002; Scheibner et al., 2017). Actually, the instruction and self-report were mostly subjective, and may not effectively ensure the participants engage in the externally focused/internally focused process, and thus, this uncertainty potentially confounds the previous findings. Our research group has proposed a task paradigm, i.e., continuous finger force feedback (Dong et al., 2012). In this paradigm, the attentional process of internal focus/external focus could be generated through the real/sham feedback condition. In the real feedback condition, visual feedback of the finger force severed as the external focus for participants to maintain the finger force, and in contrast, participants, in the sham feedback condition, should maintain the finger force according to their sensory feeling and memory. This allows generating the externally focused and internally focused processes through similar tasks, and behavioral data for each process could be acquired respectively. With this paradigm, it was observed that participants controlled the finger force at the same level no matter they engage in the attentional process of external focus/internal focus, but the internal focus induced greater behavioral variation than the external focus (Dong et al., 2012). Nevertheless, the brain mechanism underlying these behavioral findings remains to be understood.

Functional magnetic resonance imagining (fMRI) studies informed us with the brain activity underlying sustained attention. The prominent role of a brain network, i.e., default mode network (DMN) was intensively discussed in these studies (Raichle et al., 2001; Damoiseaux et al., 2008; Uddin et al., 2010). The DMN consistently shows higher activity at rest compared to tasks requiring sustained attention (Whitfield-gabrieli and Ford, 2012; Danckert and Merrifield, 2016). Higher activity of the DMN was suggested to be related to mind wandering (Andrews-Hanna, 2012; Mittner et al., 2016) and task-independent thoughts (Mason et al., 2007; Korostil et al., 2016). Increased activity of the DMN was linearly linked to intensity of awareness of internal focus (Vanhaudenhuyse et al., 2011). These evidences support the functional role of the DMN in internally directed cognition. In contrast, recent studies showed the DMN may also engage in the attentional process of external focus (Bogler et al., 2017; Scheibner et al., 2017). Scheibner et al. (2017) employed the instructions of experimenters and the self-reports from participants to generate the external focus (sound) and the internal focus (breathing), and they observed that the meditation practice reduced the activity of the DMN in the both tests of external focus and internal focus (Scheibner et al., 2017). Bogler et al. (2017) found the activity in the DMN was higher when subjects had a relatively short reaction time in a vigilance task (Bogler et al., 2017). However, no results in the report of Scheibner et al. (2017) showed the relationship between the DMN activity and the behavioral performance. Bogler et al. (2017) did not verify their findings in the attentional process of internal focus. Therefore, whether the DMN engaged in the externally focused process remains to be further validated. It is worthy to note that the DMN consists of several node regions including ventral and medial prefrontal cortex (vmPFC), the posterior cingulate cortex (PCC) and the left/right inferior parietal lobe/angular gyrus (LIPC/AG and RIPC/AG) (Raichle et al., 2001; Broyd et al., 2009; Greicius et al., 2009). In our previous studies, these regions showed higher activity in the real feedback condition than it in the sham feedback condition when we performed voxel-based analyses with the measurements of amplitude of low frequency fluctuation and regional homogeneity (Dong et al., 2012; Zhang et al., 2015a). Since no correlation between behavioral performance and regional activity were observed, these investigations offered few evidences for understanding the functional role of the DMN in sustained attention. Actually, regions constitute DMN through the functional connectivity which was methodologically defined as the correlation between the time course of a particular brain region and other regions (Friston et al., 1993). This functional connectivity could also be observed among different brain networks (Wang et al., 2014; Breakspear, 2017). These inter-regional and inter-network interactions were believed as the fundamental support for many cognitive processes, e.g., emotional modulation, skill learning, etc. (Mahiko et al., 2015; Nusslock et al., 2019). In our previous studies, we have explored the inter-network interactions, and the connectivity among the DMN, the executive network and the left frontal-parietal network exhibited changes between the attentional processes of internal attention and external focus (Zhang et al., 2015b). However behavioral performance did not show any correlation with the connectivity among these brain networks. Thus, the inter-regional connectivity was further assessed here to clarify the

functional role of DMN in the attentional process of external focus/internal focus.

The present study examined the connectivity among the node regions of DMN for verifying whether the DMN was also related to the attentional process of external focus. We hypothesized that if the DMN was related to the externally focused process, stronger connectivity could be identified in the process of external focus, and the connectivity among the node regions of the DMN should exhibit significant correlation with behavioral performance in this process. To test these hypotheses, fMRI data from our previous study were re-analyzed (Dong et al., 2012). The attentional processes of external focus and internal focus were established through the paradigm of continuous finger force feedback, and inter-regional connectivity within the DMN and their relations to behavior performances were assessed in each process separately.

MATERIALS AND METHODS

Participants

Forty-three right-handed college students participated in the study (23 ± 3 years, range 19–25; 23 females). No participant had history of brain injury, neurological illness or psychiatric disorders. Five subjects were excluded for the malfunction of experimental equipment (three subjects, leakage from the air tube resulted in the negative value of finger force) or excessive head motion (two subjects, head motion was >2 mm translation or >2° rotation in any direction), and at last, data from 38 subjects (mean age, 22 ± 2 years; 19 females) were involved in the further analysis. All experiments conducted in this study were approved by the Institutional Review Board of National Key Laboratory of Cognitive Neuroscience, Beijing Normal University. All the subjects gave written informed consent before the experiment.

Experimental Design

The current data were from our previous studies, and four papers have been published based on the data (Dong et al., 2012; Zhang et al., 2015a,b,c). Dong et al. (2012) proposed the finger force feedback paradigm and reported the behavioral data. Zhang et al. (2015a,c) were two methodological studies and provided the methodological framework for the exploration with this new paradigm. Zhang et al. (2015b) examined the functional connectivity among several brain networks, and no inter-regional functional connectivity within any one of the brain networks was analyzed (Zhang et al., 2015b). Each participant first underwent a scanning of resting state for adapting to the fMRI environment. Then, two sessions of external focus/internal focus were performed with the order counterbalanced across all participants. Each session lasts for 8 min, and the participants had a short practice period to get familiar with the related procedure before each session. In the session of external focus, the participants pinched a pressure sensor between the right index finger and thumb. This sensor is one module of an MRI-compatible physiological multi-channel analyzer (model MP150, BIOPAC Systems, Inc., Goleta, CA, United States). The sampling frequency was 250 Hz, and the pressure sensitivity was 0.01 cm H_2O. The pressure was recorded by a sensor via an airtight tube, and the force of pressure was synchronously fed back to the participant on a projector as the external focus. At the same time, each participant was requested to continuously regulate the finger force and try to maintain it at 20 cm H_2O according to the feedback. This target force was set in order to reduce the possibility of muscular fatigue for all subjects (van Duinen et al., 2007). In the session of internal focus, participants also maintained the finger force at 20 cm H_2O, and they should maintain the finger force according to their sensory memory and experience from the practice period but not from the feedback. Participants also watched a sham feedback to keep the visual inputs consistent across different sessions, and this feedback was generated with the behavioral data of another participant in external focus session. Because the sham feedback of finger force could be easily detected, we have informed participants of this fact in advance and requested them to keep their own performance unaffected.

Data Acquisition

Scanning was performed at the MRI Center of the Beijing Normal University using a 3.0-T Siemens whole-body MRI scanner. A single-shot T2*-weighted, gradient-echo, EPI sequence was used for functional imaging acquisition with the following parameters: repetition time (TR)/echo time (TE) = 2000 ms/30 ms, flip angle = 90°, acquisition matrix = 64 × 64; field of view (FOV) = 200 mm × 200 mm and slice thickness/gap = 3.5/0.7 mm. Thirty-three axial slices parallel to the AC-PC line were obtained in an interleaved order to cover the entire cerebrum and cerebellum. Then a T1-weighted three-dimensional magnetization-prepared rapid gradient echo (MPRAGE) sequence was acquired [128 sagittal slices, slice thickness/gap = 1.33/0 mm, in-plane resolution = 256 × 192, TR/TE/inversion time (TI) = 2530/3.39/1100 ms, flip angle = 7°, FOV = 256 mm × 256 mm].

Data Preprocessing

The preprocessing was carried out using the Data Processing Assistant for Resting-State fMRI (DPARSF) (Yan and Zang, 2010), which is based on the Statistical Parametric Mapping (SPM8)[1] and Resting State fMRI Data Analysis Toolkit (REST) (Song et al., 2011)[2]. The first 10 time points were removed for signal stabilization and participant adaptation, and then, the images were corrected for the difference in slice acquisition timing and head motion, coregistered to the T1 structural image. The head motion parameter measured by Friston-24 model and signals from white matter (WM) and cerebrospinal fluid (CSF) were further regressed out as nuisance covariates, and the linear trends were removed from the time courses of the voxels in each image. Then, images were spatially normalized into the standard Montreal Neurological Institute (MNI) template (re-sampled into 3 mm × 3 mm × 3 mm) and smoothed with an 8 × 8 × 8 full-width-at-half-maximum (FWHM) Gaussian kernel.

[1]http://www.fil.ion.ucl.ac.uk/spm

[2]http://www.restfmri.net

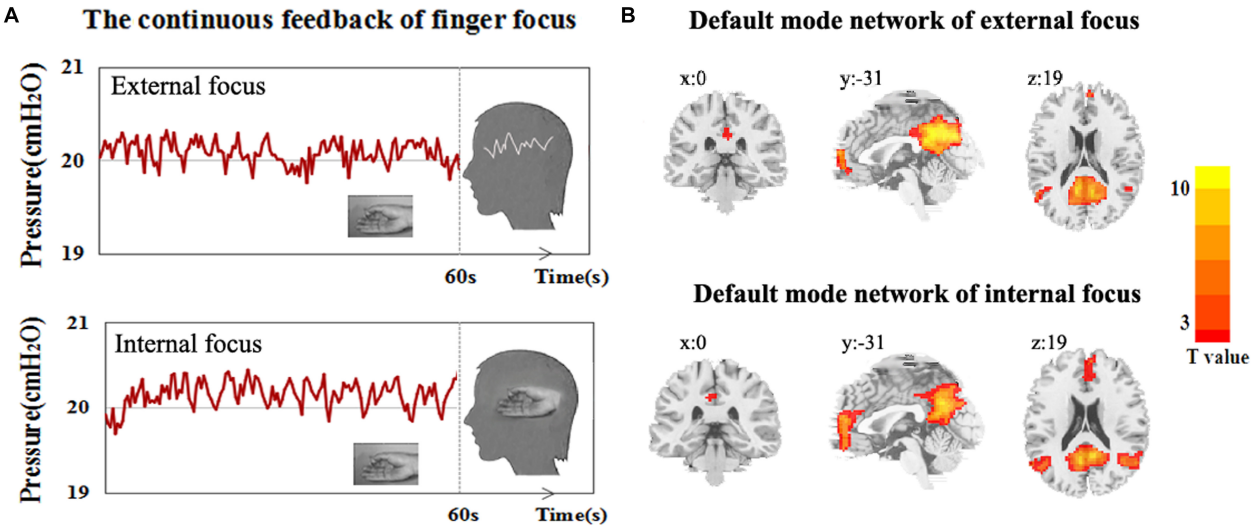

FIGURE 1 | The experiment procedure and DMN spatial map. **(A)** The experimental diagram of the attentional process of external focus/internal focus generated by continuous finger force feedback task, and **(B)** DMN spatial map of external focus/internal focus process identified with ICA.

TABLE 1 | Regions significantly recruited within DMN spatial map during the attentional process of external focus/internal focus.

Involved brain regions for DMN	L/R	BA	Peak MNI coordinates			
			x	*y*	*z*	t_{max}
The attentional process of external focus						
vmPFC	R	10	6	57	15	6.46
PCC	–	23	0	−51	33	8.28
LIPC/AG	L	19	−42	−72	36	7.32
RIPC/AG	R	39	45	−63	33	7.59
The attentional process of internal focus						
vmPFC	–	32	0	51	15	6.28
PCC	L	23	−6	−60	24	8.30
LIPC/AG	L	19	−39	−75	36	9.57
RIPC/AG	R	39	42	−63	30	7.55

The statistical threshold was set at $p < 0.001$, GRF corrected.

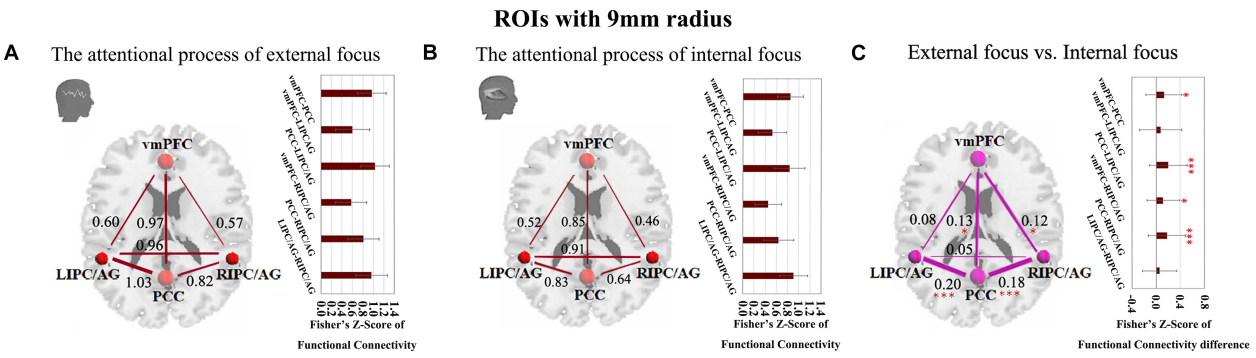

FIGURE 2 | The inter-regional connectivity between each pair of the regions within the DMN during the attentional process of external focus/internal focus. **(A)** The visualization of significant connectivity during the attentional process of external focus. **(B)** The visualization of significant connectivity during the attentional process of internal focus. In **(A,B)**, line width indicates the relative value of the connectivity. **(C)** The difference of the connectivity between the attentional processes of external focus and internal focus, and the line width indicates the relative value of the connectivity difference between the two processes. *Indicates the significant difference, $p < 0.05$; ***indicates the significant difference, $p < 0.001$, FDR corrected (Sphere radius = 9 mm).

TABLE 2 | The inter-regional connectivity within DMN of the attentional process of external focus/internal focus, and the comparison results of the inter-regional connectivity between the processes of external focus and internal focus.

Conditions	The attentional process of external focus	The attentional process of internal focus	External focus vs. internal focus	
DMN connectivity	***Fisher's Z-score, Mean ± SD***	***Fisher's Z-score, Mean ± SD***	***t(37)***	***p***
vmPFC-PCC	0.97 ± 0.27	0.85 ± 0.23	2.69	0.01*
vmPFC-LIPC/AG	0.60 ± 0.33	0.52 ± 0.26	1.41	0.17
PCC-LIPC/AG	1.03 ± 0.27	0.83 ± 0.27	3.97	0.0003***
vmPFC-RIPC/AG	0.57 ± 0.31	0.46 ± 0.24	2.68	0.01*
PCC-RIPC/AG	0.82 ± 0.30	0.64 ± 0.27	3.61	0.0009***
LIPC/AG-RIPC/AG	0.96 ± 0.30	0.91 ± 0.25	1.17	0.25

**Indicates the significant difference, $p < 0.05$; ***indicates the significant difference, $p < 0.001$, FDR corrected (Sphere radius = 9 mm).*

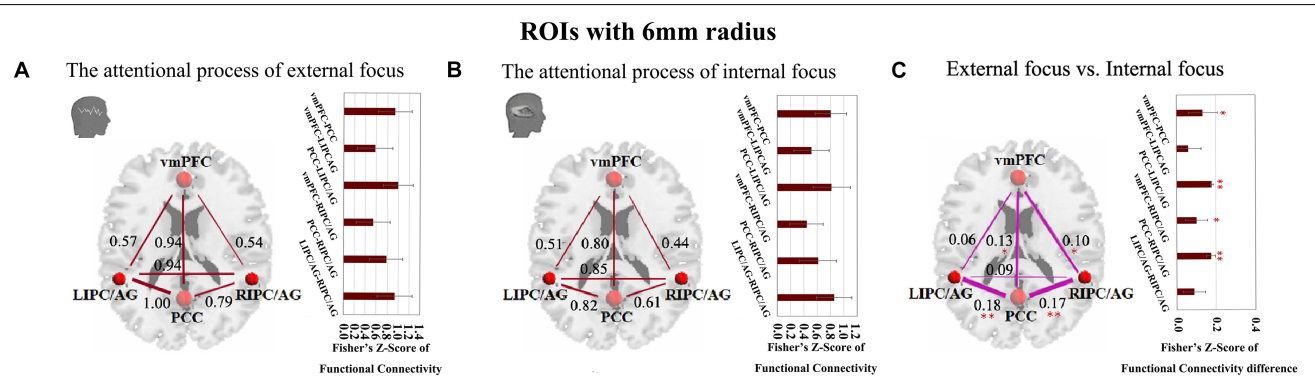

FIGURE 3 | The inter-regional connectivity between each pair of the regions within DMN during the attentional process of external focus/internal focus. **(A)** The visualization of significant connectivity during the attentional process of external focus. **(B)** The visualization of significant connectivity during the attentional process of internal focus. In **(A,B)**, line width indicates the relative value of the connectivity. **(C)** The difference of the connectivity between the attentional processes of external focus and internal focus, and the line width indicates the relative value of the connectivity difference between the two processes. *Indicates the significant difference, $p < 0.05$; **indicates the significant difference, $p < 0.005$, FDR corrected. (Sphere radius = 6 mm).

TABLE 3 | The inter-regional connectivity within the DMN of the externally and internally focused processes, and the comparison results of the inter-regional connectivity between the internally and externally focused processes.

Conditions	The attentional process of external focus	The attentional process of internal focus	External focus vs. Internal focus	
Connectivity	***Mean ± SD fisher's Z-score***	***Mean ± SD fisher's Z-score***	***t(37)***	***p***
vmPFC-PCC	0.94 ± 0.31	0.80 ± 0.23	2.55	0.02*
vmPFC-LIPC/AG	0.57 ± 0.33	0.51 ± 0.27	1.02	0.31
PCC-LIPC/AG	1.00 ± 0.27	0.82 ± 0.28	3.31	0.002**
vmPFC-RIPC/AG	0.54 ± 0.31	0.44 ± 0.25	2.41	0.02*
PCC-RIPC/AG	0.79 ± 0.30	0.61 ± 0.28	3.34	0.002**
LIPC/AG-RIPC/AG	0.94 ± 0.32	0.85 ± 0.26	1.84	0.07

**Indicates the significant difference, $p < 0.05$; **indicates the significant difference, $p < 0.005$, FDR corrected (Sphere radius = 6 mm).*

DMN Extraction With Independent Components Analysis

The preprocessed data from the external focus and internal focus sessions were combined into one single-group ICA analysis using the GIFT software[3], and the optimal component number in the analysis was estimated to be 20 according to the minimum description length (MDL) criteria (Calhoun et al., 2002). Two-step PCA was used to reduce the dimensionality of data to 20. Next, the data were decomposed by ICA using the informax algorithm (Bell and Sejnowski, 1995). To ensure the stability of the decomposition, ICASSO (Himberg et al., 2004) with 10 ICA runs were used (Ge et al., 2019), and the most stable run was selected as the final result. Then, spatially independent components (ICs) were back reconstructed for each subject, and at last, 20 ICs and the related time courses of responses for each subject were acquired. These ICs were converted to z-maps, and one-sample *t*-test was further performed to

[3]http://icatb.sourceforge.net

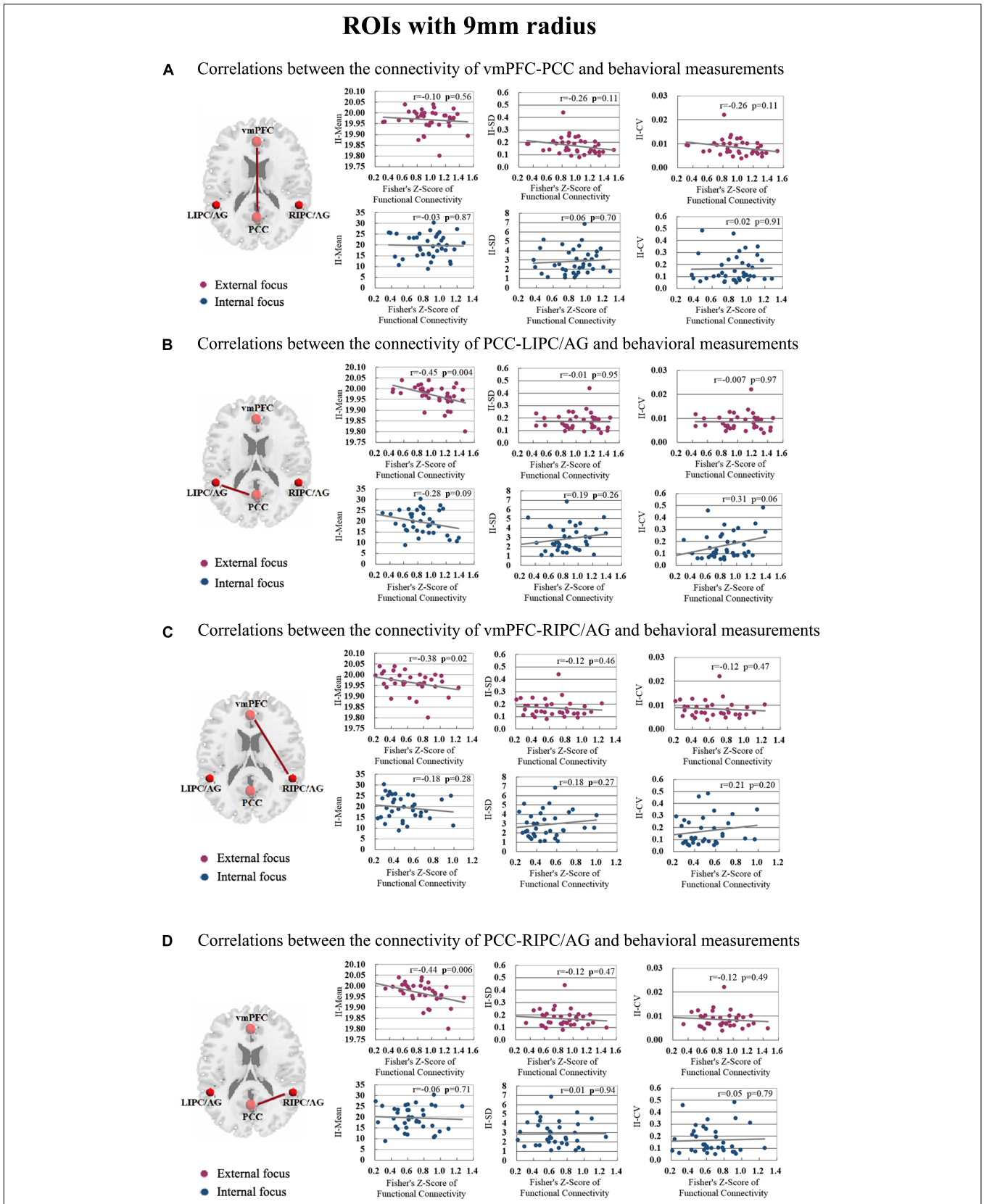

FIGURE 4 | The correlations between the behavioral measurements, i.e., II_Mean, II_SD and II_CV and the functional connectivity of **(A)** vmPFC-PCC, **(B)** PCC-LIPC/AG, **(C)** vmPFC-RIPC/AG, and **(D)** PCC-RIPC in the attentional process of external focus/internal focus. (Sphere radius = 9 mm).

TABLE 4 | The relationship between the inter-regional connectivity within DMN and behavioral measurements for the attentional process of external focus/internal focus respectively.

Behavior	The attentional process of external focus						The attentional process of internal focus					
Connectivity	II_Mean		II_SD		II_CV		II_Mean		II_SD		II_CV	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
DMN spatial map identified by ICA												
vmPFC-PCC	−0.10	0.56	−0.26	0.11	−0.26	0.11	−0.03	0.87	0.06	0.70	0.02	0.91
PCC-LIPC/AG	−0.45	0.004**	−0.01	0.95	−0.007	0.97	−0.28	0.09	0.19	0.26	0.31	0.06
vmPFC-RIPC/AG	−0.38	0.02*	−0.12	0.46	−0.12	0.47	−0.18	0.28	0.18	0.27	0.21	0.20
PCC-RIPC/AG	−0.44	0.006*	−0.12	0.47	−0.12	0.48	−0.06	0.71	0.01	0.94	0.05	0.79
DMN spatial template												
PCC-LIPC/AG	−0.50	0.001**	−0.05	0.77	−0.05	0.79	−0.1	0.55	0.03	0.86	0.08	0.61

**Indicates the significant difference, $p < 0.05$; **indicates the significant difference, $p < 0.005$ (Sphere radius = 9 mm).*

determine the group spatial map of DMN for subjects in the session of external focus/internal focus respectively ($p < 0.001$, GRF correction).

Functional Connectivity Among Regions of DMN

The functional connectivity was first analyzed based on the ICA results. Regions of the spatial map of DMN, including vmPFC, PCC, RIPC/AG and LIPC/AG were identified as the regions of interest (ROIs). A sphere with a 9-mm radius centered at the peak MNI coordinates of each ROI was defined as the seed region (external focus session, vmPFC: $x = 6$, $y = 57$, $z = 15$; PCC: $x = 0$, $y = -51$, $z = 33$; LIPC/AG: $x = -42$, $y = -72$, $z = 36$; RIPC/AG: $x = 45$, $y = -63$, $z = 33$; internal focus session, vmPFC: $x = 0$, $y = 51$, $z = 15$; PCC: $x = -6$, $y = -60$, $z = 24$; LIPC/AG: $x = -39$, $y = -75$, $z = 36$; RIPC/AG: $x = 42$, $y = -63$, $z = 30$). Then, the preprocessed image data were filtered to 0.01–0.08 Hz, and the mean time course of each seed region was extracted. Functional connectivity between each pair of two seed regions was calculated through Pearson correlation coefficient. The Fisher *z*-transformed correlation coefficients identified as the DMN connectivity were compared between the two sessions using paired *t*-test, and all of the tested results underwent the multiple comparison correction [false discovery rate (FDR) correction $q < 0.05$].

These analyses were validated by using the ROIs with a different radius (6 mm). According to the above statistical results, the DMN connectivity showing significant difference between the sessions of external focus and internal focus was identified, and the identified functional connectivity was involved in the further analysis.

Correlation Between Functional Connectivity and Behavior

The behavioral data of the external focus and internal focus sessions have been analyzed and illustrated in our previous investigation (Dong et al., 2012). Three measurements of intra-individual behavior were calculated, and the measurements include intra-individual mean finger force (II_Mean, the mean value of finger force across a whole session), intra-individual standard deviation (II_SD, the SD of the individual pinch force across a whole session) and intra-individual variation coefficient (II_CV calculated as SD/mean value of the individual pinch force across a whole session). Then, the correlation between the identified functional connectivity and each of the behavioral measurements was calculated in the external focus/internal focus session respectively.

All the results were further validated by reproducing the functional connectivity analysis based on a public DMN spatial template (Greicius et al., 2004). vmPFC, PCC, RIPC/AG, and LIPC/AG in the template were defined as the seed region. Then, the mean time course of each seed region was extracted based on the filtered image data, and functional connectivity between each pair of two seed regions was calculated with Pearson correlation coefficient. The Fisher *z*-transformed correlation coefficients were further compared between the attentional processes of external focus and internal focus using paired *t*-test, and all of the tested results were further corrected for multiple comparison (FDR correction, $q < 0.05$).

RESULTS

The DMN for the Attentional Processes of Internal Focus and External Focus

The attentional process of external focus/internal focus was generated with real/sham feedback condition (**Figure 1A**). The DMN spatial maps, identified in the attentional processes of external focus and internal focus were shown in **Figure 1B**. In both processes, the DMN spatial map involves regions of ventral and medial prefrontal cortex (vmPFC), posterior cingulated cortex (PCC)/precuneus and left and right inferior parietal cortex/angular gyrus (LIPC/AG and RIPC/AG) and the peak MNI coordinates of each region was showed in **Table 1**.

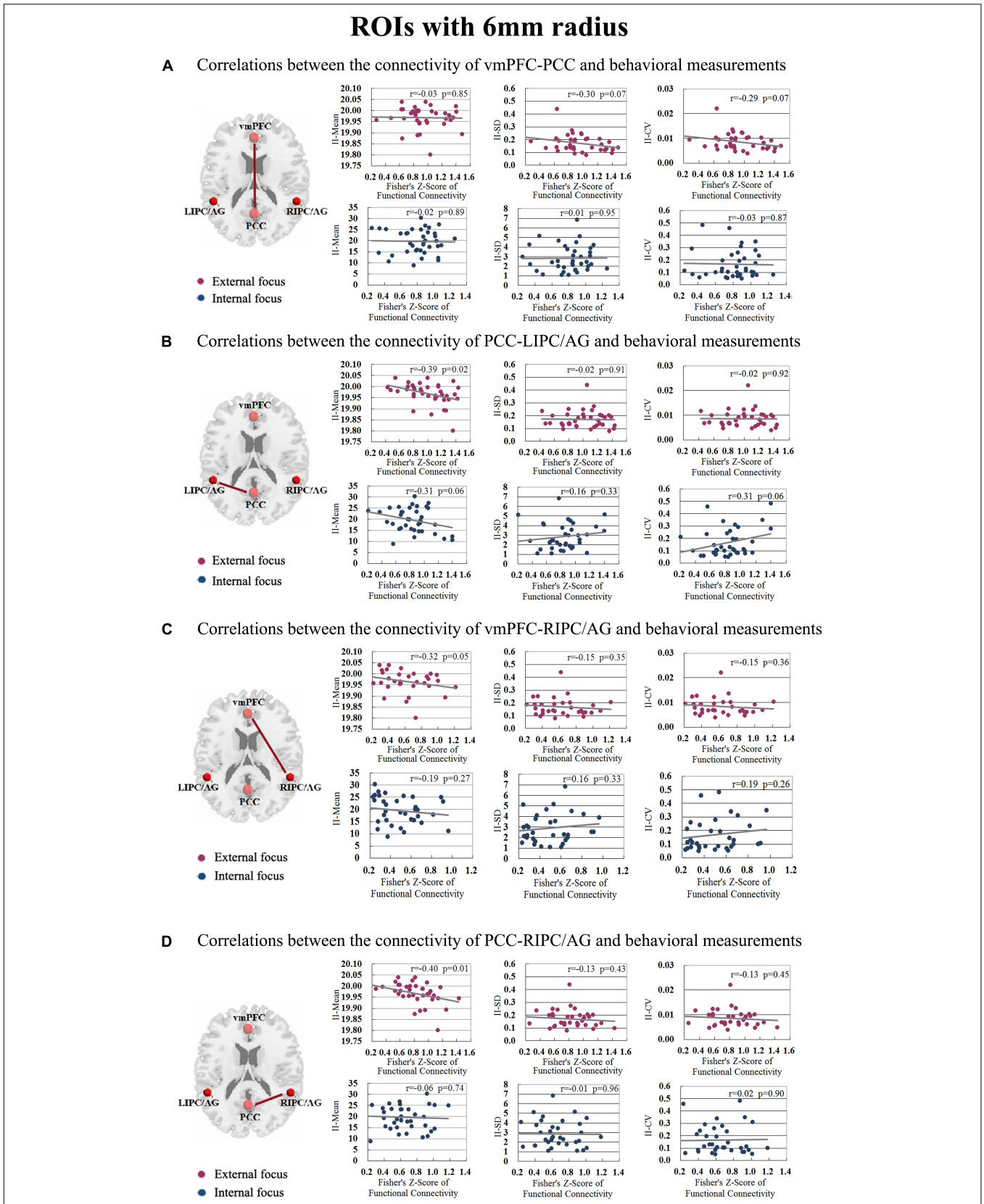

FIGURE 5 | The correlations between the behavioral measurements, i.e., II_Mean, II_SD and II_CV and the functional connectivity of **(A)** vmPFC-PCC, **(B)** PCC-LIPC/AG, **(C)** vmPFC-RIPC/AG, and **(D)** PCC-RIPC in the attentional process of external focus/internal focus (Sphere radius = 6 mm).

TABLE 5 | The relationship between the inter-regional connectivity within DMN and behavioral measurements for the attentional process of external focus/internal focus respectively.

Behavior	The attentional process of external focus						The attentional process of internal focus					
Connectivity	II_Mean		II_SD		II_CV		II_Mean		II_SD		II_CV	
	r	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*	*r*	*p*
DMN spatial map identified by ICA												
vmPFC-PCC	−0.03	0.85	−0.30	0.07	−0.29	0.07	−0.02	0.89	0.01	0.95	−0.03	0.87
PCC-LIPC/AG	−0.39	0.02*	−0.02	0.91	−0.02	0.92	−0.31	0.06	0.16	0.33	0.31	0.06
vmPFC-RIPC/AG	−0.32	0.05*	−0.15	0.35	−0.15	0.36	−0.19	0.27	0.16	0.33	0.19	0.26
PCC-RIPC/AG	−0.40	0.01*	−0.13	0.43	−0.13	0.45	−0.06	0.74	−0.009	0.96	0.02	0.90

**Indicates the significant difference, $p < 0.05$ (Sphere radius = 6 mm).*

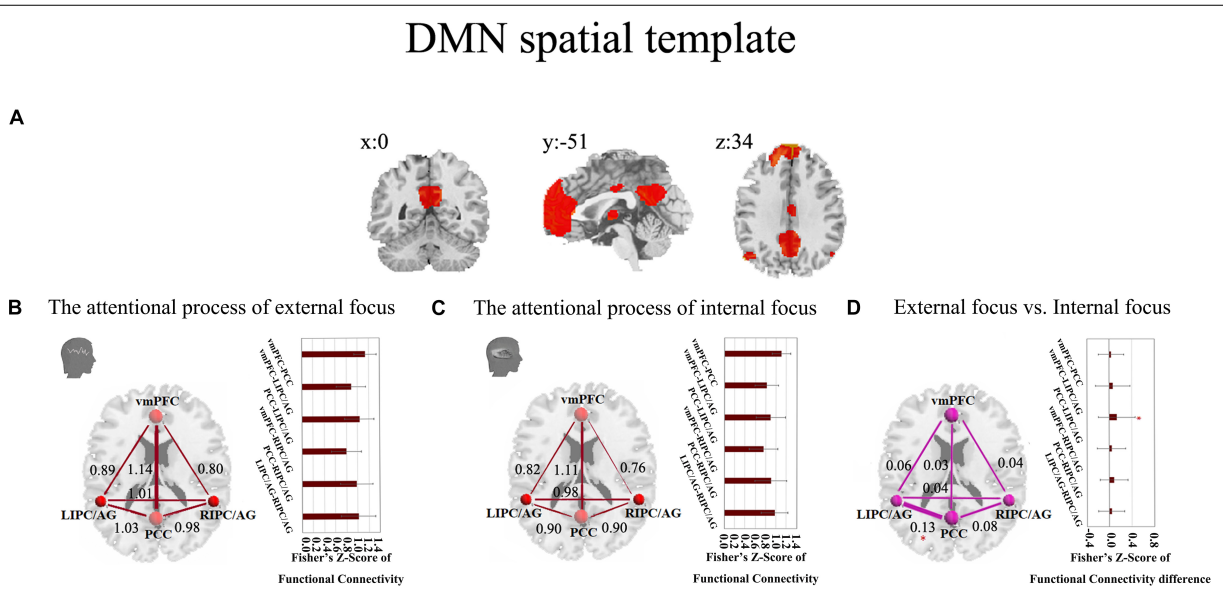

FIGURE 6 | The inter-regional connectivity between each pair of the regions within DMN during the attentional processes of external/internal focus. **(A)** The public DMN spatial template. **(B)** The visualization of significant connectivity during the attentional process of external focus. **(C)** The visualization of significant connectivity during the attentional process of internal focus. In **(B,C)**, line width indicates the relative value of the connectivity. **(D)** The difference of the connectivity between the attentional processes of external focus and internal focus and the line width indicates the relative value of the connectivity difference between the two processes. *Indicates the significant difference, $p < 0.05$, FDR corrected.

Inter-Regional Connectivity Within the DMN

The four critical regions of the DMN showed six pairs of inter-regional connectivity with each other (**Figures 2A,B**), and the inter-regional connectivity within the DMN was significant in both external focus and internal focus processes (each $t > 11.40$, $p < 0.0001$). Four pairs of the connectivity exhibited significant differences between the attentional processes of external focus and internal focus, including vmPFC-PCC, PCC-LIPC/AG, vmPFC-RIPC/AG and PCC-RIPC/AG (see details in **Figure 2C** and **Table 2**).

Since we have performed four studies based on the same data, thus, we further performed the multiple comparisons (taking into account previous correlations that we have done). Twenty comparisons between the two different attentional processes have been administrated. Using Bonferroni correction across all comparisons, the significant level for the comparisons in the current study is $p < 0.0025$ (0.05/20). Thus, the stronger connectivity of PCC-LIPC/AG and PCC-RIPC/AG for the externally focused process could withstand this multiple comparison correction (each $p < 0.001$).

Moreover, we further validated these results using the ROIs with a different radius (6 mm). The significant differences between the attentional processes of external focus and internal focus were identified in the functional connectivity of vmPFC-PCC, PCC-LIPC/AG, vmPFC-RIPC/AG and PCC-RIPC/AG (**Figure 3** and **Table 3**).

Correlation Between Inter-Regional Connectivity Within the DMN and Behavioral Measurements

Behavioral data from the external focus and internal focus sessions have been analyzed and illustrated in our previous

TABLE 6 | The inter-regional connectivity within DMN and the difference of the functional connectivity between the attentional processes of external focus and internal focus.

Conditions	The attentional process of external focus	The attentional process of internal focus	External focus vs. internal focus	
DMN Connectivity	***Fisher's Z-score, Mean ± SD***	***Fisher's Z-score, Mean ± SD***	***t(37)***	***p***
vmPFC-PCC	1.14 ± 0.20	1.11 ± 0.17	0.92	0.37
vmPFC-LIPC/AG	0.89 ± 0.26	0.82 ± 0.23	1.26	0.22
PCC-LIPC/AG	1.03 ± 0.26	0.90 ± 0.29	2.52	0.01*
vmPFC-RIPC/AG	0.80 ± 0.27	0.76 ± 0.28	0.95	0.35
PCC-RIPC/AG	0.98 ± 0.29	0.90 ± 0.31	2.01	0.05
LIPC/AG-RIPC/AG	1.01 ± 0.31	0.98 ± 0.26	0.97	0.34

**Indicates the significant difference, $p < 0.05$, FDR corrected.*

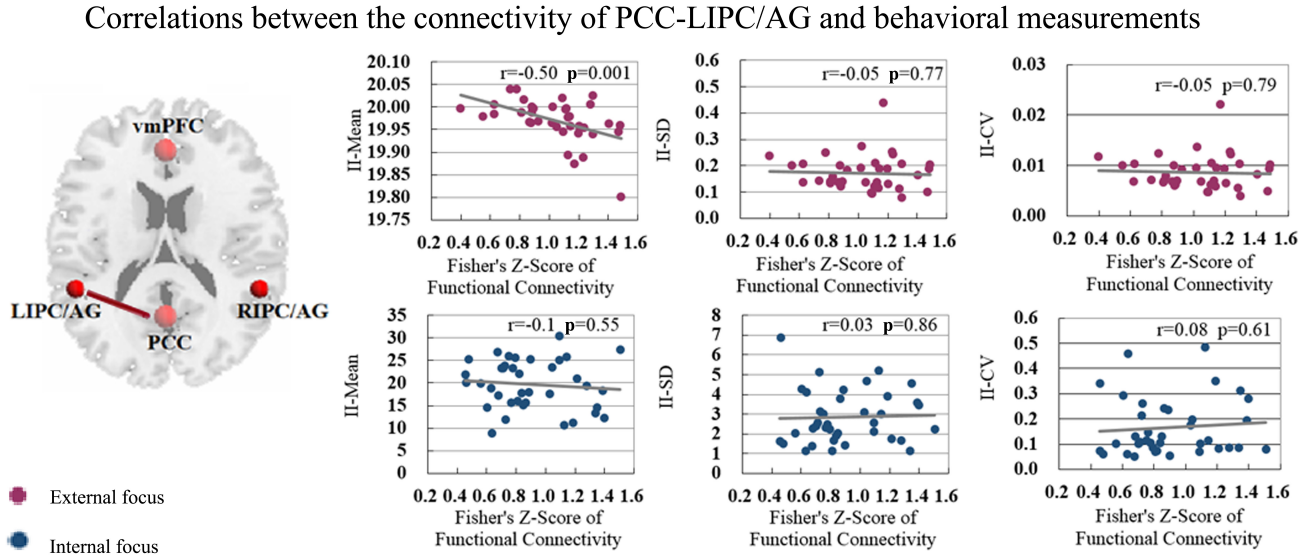

FIGURE 7 | The correlations between the behavioral measurements i.e., II_Mean, II_SD and II_CV and the functional connectivity of PCC-LIPC/AG in the attentional process of external focus/internal focus.

investigation (Dong et al., 2012). Paired *t*-tests showed that there was no significant difference in the II_Mean between the attentional processes of external focus and internal focus. II_Mean across all subjects was 19.97 ± 0.05 cm H_2O for the external focus process and 19.72 ± 5.40 cm H_2O for the internal focus process ($t = 0.29$, $p > 0.05$). However, the II_SD and II_CV in the internal focus process were markedly higher than it in the external focus process. II_SD across all subjects was 0.17 ± 0.07 cm H_2O for the external focus process and 2.85 ± 1.34 cm H_2O for the internal focus process ($t = 12.46$, $p < 0.0001$). II_CV across all subjects was 0.01 ± 0.003 cm H_2O for the real feedback and 0.17 ± 0.11 cm H_2O for the sham feedback ($t = 8.64$, $p < 0.0001$).

The correlation between significant functional connectivity within DMN spatial map and behavioral measurement was showed in **Figure 4**. There was a significant correlation between the functional connectivity of PCC-LIPC/AG and the II-Mean of finger force in external focus process not internal focus process, and there was no significant difference between the functional connectivity of PCC-LIPC/AG and the II_SD or II_CV of finger force (**Table 4**). Moreover, these results were validated using the ROIs with a different radius (6 mm) (see details in **Figure 5** and **Table 5**).

Since we have performed four studies based on the same data, thus, we further performed the multiple comparisons (taking into account previous correlations that we have done). Hundred and one correlations between the fMRI data and behavioral data have been assessed. Using Bonferroni correction across all correlations, no correlation results could withstand this multiple comparison correction (Bonferroni correction across all correlations, $p < 0.0005$).

These results were further validated with the DMN spatial template (**Figure 6A**). The six pairs of functional connectivity for the attentional processes of external focus and internal focus were show in **Figures 6B,C**, and only the functional connectivity of PCC-LIPC/AG reserved the significant differences between the attentional processes of external focus and internal focus (see details in **Figure 6D** and **Table 6**).

The functional connectivity of PCC-LIPC/AG also showed significant correlation with the behavioral measurement of II-Mean in the attentional process of external focus (**Figure 7** and **Table 4**).

DISCUSSION

The present study explored the functional significance of the DMN in sustained attention. The two processes were generated through a new paradigm, i.e., continuous feedback of finger force, and the functional connectivity among the node regions of the DMN was assessed in each process respectively. Two intriguing results were obtained: (1) the functional connectivity of PCC-LIPC/AG was significant stronger in the externally focused process than it in internally focused process, and this difference was reproduced in the validation analyses with different ROI radius and the public DMN spatial template; (2) the functional connectivity of PCC-LIPC was correlated with the behavioral measurement, II_Mean in the externally focused process. These findings potentially offer new insights into the functional significance of the DMN in the attentional processes of internal focus and external focus.

Since Chun et al. (2011) dissociated the sustained attention into the two attentional processes of internal focus and external focus (Chun et al., 2011), many studies attempt to identify the difference of behavioral performance and brain mechanism between the two attentional processes (Poolton et al., 2006; Stawarczyk et al., 2011; Shafizadeh et al., 2013; Schücker et al., 2014). Instructions and self-reports were the strategies mostly employed in these investigations to generate the attentional process of external focus/internal focus (Wulf et al., 2002; Ruocco and Direkoglu, 2013; Scheibner et al., 2017). However, these strategies could not be assessed by the objective behavioral measurements, thus it is difficulty to ensure subjects engage in these required processes. The paradigm of continuous feedback of finger force was employed in the present study, and controlling finger force by feedback and controlling finger force by the sensory memory were used to generate the external focus/internal focus. The finger force as the behavioral performance was recorded, and this paradigm potentially provide objective assessments for future studies on the attentional process of internal focus/external focus.

As hypothesized, we observed stronger functional connectivity within DMN in the attentional process of external focus. Previous studies suggested the involvement of the DMN in the internally directed cognition, e.g., mind wandering, task independent thoughts etc. (Jang et al., 2011; Benedek et al., 2016; Mittner et al., 2016). Here, the external focus increased the inter-regional connectivity of PCC-LIPC/AG as compared with the internal focus, and this result was still reserved in the validation analysis. Both PCC and LIPC were suggested to be related to working memory and information integrating (Ye and Zhou, 2009; Huang et al., 2018). In the current study, the functional connectivity of PCC-LIPC showed significant correlation with the behavioral measurement, II_Mean in the externally focused process. This attentional process requires processing and integrating the feedback information. Thus, PCC-LIPC within DMN may directly engage in the process of external focus, and probably regulated the behavioral performance related to the external focus. The dominant proposal argues that the DMN is mainly responsible for mind wandering, task independent thoughts (Andrews-Hanna, 2012; Seli et al., 2016; Scheibner et al., 2017; Bocharov et al., 2018). Thus, the functional role of DMN in sustained attention may be more than responding to the internally directed cognition.

Several limitations exist in current study. First, we observed that the stronger inter-regional connectivity of the DMN exhibited correlations with the behavioral measurement, II_Mean of the finger force in the externally focused process. However, these correlation results could not withstand the multiple comparison correction (each $p < 0.01$), if we took our previous studies based on the same data into account (Bonferroni correction across all correlations of the previous studies, $p < 0.0005$), thus the relationship between the inter-regional correlation and the behavioral performance in the process of external focus should be further verified. Second, The behavioral measurements, II_SD and II_CV were more meaningful for the assessment of the fluctuation of sustained attention (Liu et al., 2017). Whether the fluctuation of sustained attention was associated with the DMN remains to be understood. Third, for the internal focus, the relationship between the brain activity and behavioral performance in the internally focused process requires to be established in future studies, and we believed this is a critical issue for understanding of the brain mechanism underlying sustained attention.

CONCLUSION

The present study explored the functional significance of the DMN in the attentional processes of external focus and internal focus. The external focus could increase the inter-regional connectivity, PCC-LIPC/AG of the DMN, and this connectivity within the DMN was possible the reason of regulating the behavioral performance in the externally focused process; These findings offered new evidences to support the engagement of the DMN in the attentional process of external focus. Thus, the functional significance of the DMN was more than the internally directed cognition, and the continuous feedback of finger force, as an objective assessing paradigm for sustained attention with external focus and internal focus deserves more concerns in future studies.

DATA AVAILABILITY STATEMENT

The datasets analyzed in this manuscript are not publicly available. Requests to access the datasets should be directed to kevinhangbnu@foxmail.com.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Institutional Review Board of National Key Laboratory of Cognitive Neuroscience, Beijing Normal University. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

HZ and HL conceived and designed the experiment. Z-WZ and X-QL performed the data analysis. HL, Y-FZ, Y-TF, and YG provided advice on the analysis and interpretation of the results. Z-WZ, X-QL, and HZ wrote the manuscript.

FUNDING

This work was supported by the National Key R&D Program of China (2018YFC1312600 and 2018YFC1312603), National Natural Science Foundation of China (31471084 and 81520108016), and Zhejiang Fundamental Public Welfare Research Program (LGJ19C090001). HZ was supported by the Scientific Research Staring Foundation from Hangzhou Normal University. Z-WZ was supported by the project of provincial advantage discipline construction, talent cultivation, and research innovation from College of Education, Hangzhou Normal University.

REFERENCES

Andrews-Hanna, J. R. (2012). The brain's default network and its adaptive role in internal mentation. *Neuroscientist* 18, 251–270. doi: 10.1177/1073858411403316

Bell, A. J., and Sejnowski, T. J. (1995). An information-maximization approach to blind separation and blind deconvolution. *Neural Comput.* 7, 1129–1159. doi: 10.1162/neco.1995.7.6.1129

Benedek, M., Jauk, E., Beaty, R. E., Fink, A., Koschutnig, K., and Neubauer, A. C. (2016). Brain mechanisms associated with internally directed attention and self-generated thought. *Sci. Rep.* 6:22959. doi: 10.1038/srep22959

Bocharov, A. V., Knyazev, G. G., Savostyanov, A. N., Astakhova, T. N., and Tamozhnikov, S. S. (2018). EEG dynamics of spontaneous stimulus-independent thoughts. *Cogn. Neurosci.* 10, 77–87. doi: 10.1080/17588928.2018.1534820

Bogler, C., Vowinkel, A., Zhutovsky, P., and Haynes, J. D. (2017). Default network activity is associated with better performance in a vigilance task. *Front. Hum. Neurosci* 11:623. doi: 10.3389/fnhum.2017.00623

Breakspear, M. (2017). Dynamic models of large-scale brain activity. *Nat. Neurosci.* 20, 340–352. doi: 10.1038/nn.4497

Broyd, S. J., Demanuele, C., Debener, S., Helps, S. K., James, C. J., and Sonuga-Barke, E. J. S. (2009). Default-mode brain dysfunction in mental disorders: a systematic review. *Neurosci. Biobehav. Rev.* 33, 279–296. doi: 10.1016/j.neubiorev.2008.09.002

Burton, B. K., Vangkilde, S., Petersen, A., Skovgaard, L. T., Jepsen, J. R., Hemager, N., et al. (2018). Sustained attention and interference control among 7-year-old children with a familial high risk of schizophrenia or bipolar disorder-a nationwide observational cohort study. *Biol. Psychiatr.* 3, 704–712. doi: 10.1016/j.bpsc.2018.04.012

Calhoun, V. D., Adali, T., Pearlson, G. D., and Pekar, J. J. (2002). A method for making group inferences from functional MRI data using independent component analysis. *Hum. Brain Mapp.* 16, 131–131. doi: 10.1002/hbm.10044

Chun, M. M., Golomb, J. D., and Turk-Browne, N. B. (2011). A taxonomy of external and internal attention. *Annu. Rev. Psychol.* 62, 73–101. doi: 10.1146/annurev.psych.093008.100427

Coull, J. T., Frith, C. D., Frackowiak, R. S. J., and Grasby, P. M. (1996). A fronto-parietal network for rapid visual information processing: a PET study of sustained attention and working memory. *Neuropsychologia* 34, 1085–1095. doi: 10.1016/0028-3932(96)00029-2

Damoiseaux, J. S., Beckmann, C. F., Arigita, E. J. S., Barkhof, F., Scheltens, P., Stam, C. J., et al. (2008). Reduced resting-state brain activity in the "default network" in normal aging. *Cereb. Cortex* 18, 1856–1864. doi: 10.1093/cercor/bhm207

Danckert, J., and Merrifield, C. (2016). Boredom, sustained attention and the default mode network. *Exp. Brain Res.* 236, 2507–2518. doi: 10.1007/s00221-016-4617-5

Dong, Z. Y., Liu, D. Q., Wang, J., Qing, Z., Zang, Z. X., Yan, C. G., et al. (2012). Low-frequency fluctuation in continuous real-time feedback of finger force: a new paradigm for sustained attention. *Neurosci. Bull.* 28, 456–467. doi: 10.1007/s12264-012-1254-2

Fortenbaugh, F. C., Degutis, J., and Esterman, M. (2017). Recent theoretical, neural, and clinical advances in sustained attention research. *Ann. N. Y. Acad. Sci.* 1396, 70–91. doi: 10.1111/nyas.13318

Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S. J. (1993). Functional connectivity: the principal-component analysis of large (PET) data sets. *J. Cereb. Blood Flow Metab.* 13, 5–14. doi: 10.1038/jcbfm.1993.4

Ge, R., Downar, J., Blumberger, D. M., Daskalakis, Z. J., Lam, R. W., and Vila-Rodriguez, F. (2019). Structural network integrity of the central executive network is associated to the therapeutic effect of rTMS in treatment resistant depression. *Prog. Neuropsychopharmacol. Biol. Psychiatr.* 92, 217–225. doi: 10.1016/j.pnpbp.2019.01.012

Greicius, M. D., Srivastava, G., Reiss, A. L., and Menon, V. (2004). Default-mode network activity distinguishes Alzheimer's disease from healthy aging: evidence from functional MRI. *Proc. Natl. Acad. Sci. U.S.A.* 101, 4637–4642. doi: 10.1073/pnas.0308627101

Greicius, M. D., Supekar, K., Menon, V., and Dougherty, R. F. (2009). Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb. Cortex* 19, 72–78. doi: 10.1093/cercor/bhn059

Himberg, J., Hyvärinen, A., and Esposito, F. (2004). Validating the independent components of neuroimaging time series via clustering and visualization. *Neuroimage* 22, 1214–1222. doi: 10.1016/j.neuroimage.2004.03.027

Huang, Y., Hullfish, J., Ridder, D. D., and Vanneste, S. (2018). Meta-analysis of functional subdivisions within human posteromedial cortex. *Brain Struct. Funct.* 224, 435–452. doi: 10.1007/s00429-018-1781-3

Jang, J. H., Jung, W. H., Kang, D. H., Byun, M. S., Kwon, S. J., Choi, C. H., et al. (2011). Increased default mode network connectivity associated with meditation. *Neurosci. Lett.* 487, 358–362. doi: 10.1016/j.neulet.2010.10.056

Korostil, M., Fatima, Z., Kovacevic, N., Menon, M., and McIntosh, A. R. (2016). Task-independent effects are potential confounders in longitudinal imaging studies of learning in schizophrenia. *Neuroimage* 10, 159–171. doi: 10.1016/j.nicl.2015.11.016

Liu, L., Cheng, J., Su, Y., Ji, N., Gao, Q., Li, H. M., et al. (2015). Deficiency of sustained attention in ADHD and its potential genetic contributor MAOA. *J. Atten. Disord.* 22, 878–885. doi: 10.1177/1087054715574832

Liu, Z. Z., Qu, H. J., Tian, Z. L., Han, M. J., Fan, Y., Ge, L. Z., et al. (2017). Reproducibility of frequency-dependent low frequency fluctuations in reaction time over time and across tasks. *PLoS One* 12:e0184476. doi: 10.1371/journal.pone.0184476

Mahiko, K., George, M. L. D., Haakon, E., Jonathan, S., and Satoru, H. (2015). Shaped by the past: the default mode network supports cognition that is independent of immediate perceptual input. *PLoS One* 10:e0132209. doi: 10.1371/journal.pone.0132209

Mansell, W., Clark, D. M., and Ehlers, A. (2003). Internal versus external attention in social anxiety: an investigation using a novel paradigm. *Behav. Res. Ther.* 41, 555–572. doi: 10.1016/s0005-7967(02)00029-3

Mason, M. F., Norton, M. I., Van Horn, J. D., Wegner, D. M., Grafton, S. T., and Macrae, C. N. (2007). Wandering minds: the default network and stimulus-independent thought. *Science* 315, 393–395. doi: 10.1126/science.1131295

Mittner, M., Hawkins, G. E., Boekel, W., and Forstmann, B. U. (2016). A neural model of mind wandering. *Trends Cogn. Sci.* 20, 570–578. doi: 10.1016/j.tics.2016.06.004

Nusslock, R., Brody, G. H., Armstrong, C. C., Carroll, A. L., Sweet, L. H., Yu, T., et al. (2019). Higher peripheral inflammatory signaling associated with lower resting-state functional brain connectivity in emotion regulation and central executive networks. *Biol. Psychiatry* 86, 153–162. doi: 10.1016/j.biopsych.2019.03.968

Poolton, J. M., Maxwell, J. P., Masters, R. S. W., and Raab, M. (2006). Benefits of an external focus of attention: common coding or conscious processing? *J. Sports Sci.* 24, 89–99. doi: 10.1080/02640410500130854

Raichle, M. E., Macleod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682.

Ruocco, A. C., and Direkoglu, E. (2013). Delineating the contributions of sustained attention and working memory to individual differences in mindfulness. *Pers. Individ. Dif.* 54, 226–230. doi: 10.1016/j.paid.2012.08.037

Scheibner, H. J., Bogler, C., Gleich, T., Haynes, J. D., and Bermpohl, F. (2017). Internal and external attention and the default mode network. *Neuroimage* 148, 381–389. doi: 10.1016/j.neuroimage.2017.01.044

Schücker, L., Knopf, C., Strauss, B., and Hagemann, N. (2014). An internal focus of attention is not always as bad as its reputation: how specific aspects of internally focused attention do not hinder running efficiency. *J. Sport Exerc. Psychol.* 36, 233–243. doi: 10.1123/jsep.2013-0200

Seli, P., Risko, E. F., Smilek, D., and Schacter, D. L. (2016). Mind-wandering with and without intention. *Trends Cogn. Sci.* 20, 605–617. doi: 10.1016/j.tics.2016.05.010

Shafizadeh, M., Platt, G. K., and Bahram, A. (2013). Effects of focus of attention and type of practice on learning and self-efficacy in dart throwing. *Percept. Mot. Skills* 117, 1224–1234.

Song, X. W., Dong, Z. Y., Long, X. Y., Li, S. F., Zuo, X. N., Zhu, C. Z., et al. (2011). Rest: a toolkit for resting-state functional magnetic resonance imaging data processing. *PLoS One* 6:e25031. doi: 10.1371/journal.pone.0025031

Stawarczyk, D., Majerus, S., Maquet, P., and D'Argembeau, A. (2011). Neural correlates of ongoing conscious experience: both task-unrelatedness and stimulus-independence are related to default network activity. *PLoS One* 6:e16997. doi: 10.1371/journal.pone.0016997

Uddin, L. Q., Kelly, A. M., Biswal, B. B., Castellanos, F. X., and Milham, M. P. (2010). Functional connectivity of default mode network components: correlation, anticorrelation, and causality. *Hum. Brain Mapp.* 30, 625–637. doi: 10.1002/hbm.20531

van Duinen, H., Renken, R., Maurits, N., and Zijdewind, I. (2007). Effects of motor fatigue on human brain activity, an fMRI study. *Neuroimage* 35, 1438–1449. doi: 10.1016/j.neuroimage.2007.02.008

Vanhaudenhuyse, A., Demertzi, A., Schabus, M., Noirhomme, Q., Bredart, S., Boly, M., et al. (2011). Two distinct neuronal networks mediate the awareness of environment and of self. *J. Cogn. Neurosci.* 23, 570–578. doi: 10.1162/jocn.2010.21488

Wang, Z., Dai, Z., Gong, G., Zhou, C., and He, Y. (2014). Understanding structural-functional relationships in the human brain. *Neuroscientist* 21, 290–305. doi: 10.1177/1073858414537560

Whitfield-gabrieli, S., and Ford, J. M. (2012). Default mode network activity and connectivity in psychopathology. *Ann. Rev. Clin. Psychol.* 8, 49–76. doi: 10.1146/annurev-clinpsy-032511-143049

Wulf, G., McConnel, N., Gärtner, M., and Schwarz, A. (2002). Enhancing the learning of sport skills through external-focus feedback. *J. Mot. Behav.* 34, 171–182. doi: 10.1080/00222890209601939

Yan, C. G., and Zang, Y. F. (2010). DPARSF: a MATLAB toolbox for "pipeline" data analysis of resting-state fMRI. *Front. Syst. Neurosci.* 4:13.

Ye, G., Yin, G. Z., Tang, Z., Fu, J. L., Chen, J., Chen, S. S., et al. (2018). Association between increased serum interleukin-6 levels and sustained attention deficits in patients with major depressive disorder. *Psychol. Med.* 48, 2508–2514. doi: 10.1017/S0033291718000090

Ye, Z., and Zhou, X. (2009). Conflict control during sentence comprehension: fMRI evidence. *Neuroimage* 48, 280–290. doi: 10.1016/j.neuroimage.2009.06.032

Zhang, H., Gao, Z. Z., and Zang, Y. F. (2015a). An fMRI study of local synchronization in different subfrequency bands during the continuous feedback of finger force. *BioMed. Res. Int.* 2015, 1–8. doi: 10.1155/2015/273126

Zhang, H., Zhang, H., and Zang, Y. F. (2015b). Functional connectivity among brain networks in continuous feedback of finger force. *Neuroscience* 289, 134–143. doi: 10.1016/j.neuroscience.2014.12.075

Zhang, H., Zhang, L. J., and Zang, Y. F. (2015c). Fluctuation amplitude and local synchronization of brain activity in the ultra-low frequency band: an fMRI investigation of continuous feedback of finger force. *Brain Res.* 1629, 104–112. doi: 10.1016/j.brainres.2015.10.023

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.

Copyright © 2019 Zhou, Lan, Fang, Gong, Zang, Luo and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 15 October 2019
doi: 10.3389/fnins.2019.01088

Musical Role Asymmetries in Piano Duet Performance Influence Alpha-Band Neural Oscillation and Behavioral Synchronization

Auriel Washburn[1,2], Irán Román[1], Madeline Huberth[1], Nick Gang[1], Tysen Dauer[1], Wisam Reid[1†], Chryssie Nanou[1], Matthew Wright[1] and Takako Fujioka[1,3]*

[1] Center for Computer Research in Music and Acoustics, Department of Music, Stanford University, Stanford, CA, United States, [2] Department of Computer Science and Engineering, University of California, San Diego, San Diego, CA, United States, [3] Wu Tsai Neurosciences Institute, Stanford University, Stanford, CA, United States

OPEN ACCESS

Edited by:
Daya Shankar Gupta, Camden County College, United States

Reviewed by:
Caroline Palmer, McGill University, Canada
Marc Leman, Ghent University, Belgium
Laura Bishop, Austrian Research Institute for Artificial Intelligence, Austria

***Correspondence:**
Auriel Washburn
awashburn@ucdavis.edu

†Present address:
Wisam Reid, Speach and Hearing Bioscience and Technology, Harvard Medical School, Boston, MA, United States

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *01 May 2019*
Accepted: *27 September 2019*
Published: *15 October 2019*

Citation:
Washburn A, Román I, Huberth M, Gang N, Dauer T, Reid W, Nanou C, Wright M and Fujioka T (2019) Musical Role Asymmetries in Piano Duet Performance Influence Alpha-Band Neural Oscillation and Behavioral Synchronization. Front. Neurosci. 13:1088. doi: 10.3389/fnins.2019.01088

Recent work in interpersonal coordination has revealed that neural oscillations, occurring spontaneously in the human brain, are modulated during the sensory, motor, and cognitive processes involved in interpersonal interactions. In particular, alpha-band (8–12 Hz) activity, linked to attention in general, is related to coordination dynamics and empathy traits. Researchers have also identified an association between each individual's attentiveness to their co-actor and the relative similarity in the co-actors' roles, influencing their behavioral synchronization patterns. We employed music ensemble performance to evaluate patterns of behavioral and neural activity when roles between co-performers are systematically varied with complete counterbalancing. Specifically, we designed a piano duet task, with three types of co-actor dissimilarity, or asymmetry: (1) musical role (starting vs. joining), (2) musical task similarity (similar vs. dissimilar melodic parts), and (3) performer animacy (human-to-human vs. human-to-non-adaptive computer). We examined how the experience of these asymmetries in four initial musical phrases, alternatingly played by the co-performers, influenced the pianists' performance of a subsequent unison phrase. Electroencephalography was recorded simultaneously from both performers while playing keyboards. We evaluated note-onset timing and alpha modulation around the unison phrase. We also investigated whether each individual's self-reported empathy was related to behavioral and neural activity. Our findings revealed closer behavioral synchronization when pianists played with a human vs. computer partner, likely because the computer was non-adaptive. When performers played with a human partner, or a joining performer played with a computer partner, having a similar vs. dissimilar musical part did not have a significant effect on their alpha modulation immediately prior to unison. However, when starting performers played with a computer partner with a dissimilar vs. similar part there was significantly greater alpha synchronization. In other words, starting players attended less to the computer partner playing a similar accompaniment, operating in a solo-like mode. Moreover, this alpha difference based on melodic similarity was related to a difference in note-onset adaptivity, which was in turn correlated with performer trait empathy. Collectively

our results extend previous findings by showing that musical ensemble performance gives rise to a socialized context whose lasting effects encompass attentiveness, perceptual-motor coordination, and empathy.

Keywords: EEG, neural oscillation, alpha oscillations, perceptual-motor coordination, role asymmetries, social neuroscience, interpersonal coordination, musical performance

INTRODUCTION

As humans, we face situations every day that demand coordination of our actions with those of other individuals, often in order to achieve a shared goal. Fast-paced and dynamically adaptive sensorimotor interaction can be seen for example, in someone rushing through a crowded airport to catch a flight or a team of paramedics working together to respond to a medical emergency. Here, individuals have to constantly perceive the ongoing actions of others in order to efficiently organize and perform their own actions.

Research investigating the behavioral dynamics that occur between an individual's actions and the environmental events they perceive has provided valuable insight into how behavioral coordination is achieved (Schmidt and O'Brien, 1997; Richardson et al., 2005; Schmidt et al., 2007). Specifically, research on perceptual-motor coordination has demonstrated that individuals often naturally synchronize and coordinate their limb and body movements with periodic environmental events via visual (e.g., Giese et al., 1996), haptic (e.g., Jeka et al., 1998), or auditory (e.g., Repp and Penel, 2004; Repp, 2006) information. A large number of studies have demonstrated that actor-environment coordination is governed by dynamical processes of entrainment, which generally involve close temporal synchronization to an external rhythm (e.g., Kelso et al., 1990; Wimmers et al., 1992; Schmidt and Turvey, 1994; Byblow et al., 1995; Russell and Sternad, 2001; Wilson et al., 2005). Coordinated human joint action contains many of the same characteristics observed within actor-environment coordination. However, the bi-directional coupling inherent to interpersonal coordination commonly results in a mutual influence between interacting individuals. As a result, the patterns of interaction exhibited during interpersonal perceptual-motor coordination are often dynamic.

As noted by Keller (2008), ensemble music performance highlights the ability of humans to achieve temporally precise interpersonal coordination while also being flexible. Keller proposes that three fundamental skills support this kind of interactive behavior: anticipation, the perception of self and other behavior in relation to the joint goal, and adaptation. The relative symmetries and asymmetries between co-actors appear to be one of the primary factors that influence these processes and ultimately shape musical interaction as well as interpersonal interaction in general.

Relative asymmetries between co-acting individuals can arise from a unidirectional informational coupling between co-actors such that one actor receives information about the other's behavior but not vice versa (e.g., Goebl and Palmer, 2009; Washburn et al., 2015). There can also be explicit asymmetries in the intrinsic behavioral component dynamics (i.e., resonant limb/movement frequencies, see Washburn et al., 2014). For instance, pianists who exhibit similar preferred tempi during solo performances achieve better temporal synchronization and exhibit greater adaptation to each other during duet performance than pianists who have more divergent preferred solo performance rates (Loehr and Palmer, 2011; Zamm et al., 2015). While these sources of informational and physical asymmetry clearly play a role in shaping joint action, "functional asymmetries," contextually relevant differences in co-actor roles that can emerge with or without explicit instruction or intention (Richardson et al., 2016), are likely the most common type of asymmetry in everyday interpersonal interaction.

For example, Demos et al. (2017) found that introducing a confederate duet partner to participants as an experimenter vs. a fellow participant introduced an asymmetry in social status. Although this had a minimal effect on temporal coordination during the duet task, participants perceived their synchronization with the "experimenter" as much more successful. The researchers suggest that this effect may have occurred because participants believed that the experimenter's part was especially important and therefore paid more attention to the confederate's performance than participants who thought the confederate was another participant.

People also typically conceive that the members of a string quartet have distinct roles corresponding to their part/instrument that are generally related in a hierarchical fashion (i.e., the first violinist is the leader). Such explicitly prescribed functional asymmetries may indeed guide behavior of the whole group. By experimentally manipulating the leader-follower roles within string quartets such that each instrumentalist had the opportunity to act as the leader of the ensemble, Chang et al. (2017) were able to examine the magnitude and direction of information flow between assigned leaders and assigned followers. Their findings indicated that for a given quartet performance the influence of leader behavior on follower behavior was greater than the influence of followers on the leader, as well as the influence of followers on each other.

However, roles and relationships between group members are often not as static as instrument-specific roles would suggest, depending on how music is written, interpreted, and performed at a moment-to-moment basis. For instance, sometimes someone other than the first violinist will provide cues to the members of the group, acting as a "leader." At other times, the first violinist might engage in repeated turn-taking with another group member, resulting in periods of relative musical role symmetry between the two performers in that their contributions are balanced and equal. Thus, for ongoing interactions that involve numerous opportunities for information exchange between

co-actors asymmetries seem to vary dynamically and are likely to be shaped by multiple factors. Wing et al. (2014) demonstrated this by examining the emergence of functional asymmetries between quartet members in two separate professional string quartets. They found that when performing the same piece of music, the members of the two separate quartets exhibited unique patterns of symmetry and asymmetry at the beat-to-beat timescale. It therefore appears that instrument-specific-role is not the only functional asymmetry determining temporal coordination on the beat-to-beat timescale. What remains unclear are the ways in which different relational factors in music and performers (e.g., instrument, rhythmic similarity, melodic similarity, performer personality traits) might influence the temporal asymmetry in co-performer activity, as well as how these factors interact to shape this relationship.

Within the past decade researchers have conducted a large number of empirical studies aimed at establishing how neural activity supports the emergence and maintenance of coordinative patterns in joint action. The use of electroencephalography (EEG), and its magnetic-counterpart magnetoencephalography (MEG), allow neuroscientists to observe neural activity with high temporal resolution. In particular, the modulation of spontaneously occurring neural oscillations is thought to constitute one of the principal mechanisms for the dynamic coordination of functions across the brain. Fronto-central alpha rhythms (8–12 Hz) along with central beta-band oscillations (∼20 Hz) are sometimes referred to as "mu rhythms" and appear to play a large role in sensorimotor activities. Specifically, their suppression, or event-related desynchronization (ERD), is observed during voluntary movement (Pfurtscheller, 1992; Salmelin et al., 1995; Babiloni et al., 1999; Taniguchi et al., 2000; Jurkiewicz et al., 2006). Pfurtscheller et al. (1997) related this desynchronization to movement initiation, noting that it began prior to movement onset and was followed by event-related synchronization (ERS), or a return to baseline activity, 2 s following a movement onset.

Interestingly, ERD in central alpha rhythms is also exhibited during imagined movements (Pfurtscheller et al., 2006). In fact, alpha ERD was first detected in individuals watching films of biological motion (Gastaut and Bert, 1954). This effect has been replicated in several subsequent studies (Cochin et al., 1998, 2001; Hari et al., 1998; Martineau and Cochin, 2003; Holz et al., 2008; Arnstein et al., 2011). In their original study, Gastaut and Bert (1954) also noticed that the magnitude of alpha desynchronization increased in relation to how much an individual identified with the actor in a film. Additional work has shown that alpha desynchronization during action observation is modulated by the observer's action experience. For example, individuals who are given the opportunity to interact with a novel tool or object show greater alpha desynchronization than participants who don't have the same direct experience when they observe someone else engage with the tool or object (Cannon et al., 2014; Quandt and Marshall, 2014). Professional athletes also display patterns of alpha desynchronization that are distinct from non-athletes when observing videos of their area of expertise (Orgs et al., 2008; Babiloni et al., 2009, 2010).

Given the notable associations between alpha rhythm desynchronization and both voluntary movement and action observation, it is not surprising that these oscillations are also responsive to social, interactive behaviors. Recent dual-EEG studies examining the oscillatory neural activity of co-acting individuals have shown both within- and between-brain coherence in frontal and central alpha rhythms during cooperative, coordinative interaction (e.g., Cui et al., 2012; Sanger et al., 2013). This kind of dual-EEG recording, or "hyperscanning," has also allowed researchers to identify many different factors associated with alpha modulation in each of the co-actors involved in an interaction. Tognoli et al. (2007), for example, observed desynchronization of right centro-parietal alpha rhythms when participants engaged in a simple finger-tapping task in a social context, suggesting that this decrease in oscillatory power may support somatosensory awareness of a perceived co-actor.

Findings connecting alpha desynchronization to action, action-observation and interactive behaviors have also linked alpha desynchronization to the theoretical human mirror-neuron system (MNS) (Iacoboni and Dapretto, 2006; Oberman et al., 2007; Perry and Bentin, 2009; Frenkel-Toledo et al., 2014; Hobson and Bishop, 2016). A number of studies that relate alpha modulation to MNS activity have revealed desynchronization in right centro-parietal areas during social interaction (e.g., Tognoli et al., 2007; Dumas et al., 2010, 2012; Naeem et al., 2012a,b). However, alpha ERD has also been observed in a variety of other regions during interactive behavior including left centro-parietal, frontal, central, and central midline areas (Lachat et al., 2012; Ménoret et al., 2014; Konvalinka et al., 2014; Ahn et al., 2018). This emphasizes the importance of alpha's role in supporting social interactive behaviors through domain-general regulatory functions rather than domain- and location-specific sensorimotor processes. Of particular interest here regarding the functional significance of alpha is its apparent role in regulating the dynamic desynchronization and selection of cortical states (e.g., Jin et al., 2006; Klimesch et al., 2007; Klimesch, 2012). This links alpha closely to mental states of alertness, expectancy and attention (e.g., Klimesch et al., 1998; Pfurtscheller, 2003; Perry and Bentin, 2009), and to the temporal coordination of intrapersonal and interpersonal behavior often supported by these states in the context of interpersonal interaction. This is consistent with Novembre et al. (2016) observation that alpha desynchronization occurred during periods of strong temporal entrainment between co-actors during piano duet performance. Existing findings therefore indicate that alpha ERD is likely to occur frequently in both individuals involved in a joint action task.

Interestingly, the evolution of alpha desynchronization during temporal coordination is related to asymmetries in co-actor roles (Konvalinka et al., 2014). Specifically, for pairs engaged in a synchronized finger-tapping task, the individual who exhibited less adaptive, or more leader-like, behavior generally displayed greater alpha desynchronization in frontal brain regions. This demonstrates that the dynamics of alpha desynchronization are sensitive to subtle, emergent asymmetries between interacting individuals. Relatedly, Sanger et al. (2013) observed that the coherence between frontal alpha oscillations in co-performing guitarists was stronger for leader-to-follower coupling than for

follower-to-leader coupling. In Sanger et al. (2013) study the roles of leader and follower were explicitly assigned to co-performers prior to duet performance.

Hari and Kujala (2009), among others, have pointed out the apparent links between alpha ERD and the neural processes that support action observation and ultimately interpersonal interaction, including motor imitation, emotional contagion, and empathy. Babiloni et al. (2012) were the first to empirically investigate the relationship between alpha ERD and empathy by having saxophonists observe their own previously recorded ensemble performances. The findings from this study revealed that the musician with the highest score on the self-report empathy measure showed widespread alpha desynchronization during performance observation. Perspective-taking also appears to be related to perceptual accuracy in the context of action-observation. Work by Engel and Keller (2011), for example, has shown that for individuals viewing point light displays (PLDs) of improvised or imitated musical performances, the person's accuracy for identifying the type of performance was positively correlated with scores on a self-report measure of perspective-taking. Relatedly, Pecenka and Keller (2011) observed that individuals who reported higher levels of perspective-taking behavior showed greater anticipation of tempo-changing metronome sequences. Collectively, existing work on empathy and social interaction reveals strong links between trait empathy and both (1) alpha desynchronization, and (2) behavioral coordination characteristics.

In the current study, we used the context of piano duet performance to evaluate patterns of behavioral and neural activity under conditions of asymmetry between co-performers. To do this, we experimentally introduced three different task-specific asymmetries: (1) musical role (starting vs. joining), (2) task similarity (similar vs. dissimilar melodic parts), and (3) performer animacy (human-to-human vs. human-to-non-adaptive computer). The musical tasks used in the current study were piano duets, played by two players face-to-face, each with an electronic keyboard (see **Figure 1**). The piano duet scores were composed by our research team, consisting of simple short melodies designed so that each of the three experimentally introduced asymmetries were experienced during the initial portion each trial prior to the final unison period. In this final unison period of four notes, co-performers always had identical musical tasks.

We designed this experimental task to support a multi-faceted study which allowed our team to perform two separate investigations. In the investigation presented here, we evaluated neural alpha and behavioral activity during the final unison measure of the duets to determine how the asymmetries in the preceding portion of the task served as a priming context and shaped the interaction of subsequent unison performance. The other investigation, described in Huberth et al. (2019), was focused on performers' neural responses to outcome expectation violations when altered pitch feedback was experimentally introduced within the first four measures of the turn-taking duet performances. Specifically, the authors compared the feedback-related negativity (FRN) and P3 complex within each individual's EEG data in response to altered pitches in one's own part vs. the co-performer's part. Generally, performers exhibited greater responses to alterations corresponding to their own part, especially when their part was melodically similar to their co-performer's. Our current investigation is based on behavioral and EEG data from only the fifth and final measure of the duets where no altered pitch feedback was used and the notes performers had to play stayed identical across the different conditions experienced in the preceding part of the duet. Thus, the task was designed to accommodate two investigations with one round of data collection with each investigation ultimately involving unique data analysis.

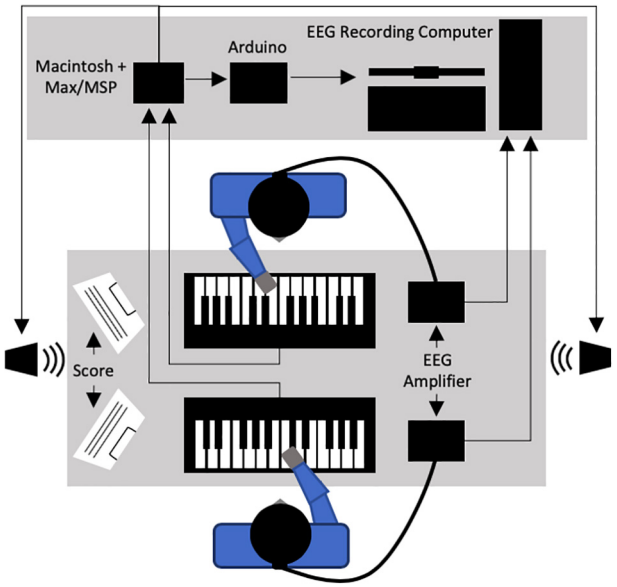

FIGURE 1 | General experimental set-up for the current study.

Based on the above literature, we had several hypotheses about the reactivity of alpha activity and performers' coordinative behaviors in our task. First, however, it is important to note that differences in starter and joiner movement onset immediately before the unison prevented us from directly comparing the effects of performing the starting vs. joining role on note-onset synchronization or alpha desynchronization. Because starting performers played three notes immediately prior to the unison part, they would show more pronounced alpha desynchronization during the unison. Second, we hypothesized that performer animacy would have a significant effect on the stability of note-onset synchronization. This expectation was informed by Fairhurst et al.'s (2013) findings that the stability of human tapping with a virtual partner was influenced by the adaptability of the partner, with both low and high adaptability leading to low synchronization stability. Based on work by Billeke et al. (2014) we also expected that individuals might show greater neural responsiveness in the alpha oscillations during interaction with a human partner compared to a computer partner.

As for the task similarity, we are not aware of existing studies illustrating the effects of any type of task similarity on neural activity during joint action. On the one hand, the finding

that action familiarity has an effect on neural measures during action imagery and observation as well as interaction (Orgs et al., 2008; Babiloni et al., 2009, 2010; Cannon et al., 2014; Quandt and Marshall, 2014; Novembre et al., 2016) indicates that greater melodic similarity between parts may engage more neural resources in duet performing musicians. On the other hand, Skewes et al. (2015) observed that task asymmetries for a dyadic aim and click task had minimal effects on co-actor synchronization, except in cases where the level of difficulty was substantially different between co-actors. Thus, we expected that the effect of having similar or dissimilar musical tasks might be the most subtle of the three asymmetries we manipulated in the current study. However, as shown in Babiloni et al. (2012), we also saw this manipulation as highly relevant to trait empathy and an important factor for understanding asymmetries in musical interaction and social interaction in general and were interested to see how it might shape coordinative activity in interaction with the other two asymmetries we examined.

Lastly, we would like to note that our investigation on the effects of task-specific asymmetries on temporal coordination behaviors during an experimentally-controlled musical duet task has implications for both musical and non-musical everyday interactions. Our study design allowed us to unambiguously examine how some of the asymmetries that are common to ensemble music performance might interact to shape musical interaction, which contains various relationships between performers that emerge and change dynamically as music unfolds. Notably, our study also has implications for interaction dynamics beyond musical performance. Any time multiple individuals coordinate to achieve a shared goal there will be some discrepancies between individuals in their attentiveness and adaptivity to one another as well as in their behavioral timing. Asymmetries between individuals can lead to these kinds of discrepancies, and the evolution of discrepancies over time will shape the course of interaction. Returning to our example of the paramedic team, one can imagine how the collective activity of the team will be influenced by one member arriving before the others, the need for team members to address an individual's multiple, similar or different injuries, and the need for each team member to multi-task, limiting their ability to attend and adapt quickly to team member actions. These interpersonal asymmetries and others will have a significant impact on how each member of the team experiences the actions of their team members and acts to support team success. By gaining a greater understanding of the effects of co-actor asymmetries we therefore broaden our understanding of how individuals interact to achieve collective goals.

MATERIALS AND METHODS

Participants

Twenty-four pianists (twelve pairs) were recruited from the Stanford University community for participation in this study. Nine pianists were removed from data analysis due to technical failures ($N = 3$), high behavioral error rates ($N = 4$), and excessive data artifacts ($N = 2$). Compared to the participant data analyzed in Huberth et al. (2019), we removed three additional pianists because of errors ($N = 2$) and artifacts ($N = 1$) that specifically affected our analyses. The four pianists with the high error rates comprised two pairs. Results reported here are therefore for a sample of 15 pianists ($M = 14.33$; $SD = 4.92$ years musical experience). These pianists ranged from 18 to 28 years of age.

Of the pairs recruited, two pairs knew each other and had played duets together prior to the experiment. Only one of these pairs' data was included following the removal of participant data prior to analysis as described above. The remaining pianists met for the first time during participation in the study. All pianists were right-handed except for one, who was not included in the data for analysis. The study protocol was approved by the Stanford University Institutional Review Board and participants provided written informed consent. Pianists were paid $20/h for their participation.

Apparatus

Two Yamaha Axiom-61 digital keyboards were arranged facing each other on a table within a sound-shielded room in the lab (see **Figure 1**). Two loudspeakers were used to provide auditory feedback to the performers during the study, with one placed at each end of the table. A custom module for Max/MSP 7.0.1 run on a Macintosh computer (OSX 10.9.5) was used to control all auditory feedback throughout the study. The piano timbre used throughout and the drum timbre used for introductory metronome clicks were built-in sounds from the OSX MIDI sound synthesizer, AU DLS Synth. All auditory feedback was played at a constant volume of approximately 75 dB SPL throughout the study (i.e., pianists were not able to produce changes in dynamics during performance).

The Max/MSP program was also used to generate trigger codes associated with meaningful timepoints and experimental conditions as they occurred in each task trial. This included tracking pianist performance for note accuracy and inter-onset-interval (IOI) in real time based on the current musical score. These codes were sent through an Arduino Uno to the computer used to record EEG data in order to achieve temporal mapping between each EEG recording and the event time course of the musical task being completed.

The component latencies produced by this apparatus were evaluated by comparing the onset latency of (1) a piano key press, (2) the resulting trigger code produced by Max/MSP, and (3) the associated auditory feedback using simultaneous three-channel audio recording of all three events (see Wright et al., 2004). The average key press to trigger code onset was 27 ms ($SD = 4.0$ ms), and the average delay between a trigger code and the associated auditory feedback was 21 ms ($SD = 3.3$ ms).

Stimulus and Task

We composed four unique piano duets for the current study. All duets had the same five measure structure (see **Figure 2**). This included four initial measures in which only one pianist played at a time and the partners alternated each measure. This meant that the "starting" player for a given trial played in measures one and three while the "joining" player played in measures two and four. In two of the four duets the starting and joining

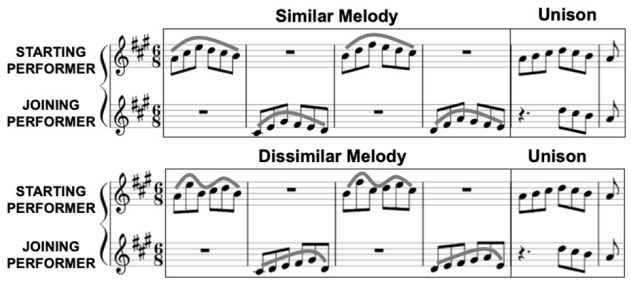

FIGURE 2 | Examples of the similar **(top)** and dissimilar **(bottom)** musical task duets composed for the current study. In each duet the starting and joining performers alternated playing in the first four measures, and played the last four notes of the fifth measure in unison. Melodic contour lines show that in the "similar" task condition the starting and joining performers played similar melodic patterns, while in the "dissimilar" task condition the patterns were distinctly different. The final unison measure was the same across all duets. All of the analyses we conducted in the current study were for neural activity and note-onset behavior during this final measure.

parts had similar melodic contours in the first four measures, while in the other two duets the contours were distinctly different (within a given player's part the two phrases they played alone always had the same contour). Each duet also included a final fifth measure, in which the starting player played the first half alone and both pianists played the remainder together in unison. The notes played in unison were identical for both the starting and joining players and were the same across all four duets. Duets were composed so that each part could be played with the right hand alone and overall hand position could remain the same throughout. Fingering numbers were provided in the score to encourage consistency between participants and ensure minimal movement.

During the study each pianist played the duets with both a human co-performer and a computer partner (audio only). The computer partner was non-adaptive and produced notes with a constant IOI of 500 ms, or a tempo of 120 bpm for the eighth note. Each experimental trial began with three isochronous metronome clicks corresponding to the eighth note, with this same IOI of 500 ms regardless of whether the trial was to be performed by two pianists or one pianist and the computer partner. Following these clicks the starting player began the first measure. After the completion of a trial, the introductory metronome clicks indicating the start of the next trial would begin after a random interval of silence (1.5–2.5 s).

In the first four measures of each duet the auditory feedback associated with key presses was sometimes altered. In other words, even though a pianist had played the correct key based on the musical score a different pitch would be presented as auditory feedback. These manipulations were introduced within both parts of each duet to evaluate the feedback-related negativity (FRN) associated with altered feedback corresponding to a pianist's own part and their partner's part as identified using EEG. In each trial, one pitch alteration occurred in each performer's part. This occurred on either the 4th or 5th note of one of the two phrases each performer played. All alterations produced in-key pitches that were ± two scale notes from the printed score note. Further information about these manipulations and the associated analyses and findings are presented in Huberth et al. (2019).

In the current study, our analysis focused solely on the fifth measure of each duet, which did not contain any altered auditory feedback. It is important to note that we fully counterbalanced the altered note position, pitch direction, and frequency of occurrence of the altered feedback such that these manipulations did not influence the data presented here. The last possible position for altered feedback (in the fourth measure, affecting the joining performer's part) occurred at least one second before the onset of the data epoch used in the current study. For the present study we established that there was no significant difference in either starter or joiner behavior when there was an alteration in the fourth measure compared to when there was not. Thus, we have collapsed data across these conditions. Using this design we were able to assess the influence of asymmetries in (1) musical role (starting vs. joining), (2) task similarity (similar vs. dissimilar melodic parts) and (3) partner animacy (human vs. computer) experienced in the preceding part of the piano duet performance on behavioral and neural activity during joint performance of the same musical sequence within the fifth measure. All analyses presented in the current paper are for the time period associated with this final unison measure.

Procedure

One pianist from a pair was randomly selected to arrive at the lab first, be prepared for EEG recording, and complete the first half of the study with the computer partner. While this individual was performing with the computer partner, the other pianist arrived in the lab and was prepared for EEG recording. When the initial pianist was finished playing with the computer partner, both pianists played together and then the second pianist completed the latter portion of the study with the computer partner. At the end of each pianist's recording session we asked them to complete two questionnaires: the 40-question version of the Cambridge Empathy Scale (Baron-Cohen and Wheelwright, 2004), and a custom 10-question measure designed to ascertain each participant's prior familiarity with their human partner and their experience of the task difficulty.

Upon arrival in the lab each pianist was introduced to the four duet scores composed for the current study. They were asked to memorize the duets as quickly as possible, but were informed that the printed scores would be placed next to their keyboard throughout the study should they need to reference them. All pianists were also asked to keep a fixed gaze at a comfortable location during trials in order to avoid excessive eye movement artifacts in the EEG recording. Before starting the first block of trials each individual was informed that throughout the study the auditory feedback associated with key presses would sometimes be altered and were asked to continue playing even if they heard an incorrect pitch (see previous section).

Each pianist played four blocks of trials with the computer partner and four blocks of trials with their human partner for a total of eight blocks. Partners played a single duet score within each block, with the starting and joining roles fixed throughout

the block. It is important to note that the difference in these roles was communicated to participants only with respect to who played first at the beginning of the trial, and was not characterized as a "leader" vs. "follower" musical relationship. Partners played the same duet score for two back-to-back blocks before switching to another score. This allowed us to counterbalance musical role so that each pianist played both the starting and joining role for each duet they played (e.g., if a person played the starting role for a duet in Block 1 the same person would play the joining role for the same duet in Block 2). This organization allowed all pianists involved in the current study to play each of the eight unique musical parts. We also counterbalanced the presentation of task similarity conditions such that each pianist played one melodically "similar" and one melodically "dissimilar" duet with both the human and computer partner. The order of presentation of the specific duet scores was counterbalanced across all participants as well.

The first two trials of every block were treated as practice trials and contained no pitch alterations. These practice trials proceeded directly into the 48 target trials for the block. As noted above, the location, direction and magnitude of altered pitch feedback within the first four measures of each trial was counterbalanced within each block (see Huberth et al., 2019 for details). We used the Max/MSP program to control sound and evaluate performance in each trial. A trial was counted as incorrect if either partner (1) pressed an incorrect key based on the score for that trial or (2) produced an IOI more than 125 ms shorter or longer than the expected IOI (500 ms). When an error occurred, the auditory feedback was stopped immediately to signal to pianists that they should prepare for the next trial. A short period of silence was maintained before the next trial began.

The length of this period was randomly selected from a uniform distribution of values representing each possible integer value in milliseconds between 1.5 and 2.5 s duration. Each trial classified as incorrect was appended to the end of the current block and had to be performed in order for the block to be complete. A single block took approximately 15 min to complete. Max/MSP recorded all note onset timing information from both keyboards as well as the success/error status of all trials in log files, which were later used for behavioral data analysis.

Electroencephalography data were collected either from a single pianist during interaction with the computer partner, or simultaneously from both pianists. All recording took place in a sound-attenuated and electromagnetically-shielded chamber within the lab. A member of the research team monitored participant compliance with stated instructions via a window from an adjacent room. Participants were encouraged to take brief breaks between blocks when needed. An experimental session for a single participant took between 3.5 and 4 h.

EEG Recording and Preprocessing

Electroencephalography data were collected using whole-head, 64-channel Neuroscan Quik-Caps (10–20 system), a SymAmpRT amplifier, and Curry 7 acquisition software (Compumedics Neuroscan Inc., El Paso, TX, United States). This included the recording of vertical and horizontal electrooculograms (EOG). Recordings were made at a 500 Hz sampling rate. Electrode impedances were kept below 10 kΩ throughout recording. Scalp electrodes on each cap were referenced to a midline electrode between CPz and Cz for recording. Prior to analysis, the raw recordings from each individual were re-referenced using the common average reference for the cap. The SymAmpRT amplifier allows for simultaneous recording from up to four caps, precluding the need for a temporal synchronization mechanism between the EEG recordings from duet partners. We processed and analyzed EEG data in MATLAB (Mathworks Inc., Natick, MA, United States) using custom scripts which incorporated routines from the Brainstorm toolbox (Tadel et al., 2011).

We removed eye artifacts from the continuous EEG data via Source Space Projection routines provided by the Brainstorm toolbox. First stereotypical eye-artifact events (blinks and movements) were detected using a single, continuous representative raw file for each participant. Then using these events a set of projectors for the participant was constructed and applied to each of the participant's trials in order to remove blink and movement artifacts. This automatically removed projectors that explained a substantial amount of variance in the participant's data (in our case more than 15% of the time). In addition, we chose to remove any additional projectors with a pattern of largely lateralized topography in order to avoid any spurious effects on the comparison of right and left electrode groupings.

We created epochs using a time window between −1.0 s before and 4.0 s after the onset of the fifth measure (i.e., total epoch duration was 5 s) from correctly-completed trials. This resulted in 48 epochs per condition per participant. Within each epoch, any channels exhibiting peak-to-peak amplitudes ±150 μV were rejected. We employed this channel rejection threshold in the current study so that we could conduct analyses on as many EEG trials as possible, and avoid artifacts resulting from any large amplitude changes. The evoked response for each condition was calculated by averaging trial epochs for that condition across participants, using a baseline period of 50 ms before the onset of the measure.

Measures and Analyses

In this study we were interested in evaluating the effects of the three co-performer asymmetries introduced experimentally on behavioral and neural activity during musical unison. In order to assess behavioral coordination we evaluated the *Note-Onset Asynchronies* between two co-performers during musical unison. To examine the neural activity exhibited by each individual before and during unison we investigated the occurrence of *Alpha Desynchronization* using EEG. We also measured each individual's trait *Empathy* so that we could investigate potential associations between empathy, movement asynchrony, and neural alpha desynchronization in our study's performance task.

Note-Onset Asynchronies

For each experimental trial we calculated a Note-Onset Asynchrony measure between players during the concluding unison segment, which consisted of the final four notes in each duet and was identical for both players and throughout

all conditions. This analysis required data for both performers within a duet. All behavioral analyses were therefore conducted on the seven complete pairs within our data set (14 participants). We first determined the note-onset time for each of the four notes in the unison segment as executed by each player (i.e., starter and joiner). We then found the asynchrony between players for each note by subtracting the key press time for one player from the corresponding key press time of the other player. We did this separately from each player's perspective so that we had two series of asynchronies values for each trial, one for the starting player and the other for the joining player. Specifically, for the starting player series we subtracted joiner timing from starter timing and vice versa for the joining player series. As a result, if the current player played a note before the other player, this would result in a negative asynchrony value for the current player and if the current player played after the partner it would result in a positive asynchrony value.

For each trial we used each player's four-note asynchrony series to calculate two measures: mean asynchrony and standard deviation of asynchrony. We then averaged these single-trial values across the trials corresponding to each musical role × task similarity × partner animacy experimental condition to gain measures of average asynchrony and asynchrony variability. In order to account for possible deviations in tempo between performers, we ultimately divided each performer's average asynchrony and asynchrony variability for each experimental condition by their average IOI in that condition. We present these measures as the percent asynchrony exhibited based on IOI.

Alpha Modulation

We computed normalized power within the alpha frequency band for each combination of musical role × task similarity × partner animacy conditions in three steps. This analysis included all 15 of the participants retained in the current study, coming from a total of eight different duet pairs.

(1) As noted above, single-measure epochs from each participant for a given condition were averaged to produce the associated evoked response for the condition. We then subtracted this within-participant condition average from each participant's own trial epochs in order to derive the induced response exhibited by each participant in that condition.
(2) We used these resultant epochs to generate time-frequency representations (TFRs) for each epoch using Morlet wavelet decomposition with 32 logarithmically-spaced bins from 1 to 60 Hz. Brainstorm routines were employed to calculate the *z*-score normalized signal power in each bin for a single epoch as the product of each wavelet coefficient and its complex conjugate. We then averaged TFRs across epochs to generate a characteristic TFR for each experimental condition. During this process we observed that for some average TFRs, normalized power seemed to be spuriously concentrated a specific TF region (i.e., very little relative power was observed for the majority of the spectrum). For these averages we made a close examination of each of the contributing epoch TFRs in order to identify the source of the seemingly artifactual relative power concentrations. Oftentimes these were the result of extreme changes in single channel behavior within a few single trials per condition or single trials in which a number of channels simultaneously exhibited noisy behavior. We removed such channels and trials. This affected an average of 2.4 trials per condition, per participant. A total of 97.78% of the original, correctly performed trials are ultimately included in the data presented here.
(3) Using the corrected TFRs we extracted the alpha-band time course in each condition by averaging normalized power across the four existing frequency bins corresponding to the alpha-band (8.2, 9.3, 10.4, 11.7 Hz). This allowed us to establish each participant's average alpha activity for a given experimental condition, which we low-pass filtered at 8 Hz to remove transients. The resultant time series was baselined using a period of 80 ms before the onset of the measure. We then calculated the average activity within each of three electrode groupings: frontal-centro-medial (fcm: F1, Fz, F2, FC1, FCz, FC2, C1, Cz, C2), parietal left (pl: C1, C3, TP7, CP5, CP3, P7, P5, P3), and parietal right (pr: C2, C4, CP4, CP6, TP8, P4, P6, P8).

Alpha ERD amplitude at the unison onset was the main feature of interest for statistical analysis in the current study. To establish our exact analysis time window we first generated the average alpha ERD waveform associated with the fifth measure, including each of our three electrode groups of interest across both starting and joining roles. We then identified the latency of the negative ERD peak (i.e., trough) closest to the expected unison onset time (i.e., 1.5 s after the onset of the fifth measure).

We established the time window of interest around the time point associated with the unison-related trough using the method suggested by Keil et al. (2014) for selecting a window without bias. The process for identifying this time window required that we first locate the positive peaks on either side of the trough. For the peak preceding and the peak succeeding the trough, we measured the absolute value of the difference in trough-peak amplitude. We then found the time point on the ERD grand average trajectory associated with an amplitude 50% of the total corresponding trough-peak amplitude. Through this process we obtained a time window for evaluating unison onset-associated activity of 1.12 to 2.36 s after the onset of the fifth measure.

Empathy Quotient

We used the 40-question Cambridge Empathy Scale (Baron-Cohen and Wheelwright, 2004) to establish each participant's trait empathy, or empathy quotient (EQ). A higher EQ is indicative of greater empathy toward others.

RESULTS

Task Performance Evaluation

Throughout our study we maintained two criteria for successful trials: (1) correct keypresses based on the score and (2) a

consistent tempo, defined as every note IOI being between 375 and 625 ms (±25% from the correct IOI of 500 ms). Trials which included violations of either criterion were classified as error trials (see Procedure section for details). To examine whether performer error rates were associated with increasing in fatigue as the study progressed, we compared the number of errors exhibited in the first vs. second half of each study block. This revealed no differences, establishing that participants did not commit significantly more errors later in the block.

In order to further examine the possible deviation of performance tempo from the expected IOI, we also conducted separate 2 (task similarity: similar, dissimilar) × 2 (partner animacy: human, computer) × 7 pair (seven unique performer pairs) mixed-model analyses of variance (ANOVAs) on the average IOI during the analysis time window for each of the starting and joining roles.

For performers assuming a starting role we found no significant interactions between variables, but we did find a significant main effect of partner, $F(1,6) = 62.50$, $p < 0.001$, $\eta_p^2 = 0.90$. As can be seen in **Table 1**, IOIs were shorter when participants played with a human partner.

For performers assuming a joining role, we found a similar main effect for the partner condition, $F(1,6) = 75.53$, $p < 0.001$, $\eta_p^2 = 0.92$. We also observed a significant main effect of pair for this role, $F(1,6) = 3.95$, $p = 0.047$, $\eta_p^2 = 0.77$. Fisher's LSD *post hoc* comparisons revealed that two pairs generally exhibited shorter IOIs than some of the other pairs during the human partner conditions. Therefore, the subsequent analysis of note onset asynchronies was normalized to each performer's average IOI value in a given condition.

Note-Onset Asynchronies

We assessed the effects of task similarity and partner animacy in the current study separately for the starting and joining conditions based on the distinct movement requirements preceding unison.

We conducted separate 2 (task similarity) × 2 (partner animacy) × 7 (pair) mixed-model ANOVAs on the average note-onset asynchrony measure normalized by each performer's average condition IOI for each of the starting and joining roles. For performers assuming a starting role, there was a significant main effect of partner animacy, $F(1,6) = 14.49$, $p = 0.007$, $\eta_p^2 = 0.67$, but no main effect of task similarity or pair or interactions between variables. We observed a similar pattern of results for performers assuming a joining role, again with a significant main effect of partner animacy, $F(1,6) = 7.70$, $p = 0.03$, $\eta_p^2 = 0.52$, but no main effect of task similarity or pair or interactions between variables. As can be seen in **Figure 3**, performer note-onsets generally arrived ahead of the computer partner but were much more closely synchronized to the human partner.

We also conducted separate 2 × 2 × 7 ANOVAs on the standard deviation of average note-onset asynchrony measure for starting and joining performers. For starting performers we found a significant main effect of partner animacy, $F(1,6) = 29.96$, $p = 0.001$, $\eta_p^2 = 0.81$, but no main effect of task similarity or pair or interactions between variables. For joining performers, we also found a main effect of partner animacy, $F(1,6) = 57.26$, $p < 0.001$, $\eta_p^2 = 0.89$, but no main effect of task similarity or pair or interactions between variables. **Figure 4** shows both starting and joining performers exhibited greater variability in note-onset asynchronies when playing with a human co-performer as compared to the computer.

TABLE 1 | Performer IOIs during Unison across Experimental Conditions.

		IOI (ms)			
		Starter		Joiner	
Partner	**Task Similarity**	***M***	***SD***	***M***	***SD***
Human	Similar	467.58	17.23	466.82	15.92
	Dissimilar	471.27	14.83	470.82	14.07
Computer	Similar	500.28	5.74	497.89	3.64
	Dissimilar	487.792	3.65	496.71	2.94

Alpha Desynchronization

We conducted separate 2 (task similarity) × 2 (partner animacy) × 3 (electrode group: fcm, pl, pr) × 8 (pair) mixed-model ANOVAs on alpha activity for each of the starting and joining roles. As we have discussed in the "Measures and Analyses" section, it was necessary for us to use just the seven full pairs for the behavioral analyses, but we included all 15 participants in the alpha analyses resulting in eight-levels for the pair factor. A time window of 1.12 to 2.36 s after the onset of the fifth measure was used to capture activity related to the unison onset which started on the fourth note of the measure, at approximately 1.5 s. **Figure 5** illustrates the time course of normalized alpha power during this fifth measure.

As can be seen in **Figure 5A**, the alpha modulation range for performers in a starting role during the unison measure was often small because they were already moving during this time window. When pianists played with a human partner these fluctuations reflected the temporal regularity of the note-onset actions as well as slightly larger desynchronizations prior to the unison onset in some conditions, but did not reveal substantial differences between having similar or dissimilar musical parts. In contrast, for starting performers playing a similar part to their computer co-performer there was a large alpha synchronization prior to the unison onset, leading to a sustained difference in alpha activity compared to performance of dissimilar musical parts across the time window of interest. A 2 × 2 × 3 × 8 ANOVA conducted on alpha modulation for performers playing the starting part revealed a significant interaction of task similarity and partner animacy, $F(1,7) = 5.82$, $p = 0.047$, $\eta_p^2 = 0.45$, but no other interactions or main effects. To explore this interaction we collapsed across the electrode groupings, released pair as a factor and conducted a simple effects analysis evaluating the effect of task similarity when interacting with (1) a human partner and (2) the computer partner. We found a significant effect of task similarity for individuals interacting with the computer partner, $F(1,14) = 5.20$, $p = 0.04$, $\eta_p^2 = 0.27$, but not with a human partner. Specifically, pianists performing with the computer partner

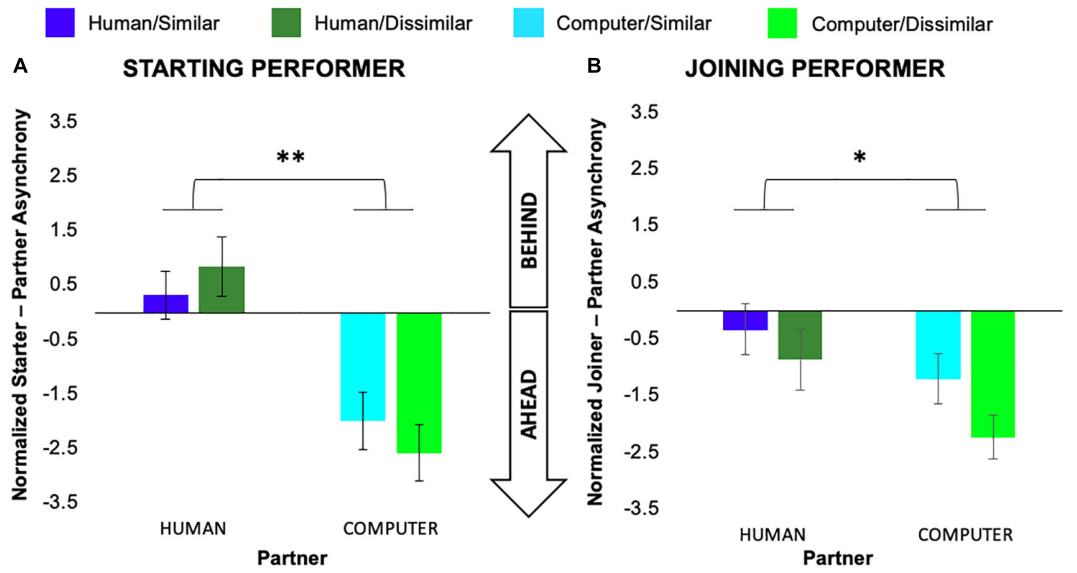

FIGURE 3 | Normalized average asynchronies exhibited by participants in the role of **(A)** starting and **(B)** joining performer. Error bars show standard error. $^{*}p < 0.05$, $^{**}p < 0.01$.

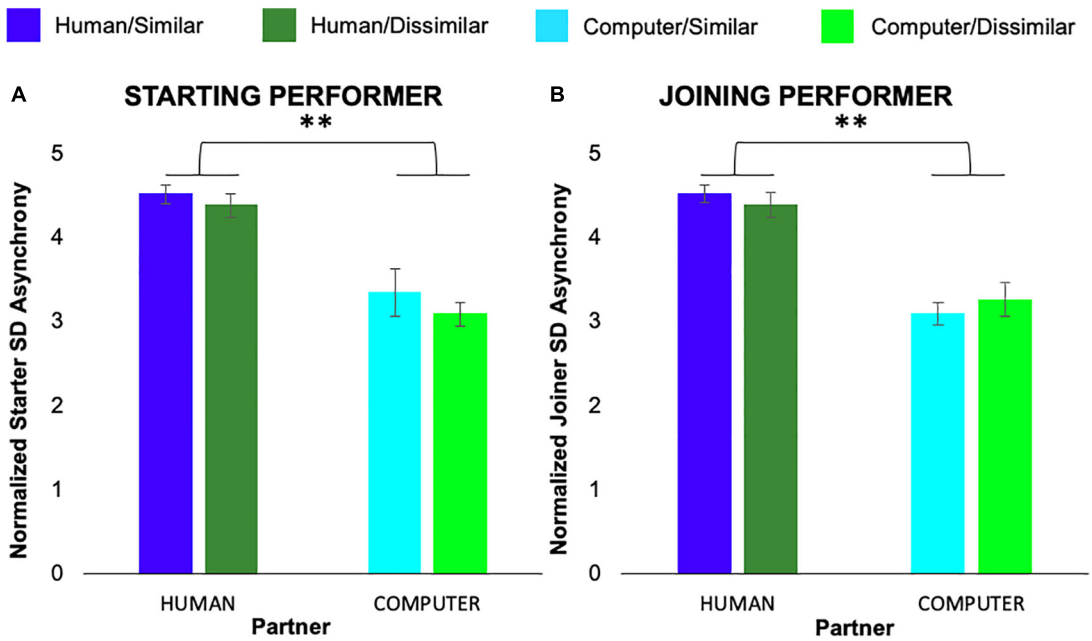

FIGURE 4 | Normalized average standard deviation of asynchronies exhibited by participants in the role of **(A)** starting and **(B)** joining performer. Error bars show standard error. $^{**}p < 0.01$.

exhibited a moderate synchronization (e.g., alpha power increase) in the "similar" task condition and a minor desynchronization (e.g., alpha power decrease) in the "dissimilar" task condition.

Figure 5B illustrates that the time course of alpha modulation in joining performers was different from that for starting performers because the joining performers finished playing measure four, rested for the first half of measure five, and then began to play again from the unison onset which occurred halfway through measure five. As a result, they often show a characteristic rebound around 0.5 s, followed by a desynchronization or return to baseline prior to beginning to play at the unison onset. A 2 × 2 × 3 × 8 ANOVA on alpha modulation in joining performers revealed no significant interactions between variables or main effects.

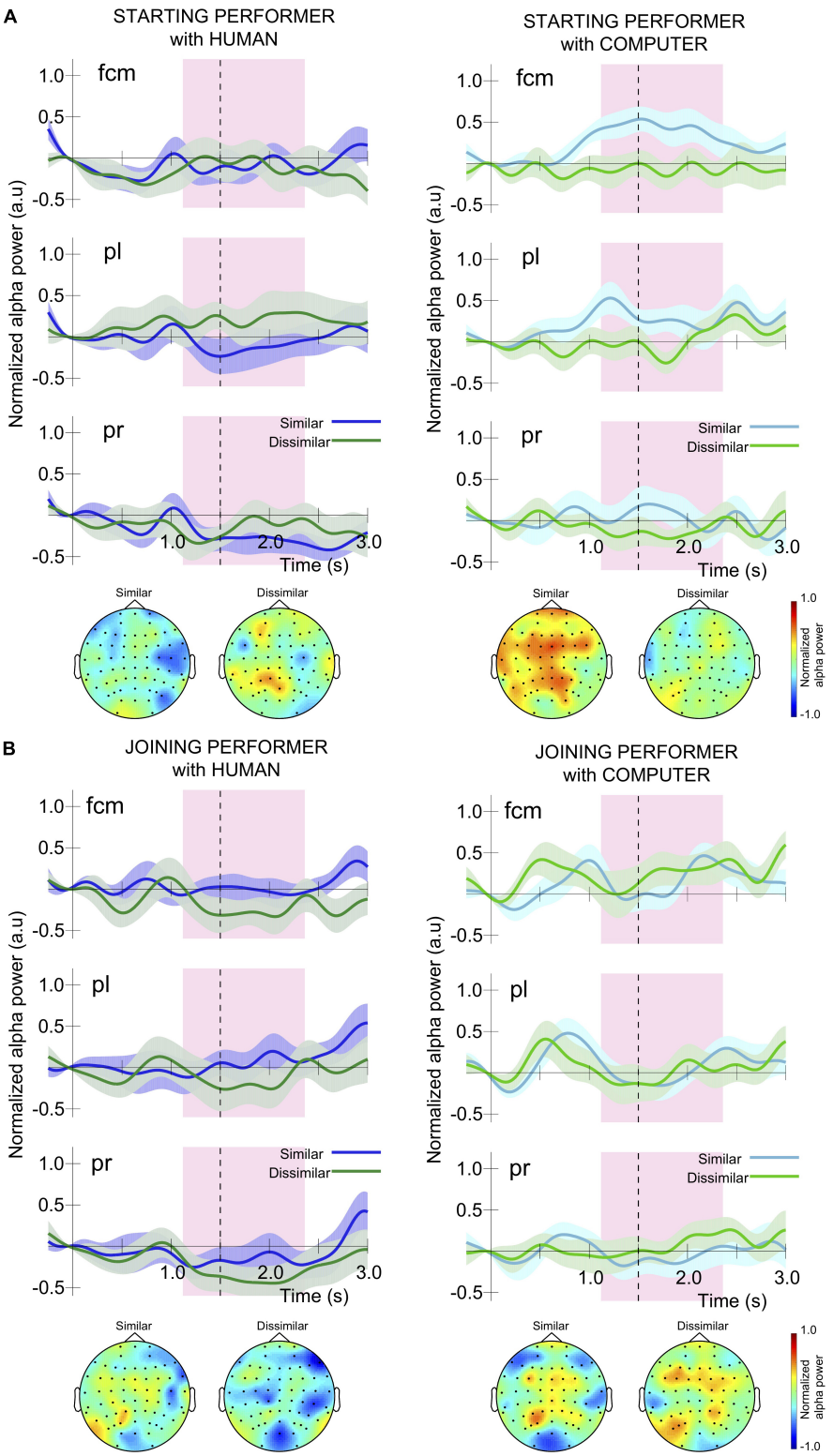

FIGURE 5 | Time course of grand average normalized alpha power around the final unison measure for participants in the role of **(A)** starting and **(B)** joining performer, in the three electrode groups, fcm, pl, and pr. The onset of the measure is designated as time 0 and the starter played three notes before the joiner started playing. The fourth note of the measure, and first note of unison, occurred around 1.5 s and is marked by a black dashed vertical line. The shaded purple rectangle corresponds to the time window around unison used for statistical comparison between conditions (1.12 to 2.36 s). The topography associated with this time window in each condition is also provided. Shading around the alpha power time course for each condition corresponds to the standard error.

Associations Between EQ, Note-Onset Asynchrony, and Alpha Modulation

We assessed possible associations between performer EQ, alpha modulation, and note-onset asynchrony behavior based on the aforementioned note-onset asynchrony and alpha modulation results. Consistent with the note-onset asynchrony and alpha modulation results presented, we analyzed data separately for individuals assuming starting vs. joining roles. Notably, we established that there was no significant difference in EQ between performers who arrived first (performing with the computer partner and then their human partner) and those who arrived second (performing with their human partner and then the computer partner).

Starting Role

Our analysis of note-onset asynchronies for performers in a starting role revealed a significant effect of performing with a human vs. computer partner on both average asynchronies and asynchrony variability. Namely, performers more frequently played notes ahead of the computer partner but exhibited greater variability of onset asynchronies with the human partner. Based on these results we first established the participant-wise human vs. computer condition difference (HvC difference), averaged across the similar vs. dissimilar task conditions, for both (1) average asynchronies and (2) asynchrony variability. We then compared each of these variables to (1) the HvC difference in alpha activity and (2) performer EQ. In comparing the note-onset asynchrony variables to the HvC difference in alpha activity we looked at associations with each of the individual electrode groups, as well as activity averaged across electrode groups. None of these correlations revealed significant associations between note-onset asynchrony behavior and alpha activity or performer EQ related to the difference in interaction with a human vs. computer partner.

Our analysis of alpha activity indicated a significant difference between the similar vs. dissimilar conditions for performers playing the starting role during interaction with the computer partner. We were interested in exploring this effect further by evaluating any potential associations between this difference and performer EQ as well as note-onset asynchrony behavior.

For each performer assuming the starting role and performing with the computer partner we first identified the difference in alpha activity between the similar vs. dissimilar conditions (SvD difference), collapsing across our three electrode groups. We also identified the SvD difference in the performer's average asynchrony, as well as the SvD difference in the standard deviation of their asynchronies (i.e., the difference in asynchrony variability). We then used three separate correlations to evaluate the associations between the participant-wise SvD difference in alpha activity between conditions and each of (1) performer EQ, (2) the SvD difference in average asynchrony, and (3) the SvD difference in asynchrony variability. We did not observe significant associations between the participant-wise SvD difference in alpha activity and (1) performer EQ or (2) the SvD difference in average asynchrony. However, we did find a moderate, although non-significant, negative correlation between the SvD difference in alpha activity and the SvD difference in asynchrony variability, $r(12) = -0.45$, $p = 0.11$. This association indicates that the greater an individual's alpha desynchronization in the "dissimilar" task condition as compared to the "similar" task condition, the greater the variability in their note-onset asynchronies for the "dissimilar" vs. the "similar" condition.

Given the possible association between SvD alpha activity and SvD asynchrony variability we also chose to evaluate the correlation between the SvD difference in asynchrony variability with the computer partner and performer EQ. This yielded a significant negative correlation, $r(12) = -0.59$, $p = 0.03$. This association indicates that individuals with a higher EQ showed greater variability in note-onset asynchronies in the "dissimilar" condition compared to the "similar" condition.

Joining Role

Our analysis of note-onset asynchrony behavior by joining performers revealed a significant difference in average asynchrony between playing with a human vs. computer partner, as well as an interaction between partner animacy and task similarity for asynchrony variability. Like we saw for starting performers, joining performers anticipated the computer partner more frequently than a human partner. They also exhibited greater asynchrony variability with human partners, with this difference being greater in the similar task condition compared to the dissimilar task condition. We calculated the participant-wise human vs. computer difference (HvC difference) in average asynchrony and compared this to (1) the HvC difference in alpha activity and (2) performer EQ. In examining the relationship between average asynchrony and alpha activity we looked at separate electrode groups, as well as activity averaged across electrode groups. None of these correlations revealed significant associations. We also calculated the participant-wise HvC difference in asynchrony variability and compared this variables to the (1) the HvC difference in alpha activity and (2) performer EQ. Neither of these associations were significant.

Our alpha activity analysis for joining performers did not reveal any significant interactions between variables or main effects so we did not examine further correlations with this data.

DISCUSSION

The current study is the first to systematically investigate how co-actor asymmetries act and interact to shape neural and behavioral activity during interpersonal interaction. Our findings indicate that asymmetries in musical role (starting vs. joining), task similarity (similar vs. dissimilar melodic parts), and performer animacy (human-to-human vs. human-to-non-adaptive computer) act to define specific interactive contexts within which performers experience characteristic relationships with their co-actor's behavior. These characteristic relationships are reflected in the temporal dynamics of neural alpha modulation and behavioral coordination that each actor exhibits during a short period of musical unison.

Our observation of increased note-onset asynchrony during performance with a computer partner indicates a general

discrepancy between human and computer timing dynamics. This is likely due to the lack of adaptability inherent to the computer partner in our study (Repp and Keller, 2008) and the natural human inclination to anticipate periodic stimuli (Mates et al., 1994). As noted above, we did not statistically compare neural or behavioral activity between the starting vs. joining role due to the difference in movement onset during the final unison phrase. However, a qualitative review of our findings as presented in **Figure 3** suggests that in the computer condition both starters and joiners generally played ahead of the computer. Again, this likely relates to a human inclination to play faster than the set computer tempo.

In the human condition, joiners typically played just slightly ahead of starters on average. This result is similar to Goebl and Palmer's (2009) observation that follower pianists tended to temporally precede leader pianists in conditions where they could hear the leader's part of a duet as well as their own. This anticipation was small when the leader also heard the follower, and greater when the leader could only hear themselves. The roles in our study were most similar to the full auditory feedback condition in their study, with the exception that starters and joiners in our study alternated playing prior to the final unison phrase. As a result, the starter may have focused more on their own part, creating some similarity between our starter vs. joiner roles and the leader-follower condition in Goebl and Palmer's study in which the leader only received auditory feedback corresponding to their own behavior. With this understanding of the relationship between co-performers, the observed temporal asynchrony is, in fact, consistent with the idea that joiners exhibited a small negative mean asynchrony (NMA) with respect to the starter after the starter played the first three notes of the phrase final unison phrase.

Interestingly, we observed reduced variability of note-onset asynchronies during interaction with the computer partner as compared to the human partner. These results are somewhat contrary to previous findings demonstrating that optimal levels of mutual adaptivity during synchronization lead to reduced variability (Fairhurst et al., 2012). We speculate that this discrepancy is driven by the fact that we evaluated coordination stability over only four key-presses, while the previous work looked at much longer sequences of synchronized taps. Given the dynamic nature of interpersonal interaction, measures of coordination averaged over a long time-window may be substantially different than coordinative patterns in a local segment.

For performers in the joining role we observed a consistent alpha synchronization followed by a desynchronization or return to baseline in alpha-band activity immediately prior to the unison. This finding could indicate that joining performers were preparing to start playing the unison, with alpha ERD reflecting movement preparation (Arroyo et al., 1993; Pfurtscheller et al., 1997; de Jong et al., 2006; Gladwin et al., 2008; Yamanaka and Yamamoto, 2010). At the same time, this may also reflect the joiner attending to their partner's (i.e., starter's) musical activity in order to achieve joint temporal coordination (Jin et al., 2006; Klimesch et al., 2007; Klimesch, 2012; Novembre et al., 2016). The time course of alpha modulation in joining performers was not significantly affected by having a similar vs. dissimilar musical task to one's co-performer. Starting performers were already playing throughout the unison measure and generally displayed small alpha modulations around the unison onset. In some conditions they also displayed distinct alpha desynchronization prior to the unison onset, independent from alpha ERD that would have been associated with the start of their playing at the beginning of the measure.

As we saw for joining performers, for starting performers playing with a human partner there was no effect of musical task similarity on alpha modulation prior to unison. However, in the condition where the co-performer was the computer partner and their parts were musically similar we actually observed ERS, constituting significantly different alpha activity from the condition in which starters had a distinctly different musical part from the computer partner. While this result is somewhat consistent with previous work revealing that individuals exhibit greater alpha ERD during interaction with a human partner vs. a computer partner (Billeke et al., 2014), it also provides a more nuanced view of the effects of role asymmetries on neural processes related to perception and action. Specifically, musical task similarity appeared to moderate the effect of playing with a computer partner but only for performers assuming a starting role in the duet. We speculate that the experience of being a starting performer during interaction with a computer partner playing a similar part is akin to that of performing a solo with a karaoke accompaniment. Among all of the performance conditions created in the current study this situation may invoke the strongest solo mindset for a performer and therefore result in the least attentiveness to co-performer behavior. This view is supported by previous studies which have related ERS in the alpha band to the inhibition of external stimuli or co-actor activity during movement (Klimesch et al., 2007; Babiloni et al., 2012; Klimesch, 2012).

Notably, there was no effect of electrode grouping on starter or joiner alpha modulation, and no interaction between electrode grouping, performer animacy, and task similarity. The occurrence of consistent alpha modulation across left and right parietal and frontal-centro-medial areas is in line with the role of alpha as facilitating long-range communication for domain-general attentional functions. Such attention processes in turn would act to support the temporal coordination of social interactive behaviors in our task. This pattern is in contrast to the right centro-parietal alpha modulation specific to the subtle timing difference between performers in a joint tapping task (Tognoli et al., 2007). Also, the lack of lateralization in the centro-parietal sites here speaks against the relation to movement-related functions because our participants used only the right hand for keyboard playing.

In the current study we did not directly compare starter vs. joiner alpha activity because movement onsets differed between the two conditions. Because Konvalinka et al. (2014) found greater frontal alpha desynchronization in leaders compared to followers, one might expect a similar pattern of difference between our starting and joining conditions separate from the difference due to distinct movement onsets. However, it is important to note that our definition of starting and joining

roles depended on the global music context rather than a local measure of adaptability of own movement interval to the partner's previous interval, employed by Konvalinka et al. (2014), which fluctuated from time to time. Indeed, a visual comparison of our starter and joiner data did not reveal noticeable differences in the magnitude of alpha suppression. Together these findings indicate that the assignment to starter and joiner roles in our study did not have the same effect on co-actor attention as the emergent functional asymmetry at the beat-to-beat level between leaders and followers in the previous work. In other words, it is quite possible that the roles defined by musical turn-taking at the phrase-structure level in our design engage the brain in a different manner than asymmetrical roles identified at a local level. Also, the musical task in our study is considerably more challenging than synchronized finger-tapping and generally required greater attention to co-actor behavior for both starters and joiners.

The effects of partner animacy that we saw on behavioral patterns of note-onset asynchrony magnitude and variability were not associated with any systematic difference in alpha modulation. However, we did detect a moderate, non-significant correlation relating alpha modulation to note-onset asynchrony variability for the starting performer when interacting with the computer partner. Namely, the difference in alpha ERD between the similar and dissimilar musical task conditions was negatively associated with the difference in note-onset asynchrony between the two conditions. This association indicates that individuals who exhibited greater alpha suppression in the "dissimilar" task condition compared to the "similar" task condition also displayed greater variability of note-onset asynchronies in the "dissimilar" vs. "similar" condition. As we have noted this effect was only approaching significance, but it suggests that there may be some connection between attention to co-actor behavior and increased variability, possibly as the result of high levels of adaptation during interpersonal coordination when task similarity between co-actors plays a role.

Interestingly, we further identified a significant correlation between the difference in asynchrony variability in the "similar" and "dissimilar" conditions and performer EQ for starting performers interacting with a computer partner. This association established that individuals with a higher EQ exhibited greater asynchrony variability in the "dissimilar" condition compared to the "similar" condition. In both Pecenka and Keller's (2011) work and our own study, increased perspective-taking or empathy therefore appears to be associated with a greater influence of external stimulus or co-actor activity on an individual's temporal pattern of behavior. During tasks which involve temporal coordination, individuals respond to this influence by continuously acting to adapt their behavior to the ongoing stimulus or co-actor activity. As a result, their behavior becomes more variable than that of an individual who doesn't exhibit the same degree of adaptivity. Thus, EQ is thought to correspond to the level of adaptivity an individual exhibits during interaction, with higher EQ correlating with greater adaptivity.

It is possible that we saw heightened variability specifically within the "dissimilar" condition because performers saw themselves as occupying a more distinct role than their co-performer, thus making their contribution to the ensemble somewhat more tangible. Our alpha ERD findings indicate that in the "similar," computer condition performers were less attentive to the co-performer part, likely because they saw themselves as occupying a solo role and did not need to adapt or be adapted to in order to achieve a successful performance. However, once performers experience their parts as distinct and complementary, they may experience a greater need for adaptation between performers to achieve the joint goal. People with high levels of perspective-taking are more likely to respond to this context with an increase in adaptivity to their co-performer, while those with lower levels are more likely to maintain a more stable pattern of behavior.

Although we did not replicate the direct association between EQ and alpha ERD presented in previous work using a musical performance observation task (Babiloni et al., 2012), our results still suggest that there are strong connections between trait empathy, alpha modulation, and behavioral coordination during interpersonal interaction. Many existing studies have established an association between social interaction and alpha modulation, often suggesting that alpha modulation in MNS regions is involved in sensorimotor processing during social interactions (Babiloni et al., 2012; Sanger et al., 2013; Konvalinka et al., 2014; Novembre et al., 2016). However, to our knowledge, no existing work has illustrated how trait empathy might be related to alpha modulation itself. Our findings suggest that if a person has high trait empathy, perspective-taking may occur more naturally. In turn, this may be associated with more dynamic alpha modulation in response to the individual's current task demands, and, depending on these conditions, more sensitivity in the action-perception of a co-actor's behavior. Our study is also the first to capture the effects of multiple levels of co-performer asymmetries simultaneously in a controlled musical performance task. By always having participants perform an identical musical sequence in the unison phrase based on their starter or joiner role we are able to demonstrate that co-performer asymmetries in the priming context have a measurable impact on subsequent alpha modulation during the unison.

Future work can build on these findings to identify how specific levels of engagement, as indicated by alpha ERD, are dynamically linked to collective performance outcomes which are influenced by certain asymmetries between co-actors. This suggests that asymmetries between co-actors could also be adjusted systematically during real world interaction in order to achieve greater actor engagement. For example, based on our own findings, when people have to interact with computer-generated actors, the design of the computer co-actor should include features of adaptation to human co-actor's behavior in order to support human engagement during interaction, especially if the two actors have similar task roles and the human initiates the task. Alternatively, when people are engaged in situations where certain asymmetries associated with lower engagement are unavoidable, other adjustments and strategies could be employed to increase engagement and improve performance. For

instance, in the context of network ensemble music performance, long audio and video delays may prevent one performer from effectively adapting to their co-actor. The other individual might then deal with this issue by re-structuring the environment in order to remain engaged with their (apparently not so successfully adaptive) co-performer at key points during the interaction. In a musical context, such a solution might be as simple as adding a written reminder to attend to another group member's behavior, or utilizing shared cues to overcome the timing discrepancies at a specific point in music. At a broader level, this could also involve training individuals to remain engaged with co-actor behavior, even when engagement is less likely to occur naturally, as enhancing an individual's engagement is likely to improve collective performance.

Our study includes several noteworthy limitations and also illuminates areas for future work. First, our findings regarding the effect of partner animacy are consistent with previous work evaluating the effects of non-adaptivity during ensemble performance (e.g., Demos et al., 2017). However, we are somewhat limited in drawing conclusions about this effect given that in our human partner condition both co-performers were always able to see each other while in our computer partner condition there was no physically or visually presented co-actor. It is possible that the visual information about co-performer behavior available in the human partner condition afforded smaller average note-onset asynchronies and may have affected alpha band activity as well. Interestingly, however, there is also existing work showing that the absence of visual information during piano duet performance is associated with a higher degree of coupling between performers (Walton et al., 2015). In this case the authors suggest that musicians become more tightly coordinated in order to increase the likelihood of cohesive performance. Future work should aim to establish a deeper understanding of the independent effects of co-performer non-adaptivity and visual information about co-performer behavior. It is worth noting, however, that our key results were observed within each partner condition. This means that even when performers experienced the same visual environment they experienced the duet performance differently depending on the musical similarity between the two parts.

Second, while we demonstrated a relationship between EQ and note-onset asynchrony variability within individual performers, our study was not designed to evaluate the effect of asymmetries between co-performers' persistent, social personality traits or behavioral characteristics. Extensions of our present work could be used to identify associations between pairwise asymmetries in characteristics like empathy and locus of control and (1) co-actor differences in behavioral and neural activity or (2) collective coordination outcomes. Recent work has revealed that individuals with expertise in couple dancing (e.g., Tango, Salsa, Swing) show enhanced neural activity when they perform the role for which they are an expert (i.e., leader vs. follower) (Chauvigné and Brown, 2018). These findings indicate that trait-level asymmetries between co-actors significantly affect each actor's neural and behavioral processes during interpersonal interaction.

Our role asymmetries existed at the musical phrase level context, which is notably shorter-term than the trait level context. However, it is important to remember that asymmetries existing at even shorter timescales are also likely to influence the temporal dynamics of interpersonal coordination. In many social interactions, musical and otherwise, co-actor asymmetries are dynamically varied. As a result, shifts in attention and related patterns of coordination are likely to fluctuate frequently. This occurs in other creative social interactions like dance and acting, especially those which allow for some degree of improvisation, as well team or unit-based scenarios like sporting events and military missions.

The continued, simultaneous use of neural and behavioral measurement techniques will allow future researchers to further investigate the ability of individuals to adapt to changing asymmetries while maintaining coordinated activity. Notably, our focus in this study was on the neural and behavioral activity of two co-performers during interaction and did not include consideration of a larger group of co-actors or the experience of a listener or observer. It would, for example, be valuable to determine whether the kind of causal influences on respiration and heart rate variability Müller and Lindenberger (2011) observed between a conductor and choir members are also related to distinct conductor vs. choir member alpha modulation. Additionally, future work exploring the association between co-performer alpha modulation and listener or observer experience will provide critical insight into how co-actor engagement is facilitated by rapid modulations of neural activity and may shape a third-party audience's perception of collaborative performance.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

This study was approved by and carried out in accordance with the recommendations of the Stanford University Institutional Review Board. All participants gave written informed consent.

AUTHOR CONTRIBUTIONS

All authors designed the study and edited the manuscript. MW created the experimental apparatus and MH and AW verified its setup. TF, IR, MH, NG, TD, WR, and CN collected the data. AW, IR, MH, and TF analyzed the data. AW and TF wrote the manuscript.

FUNDING

This research was supported by the National Science Foundation (1606904). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Science Foundation.

REFERENCES

Ahn, S., Cho, H., Kwon, M., Kim, K., Kwon, H., Kim, B. S., et al. (2018). Interbrain phase synchronization during turn-taking verbal interaction—a hyperscanning study using simultaneous EEG/MEG. *Hum. Brain Mapp.* 39, 171–188. doi: 10.1002/hbm.23834

Arnstein, D., Cui, F., Keysers, C., Maurits, N. M., and Gazzola, V. (2011). μ-suppression during action observation and execution correlates with BOLD in dorsal premotor, inferior parietal, and SI cortices. *J. Neurosci.* 31, 14243–14249. doi: 10.1523/jneurosci.0963-11.2011

Arroyo, S., Lesser, R. P., Gordon, B., Uematsu, S., Jackson, D., and Webber, R. (1993). Functional significance of the mu rhythm of human cortex: an electrophysiologic study with subdural electrodes. *Electroencephalogr. Clin. Neurophysiol.* 87, 76–87. doi: 10.1016/0013-4694(93)90114-b

Babiloni, C., Buffo, P., Vecchio, F., Marzano, N., Del Percio, C., Spada, D., et al. (2012). Brains "in concert": frontal oscillatory alpha rhythms and empathy in professional musicians. *Neuroimage* 60, 105–116. doi: 10.1016/j.neuroimage.2011.12.008

Babiloni, C., Carducci, F., Cincotti, F., Rossini, P. M., Neuper, C., Pfurtscheller, G., et al. (1999). Human movement-related potentials vs desynchronization of EEG alpha rhythm: a high-resolution EEG study. *Neuroimage* 10, 658–665. doi: 10.1006/nimg.1999.0504

Babiloni, C., Del Percio, C., Rossini, P. M., Marzano, N., Iacoboni, M., Infarinato, F., et al. (2009). Judgment of actions in experts: a high-resolution EEG study in elite athletes. *Neuroimage* 45, 512–521. doi: 10.1016/j.neuroimage.2008.11.035

Babiloni, C., Marzano, N., Infarinato, F., Iacoboni, M., Rizza, G., Aschieri, P., et al. (2010). "Neural efficiency" of experts' brain during judgment of actions: a high-resolution EEG study in elite and amateur karate athletes. *Behav. Brain Res.* 207, 466–475. doi: 10.1016/j.bbr.2009.10.034

Baron-Cohen, S., and Wheelwright, S. (2004). The empathy quotient: an investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *J. Autism Dev. Disord.* 34, 163–175. doi: 10.1023/b:jadd.0000022607.19833.00

Billeke, P., Zamorano, F., Chavez, M., Cosmelli, D., and Aboitiz, F. (2014). Functional cortical network in alpha band correlates with social bargaining. *PLoS One* 9:e109829. doi: 10.1371/journal.pone.0109829

Byblow, W. D., Chua, R., and Goodman, D. (1995). Asymmetries in coupling dynamics of perception and action. *J. Motor Behav.* 27, 123–137. doi: 10.1080/00222895.1995.9941705

Cannon, E. N., Yoo, K. H., Vanderwert, R. E., Ferrari, P. F., Woodward, A. L., and Fox, N. A. (2014). Action experience, more than observation, influences mu rhythm desynchronization. *PLoS One* 9:e92002. doi: 10.1371/journal.pone.0092002

Chang, A., Livingstone, S. R., Bosnyak, D. J., and Trainor, L. J. (2017). Body sway reflects leadership in joint music performance. *Proc. Natl. Acad. Sci. U.S.A.* 114, E4134–E4141. doi: 10.1073/pnas.1617657114

Chauvigné, L. A., and Brown, S. (2018). Role-specific brain activations in leaders and followers during joint action. *Front. Hum. Neurosci.* 12:401. doi: 10.3389/fnhum.2018.00401

Cochin, S., Barthelemy, C., Lejeune, B., Roux, S., and Martineau, J. (1998). Perception of motion and qEEG activity in human adults. *Electroencephalogr. Clin. Neurophysiol.* 107, 287–295. doi: 10.1016/s0013-4694(98)00071-6

Cochin, S., Barthelemy, C., Roux, S., and Martineau, J. (2001). Electroencephalographic activity during perception of motion in childhood. *Eur. J. Neurosci.* 13, 1791–1796. doi: 10.1046/j.0953-816x.2001.01544.x

Cui, X., Bryant, D. M., and Reiss, A. L. (2012). NIRS-based hyperscanning reveals increased interpersonal coherence in superior frontal cortex during cooperation. *Neuroimage* 59, 2430–2437. doi: 10.1016/j.neuroimage.2011.09.003

de Jong, R., Gladwin, T. E., and 't Hart, B. M. (2006). Movement-related EEG indices of preparation in task switching and motor control. *Brain Res.* 1105, 73–82. doi: 10.1016/j.brainres.2006.03.030

Demos, A. P., Carter, D. J., Wanderley, M. M., and Palmer, C. (2017). The unresponsive partner: roles of social status, auditory feedback, and animacy in coordination of joint music performance. *Front. Psychol.* 8:149. doi: 10.3389/fpsyg.2017.00149

Dumas, G., Martinerie, J., Soussignan, R., and Nadel, J. (2012). Does the brain know who is at the origin of what in an imitative interaction? *Front. Hum. Neurosci.* 6:128. doi: 10.3389/fnhum.2012.00128

Dumas, G., Nadel, J., Soussignan, R., Martinerie, J., and Garnero, L. (2010). Inter-brain synchronization during social interaction. *PLoS One* 5:e12166. doi: 10.1371/journal.pone.0012166

Engel, A., and Keller, P. E. (2011). The perception of musical spontaneity in improvised and imitated jazz performances. *Front. Psychol.* 2:83. doi: 10.3389/fpsyg.2011.00083

Fairhurst, M. T., Janata, P., and Keller, P. (2012). Being and feeling in sync with an adaptive virtual partner: brain mechanisms underlying dynamic cooperativity. *Cereb. Cortex* 23, 2592–2600. doi: 10.1093/cercor/bhs243

Fairhurst, M. T., Janata, P., and Keller, P. E. (2013). Being and feeling in sync with an adaptive virtual partner: brain mechanisms underlying dynamic cooperativity. *Cereb. Cortex* 23, 2592–2600. doi: 10.1093/cercor/bhs243

Frenkel-Toledo, S., Bentin, S., Perry, A., Liebermann, D. G., and Soroker, N. (2014). Mirror-neuron system recruitment by action observation: effects of focal brain damage on mu suppression. *NeuroImage* 87, 127–137. doi: 10.1016/j.neuroimage.2013.10.019

Gastaut, H. J., and Bert, J. (1954). EEG changes during cinematographic presentation (Moving picture activation of the EEG). *Electroencephalogr. Clin. Neurophysiol.* 6, 433–444. doi: 10.1016/0013-4694(54)90058-9

Giese, M. A., Dijkstra, T. M. H., Schöner, G., and Gielen, C. C. A. M. (1996). Identification of the nonlinear state-space dynamics of the action-perception cycle for visually induced postural sway. *Biol. Cybernet.* 74, 427–437. doi: 10.1007/s004220050254

Gladwin, T. E., 't Hart, B. M., and de Jong, R. (2008). Dissociations between motor-related EEG measures in a cued movement sequence task. *Cortex* 44, 521–536. doi: 10.1016/j.cortex.2007.10.005

Goebl, W., and Palmer, C. (2009). Synchronization of timing and motion among performing musicians. *Music Percept. Interdiscipl. J.* 26, 427–438. doi: 10.1525/mp.2009.26.5.427

Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S., and Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: a neuromagnetic study. *Proc. Natl. Acad. Sci. U.S.A.* 95, 15061–15065. doi: 10.1073/pnas.95.25.15061

Hari, R., and Kujala, M. V. (2009). Brain basis of human social interaction: from concepts to brain imaging. *Physiol. Rev.* 89, 453–479. doi: 10.1152/physrev.00041.2007

Hobson, H. M., and Bishop, D. V. M. (2016). Mu suppression–a good measure of the human mirror neuron system? *Cortex* 82, 290–310. doi: 10.1016/j.cortex.2016.03.019

Holz, E. M., Doppelmayr, M., Klimesch, W., and Sauseng, P. (2008). EEG correlates of action observation in humans. *Brain Topogr.* 21, 93–99. doi: 10.1007/s10548-008-0066-1

Huberth, M., Dauer, T., Nanou, C., Román, I., Gang, N., Reid, W., et al. (2019). Performance monitoring of self and other in a turn-taking piano duet: a dual-EEG study. *Soc. Neurosci.* 14, 449–461. doi: 10.1080/17470919.2018.1492968

Iacoboni, M., and Dapretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nat. Rev. Neurosci.* 7, 942–951. doi: 10.1038/nrn2024

Jeka, J., Oie, K., Schöner, G., Dijkstra, T., and Henson, E. (1998). Position and velocity coupling of postural sway to somatosensory drive. *J. Neurophysiol.* 79, 1661–1674. doi: 10.1152/jn.1998.79.4.1661

Jin, Y., O'Halloran, J. P., Plon, L., Sandman, C. A., and Potkin, S. G. (2006). Alpha EEG predicts visual reaction time. *Int. J. Neurosci.* 116, 1035–1044. doi: 10.1080/00207450600553232

Jurkiewicz, M. T., Gaetz, W. C., Bostan, A. C., and Cheyne, D. (2006). Post-movement beta rebound is generated in motor cortex: evidence from neuromagnetic recordings. *Neuroimage* 32, 1281–1289. doi: 10.1016/j.neuroimage.2006.06.005

Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E. S., Luck, S. J., et al. (2014). Committee report: publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology* 51, 1–21. doi: 10.1111/psyp.12147

Keller, P. E. (2008). "Joint action in music performance," in *Enacting Intersubjectivity: A Cognitive and Social Perspective to the Study of Interactions*, eds F. Morganti, A. Carassa, and G. Riva, (Amsterdam: IOS Press), 205–221.

Kelso, J. A. S., DelColle, J. D., and Schöner, G. (1990). Action-perception as a pattern formation process. *Attent. Perform.* 5, 139–169. doi: 10.1007/s11571-006-9001-x

Klimesch, W. (2012). Alpha-band oscillations, attention, and controlled access to stored information. *Trends Cogn. Sci.* 16, 606–617. doi: 10.1016/j.tics.2012.10.007

Klimesch, W., Doppelmayr, M., Russegger, H., Pachinger, T., and Schwaiger, J. (1998). Induced alpha band power changes in the human EEG and attention. *Neurosci. Lett.* 244, 73–76. doi: 10.1016/s0304-3940(98)00122-0

Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition–timing hypothesis. *Brain Res. Rev.* 53, 63–88. doi: 10.1016/j.brainresrev.2006.06.003

Konvalinka, I., Bauer, M., Stahlhut, C., Hansen, L. K., Roepstorff, A., and Frith, C. D. (2014). Frontal alpha oscillations distinguish leaders from followers: multivariate decoding of mutually interacting brains. *Neuroimage* 94, 79–88. doi: 10.1016/j.neuroimage.2014.03.003

Lachat, F., Hugueville, L., Lemaréchal, J.-D., Conty, L., and George, N. (2012). Oscillatory brain correlates of live joint attention: a dual-EEG study. *Front. Hum. Neurosci.* 6:156. doi: 10.3389/fnhum.2012.00156

Loehr, J. D., and Palmer, C. (2011). Temporal coordination between performing musicians. *Q. J. Exp. Psychol.* 64, 2153–2167. doi: 10.1080/17470218.2011.603427

Martineau, J., and Cochin, S. (2003). Visual perception in children: human, animal and virtual movement activates different cortical areas. *Int. J. Psychophysiol.* 51, 37–44. doi: 10.1016/s0167-8760(03)00151-x

Mates, J., Müller, U., Radil, T., and Pöppel, E. (1994). Temporal integration in sensorimotor synchronization. *J. Cogn. Neurosci.* 6, 332–340. doi: 10.1162/jocn.1994.6.4.332

Ménoret, M., Varnet, L., Fargier, R., Cheylus, A., Curie, A., Des Portes, V., et al. (2014). Neural correlates of non-verbal social interactions: a dual-EEG study. *Neuropsychologia* 55, 85–97. doi: 10.1016/j.neuropsychologia.2013.10.001

Müller, V., and Lindenberger, U. (2011). Cardiac and respiratory patterns synchronize between persons during choir singing. *PLoS One* 6:e24893. doi: 10.1371/journal.pone.0024893

Naeem, M., Prasad, G., Watson, D. R., and Kelso, J. S. (2012a). Electrophysiological signatures of intentional social coordination in the 10–12 Hz range. *Neuroimage* 59, 1795–1803. doi: 10.1016/j.neuroimage.2011.08.010

Naeem, M., Prasad, G., Watson, D. R., and Kelso, J. S. (2012b). Functional dissociation of brain rhythms in social coordination. *Clin. Neurophysiol.* 123, 1789–1797. doi: 10.1016/j.clinph.2012.02.065

Novembre, G., Sammler, D., and Keller, P. (2016). Neural alpha oscillations index the balance between self-other integration and segregation in real-time joint action. *Neuropsychologia* 89, 414–425. doi: 10.1016/j.neuropsychologia.2016.07.027

Oberman, L. M., Pineda, J. A., and Ramachandran, V. S. (2007). The human mirror neuron system: a link between action observation and social skills. *Soc. Cogn. Affect. Neurosci.* 2, 62–66. doi: 10.1093/scan/nsl022

Orgs, G., Dombrowski, J. H., Heil, M., and Jansen-Osmann, P. (2008). Expertise in dance modulates alpha/beta event-related desynchronization during action observation. *Eur. J. Neurosci.* 27, 3380–3384. doi: 10.1111/j.1460-9568.2008.06271.x

Pecenka, N., and Keller, P. E. (2011). The role of temporal prediction abilities in interpersonal sensorimotor synchronization. *Exp. Brain Res.* 211, 505–515. doi: 10.1007/s00221-011-2616-0

Perry, A., and Bentin, S. (2009). Mirror activity in the human brain while observing hand movements: a comparison between EEG desynchronization in the mu-range and previous fMRI results. *Brain Res.* 1282, 126–132. doi: 10.1016/j.brainres.2009.05.059

Pfurtscheller, G. (1992). Event-related synchronization (ERS): an electrophysiological correlate of cortical areas at rest. *Electroencephalogr. Clin. Neurophysiol.* 83, 62–69. doi: 10.1016/0013-4694(92)90133-3

Pfurtscheller, G. (2003). Induced oscillations in the alpha band: functional meaning. *Epilepsia* 44, 2–8. doi: 10.1111/j.0013-9580.2003.12001.x

Pfurtscheller, G., Brunner, C., Schlögl, A., and Da Silva, F. L. (2006). Mu rhythm (de) synchronization and EEG single-trial classification of different motor imagery tasks. *NeuroImage* 31, 153–159. doi: 10.1016/j.neuroimage.2005.12.003

Pfurtscheller, G., Neuper, C., Andrew, C., and Edlinger, G. (1997). Foot and hand area mu rhythms. *Int. J. Psychophysiol.* 26, 121–135. doi: 10.1016/s0167-8760(97)00760-5

Quandt, L. C., and Marshall, P. J. (2014). The effect of action experience on sensorimotor EEG rhythms during action observation. *Neuropsychologia* 56, 401–408. doi: 10.1016/j.neuropsychologia.2014.02.015

Repp, B. H. (2006). Does an auditory distractor sequence affect self-paced tapping? *Acta Psychol.* 12, 81–107. doi: 10.1016/j.actpsy.2005.06.006

Repp, B. H., and Keller, P. (2008). Sensorimotor synchronization with adaptively timed sequences. *Hum. Move. Sci.* 27, 423–456. doi: 10.1016/j.humov.2008.02.016

Repp, B. H., and Penel, A. (2004). Rhythmic movement is attracted more strongly to auditory than to visual rhythms. *Psychol. Res.* 68, 252–270.

Richardson, M. J., Marsh, K. L., and Schmidt, R. C. (2005). Effects of visual and verbal interaction on unintentional interpersonal coordination. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 62–79.4.

Richardson, M. J., Washburn, A., Kallen, R. W., and Harrison, S. J. (2016). "Symmetry and the behavioral dynamics of social coordination," in *Interpersonal Coordination and Performance in Social Systems*, eds P. Passos, and K. Davis, (Abingdon: Routledge), 65–81.

Russell, D. M., and Sternad, D. (2001). Sinusoidal visuomotor tracking: intermittent servo-control or coupled oscillations? *J. Motor Behav.* 33, 329–349. doi: 10.1080/00222890109601918

Salmelin, R., Hämäläinen, M., Kajola, M., and Hari, R. (1995). Functional segregation of movement-related rhythmic activity in the human brain. *Neuroimage* 2, 237–243. doi: 10.1006/nimg.1995.1031

Sanger, J., Muller, V., and Lindenberger, U. (2013). Directionality in hyperbrain networks discriminates between leaders and followers in guitar duets. *Front. Hum. Neurosci.* 7:234. doi: 10.3389/fnhum.2013.00234

Schmidt, R. C., and O'Brien, B. (1997). Evaluating the dynamics of unintended interpersonal coordination. *Ecol. Psychol.* 9, 189–206. doi: 10.1207/s15326969eco0903_2

Schmidt, R. C., Richardson, M. J., Arsenault, C., and Galantucci, B. (2007). Visual tracking and entrainment to an environmental rhythm. *J. Exp. Psychol. Hum. Percept. Perform.* 33, 860–870. doi: 10.1037/0096-1523.33.4.860

Schmidt, R. C., and Turvey, M. T. (1994). Phase-entrainment dynamics of visually coupled rhythmic movements. *Biol. Cybernet.* 70, 369–376. doi: 10.1007/s004220050040

Skewes, J. C., Skewes, L., Michael, J., and Konvalinka, I. (2015). Synchronised and complementary coordination mechanisms in an asymmetric joint aiming task. *Exp. Brain Res.* 233, 551–565. doi: 10.1007/s00221-014-4135-2

Tadel, F., Baillet, S., Mosher, J. C., Pantazis, D., and Leahy, R. M. (2011). Brainstorm: a user-friendly application for MEG/EEG analysis. *Comput. Intell. Neurosci.* 2011, 1–13. doi: 10.1155/2011/879716

Taniguchi, M., Kato, A., Fujita, N., Hirata, M., Tanaka, H., Kihara, T., et al. (2000). Movement-related desynchronization of the cerebral cortex studied with spatially filtered magnetoencephalography. *Neuroimage* 12, 298–306. doi: 10.1006/nimg.2000.0611

Tognoli, E., Lagarde, J., DeGuzman, G. C., and Kelso, J. S. (2007). The phi complex as a neuromarker of human social coordination. *Proc. Natl. Acad. Sci. U.S.A.* 104, 8190–8195. doi: 10.1073/pnas.0611453104

Walton, A., Richardson, M. J., Langland-Hassan, P., Chemero, A., and Washburn, A. (2015). "Musical improvisation: multi-scaled spatiotemporal patterns of coordination," in *Proceedings of the 37th Annual Meeting of the Cognitive Science Society*, eds D. C. Noelle, R. Dale, A. S. Warlaumont, J. Yoshimi, T. Matlock, C. D. Jennings, et al. (Austin, TX: Cognitive Science Society), 2595–2600.

Washburn, A., Coey, C. A., Romero, V., and Richardson, M. J. (2014). Visual multifrequency entrainment: can 1: 2, 2: 3, and 3: 4 coordination occur spontaneously? *J. Motor Behav.* 46, 247–257. doi: 10.1080/00222895.2014.893980

Washburn, A., Kallen, R. W., Coey, C. A., Shockley, K., and Richardson, M. J. (2015). Harmony from chaos? Perceptual-motor delays enhance behavioral anticipation in social interaction. *J. Exp. Psychol. Hum. Percept. Perform.* 41, 1166–1177. doi: 10.1037/xhp0000080

Wilson, A. D., Collins, D. R., and Bingham, G. P. (2005). Perceptual coupling in rhythmic movement coordination: stable perception leads to stable action. *Exp. Brain Res.* 164, 517–528. doi: 10.1007/s00221-005-2272-3

Wimmers, R. H., Beek, P. J., and van Wieringen, P. C. (1992). Phase transitions in rhythmic tracking movements: a case of unilateral coupling. *Hum. Move. Sci.* 11, 217–226. doi: 10.1016/0167-9457(92)90062-g

Wing, A. M., Endo, S., Bradbury, A., and Vorberg, D. (2014). Optimal feedback correction in string quartet synchronization. *J. R. Soc. Interface* 93, 20131125. doi: 10.1098/rsif.2013.1125

Wright, M., Cassidy, R. J., and Zbyszyński, M. F. (2004). "Audio and gesture latency measurements on Linux and OSX," in *Proceedings of the International Computer Music Conference*, (Miami, FL: International Computer Music Association), 423–429.

Yamanaka, K., and Yamamoto, Y. (2010). Lateralised EEG power and phase dynamics related to motor response execution. *Clin. Neurophysiol.* 121, 1711–1718. doi: 10.1016/j.clinph.2010.03.027

Zamm, A., Pfordresher, P. Q., and Palmer, C. (2015). Temporal coordination in joint music performance: effects of endogenous rhythms and auditory feedback. *Exp. Brain Res.* 233, 607–615. doi: 10.1007/s00221-014-4140-5

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.

Copyright © 2019 Washburn, Román, Huberth, Gang, Dauer, Reid, Nanou, Wright and Fujioka. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 15 November 2019
doi: 10.3389/fnins.2019.01220

Tracking the Dynamic Functional Network Interactions During Goal-Directed Auditory Tasks by Brain State Clustering

Gaoyan Zhang, Yuexuan Li and Jinliang Zhang*

College of Intelligence and Computing, Tianjin Key Laboratory of Cognitive Computing and Application, Tianjin University, Tianjin, China

Both perceiving and processing external sound stimuli as well as actively maintaining and updating relevant information (i.e., working memory) are critical for communication and problem solving in everyday acoustic environments. The translation of sensory information into perceptual decisions for goal-directed tasks hinges on dynamic changes in neural activity. However, the underlying brain network dynamics involved in this process are not well specified. In this study, we collected functional MRI data of participants engaging in auditory discrimination and auditory working memory tasks. Independent component analysis (ICA) was performed to extract the brain networks involved and the sliding-window functional connectivity (FC) among networks was calculated. Next, a temporal clustering technique was used to identify the brain states underlying auditory processing. Our results identified seven networks configured into four brain states. The number of brain state transitions was negatively correlated with auditory discrimination performance, and the fractional dwell time of State 2-which included connectivity among the triple high-order cognitive networks and the auditory network (AN)-was positively correlated with working memory performance. A comparison of the two tasks showed significant differences in the connectivity of the frontoparietal, default mode, and sensorimotor networks (SMNs), which is consistent with previous studies of the modulation of task load on brain network interaction. In summary, the dynamic network analysis employed in this study allowed us to isolate moment-to-moment fluctuations in inter-network synchrony, find network configuration in each state, and identify the specific state that enables fast, effective performance during auditory processing. This information is important for understanding the key neural mechanisms underlying goal-directed auditory tasks.

Keywords: dynamic network interaction, functional MRI, goal-directed auditory tasks, independent component analysis, brain state clustering

OPEN ACCESS

Edited by:
Andreas Bahmer, Universitätsklinikum Würzburg, Germany

Reviewed by:
Paola Valsasina, San Raffaele Scientific Institute (IRCCS), Italy
Joan Guàrdia-Olmos, University of Barcelona, Spain

****Correspondence:***
Gaoyan Zhang
zhanggaoyan@tju.edu.cn

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *29 May 2019*
Accepted: *29 October 2019*
Published: *15 November 2019*

Citation:
Zhang G, Li Y and Zhang J (2019) Tracking the Dynamic Functional Network Interactions During Goal-Directed Auditory Tasks by Brain State Clustering. Front. Neurosci. 13:1220. doi: 10.3389/fnins.2019.01220

INTRODUCTION

Both perceiving and processing external sound stimuli as well as actively maintaining and updating relevant information (i.e., working memory) are critical for communication and problem solving in everyday acoustic environments (Huang et al., 2013). The translation of sensory information into perceptual decisions for goal-directed tasks hinges on dynamic changes in neural activity

(Kopell et al., 2014). However, the dynamic changes in brain networks involved in this process are not well specified.

Functional connectivity (FC) is generally used to evaluate interactions in the brain, and it usually refers to the degree of co-variation between spatially distributed signals emanating from the brain. Interactions include FC among different brain regions that constitutes a local brain network and FC among different brain networks that constitute the large-scale brain network. Resting state fMRI measurements have shown that a brain network of auditory modality-specific areas in the temporal lobe participate in auditory processing (Damoiseaux et al., 2006). Task fMRI studies based on different cognitive loads have reported that distinct cortical networks were activated by auditory attention and working memory load (Huang et al., 2013), and FC between the supratemporal plane (STP) and inferior parietal lobule (IPL) in the auditory network (AN) was modulated when discriminating and actively maintaining different pitch-varying sounds (Hakkinen and Rinne, 2018). Another study on auditory word processing based on FC analyses demonstrated that auditory processing recruited the language network (LN), the dorsal attention network (DAN), and the default mode network (DMN). This study also found that intra-network connectivity was stronger in one language than in another (Jung et al., 2018).

Although previous studies have suggested that multiple brain networks are involved in processing auditory goal-directed tasks, it should be noted that these FC studies are commonly conducted based on the hypothesis that FC in the human brain is stable. Correspondingly, the network dynamics during the auditory process are unclear. Recent work has increasingly found that FC is dynamic and evolves in biologically meaningful ways at temporal scales ranging from years to seconds (Gonzalez-Castillo and Bandettini, 2018). At shorter temporal scales, FC patterns computed over tens of seconds contain sufficient information to determine the tasks in which subjects are actively engaged (Shirer et al., 2012; Gonzalez-Castillo et al., 2015). A study that used magnetoencephalographic signals to assess human listeners judging acoustic stimuli composed of carefully titrated clouds of tone sweeps, suggested that global network communication during perceptual decision-making was implemented in the human brain by large-scale couplings between beta-band neural oscillations (Alavash et al., 2017). However, how large-scale functional network interactions change dynamically in the temporal domain and how different cognitive loads modulate dynamic functional network connectivity (FNC) in auditory tasks is still unclear. Further investigation of these unsolved questions is important to improve our understanding of how these processes support goal-directed functioning in everyday acoustic environments.

The recent development of time-resolved analyses of functional neuroimaging data provide a unique opportunity to examine time-varying reconfigurations in global network structure (Shine et al., 2016). Many studies now use independent component analysis (ICA) to extract brain networks and assess dynamic changes in connectivity strength among networks to explore the neural mechanisms underlying development and brain disease (Faghiri et al., 2018). In this study, we used this method to track the dynamic changes in FNC during different auditory tasks. We also assessed the modulation of task load on FNC and its correlation with cognitive behaviors. We believe that this dynamic FNC analysis may reveal detailed information regarding brain dynamics during auditory goal-directed tasks.

MATERIALS AND METHODS

Participants

Twenty college students (mean age: 22.5 years, age range: 20–24, 10 female, right handed) participated in this study. They all had normal hearing, with no history of neurological disorders.

Experiments

The whole experiment included one auditory discrimination run and one auditory working memory run, with a total length of 402s for each run. Both runs started with an 8-s fixation, followed by eight 36-s sound blocks interleaved with eight 12-s resting blocks. The eight task block included four sound-source categories (two living categories of animal sounds and human sounds and two non-living sound-source categories of traffic sounds and tool sounds) intersected with two directions (left and right) (Engel et al., 2009). In each sound block, 12 sound samples (with same category and direction) were randomly presented, and each lasted for 2.5 s with an inter-sample-interval of 0.5 s. In the auditory discrimination task, participants were asked to judge whether the current sound samples were same as the first sound samples in that block (0-back). For the sound blocks in the auditory working memory task, a 2-back paradigm was used; here, participants were instructed to judge whether a current sound sample was same as the one given two samples before.

Data Collection

Imaging data were acquired using a 3.0-T SIEMENS MRI scanner. An eight-channel head coil was used during scanning. Foam pads and earplugs were used for all participants to reduce head motion and scanner noise. To prevent visual input from distracting participants from the auditory task, eyeshades were worn by participants during testing. T2*-weighted images were acquired using a gradient echo-planar imaging (EPI) sequence with the following parameters: repetition time (TR) = 2000 ms, echo time (TE) = 30 ms, voxel size = 3.1 × 3.1 × 4.0 mm^3, matrix size = 64 × 64, slices = 33, slice thickness = 4 mm, slice gap = 0.6 mm. T1-weighted anatomical images were acquired using a three-dimensional magnetization-prepared rapid acquisition gradient echo (3D MPRAGE) sequence with the following parameters: TR = 1900 ms, TE = 2.52 ms, time of inversion (TI) = 1100 ms, voxel size = 1 × 1 × 1 mm^3, matrix size = 256 × 256. Participants perceived auditory stimuli through the earphones of the VisuaStim Digital MRI Compatible fMRI system.

Data Preprocessing

The DPABI toolbox[1] was employed for data preprocessing. For each run, the first four images were removed to minimize magnetic saturation effect. Slice timing and head motion correction were performed for the remaining functional images. The translation and rotation parameters of head motion were less than 2 mm and 2°. We also calculated the framewise displacements using a method reported in a previous study (Jenkinson et al., 2002). The framewise displacements were 0.04 ± 0.01 and 0.04 ± 0.02 for the two runs, demonstrating the head motion across frames was controlled well. Next, structural T1-weighted images were co-registered to the mean functional image, and then normalized to Montreal Neurological Institute (MNI) space using a non-linear registration. EPI data were spatially normalized to MNI space with warping parameters estimated from coregistered, high-resolution T1 images, and voxel size was re-sampled as $3 \times 3 \times 3$ mm^3. The normalized data were then smoothed with a 6-mm full-width half-maximum Gaussian kernel to improve the signal-to-noise ratio. After that, experimental paradigm convolved with the canonical hemodynamic response function was used as a regressor in a general linear model to calculate the brain activation map in each task. Six head motion parameters and their derivatives was used as covariates. For the FC analysis, six head motion parameters and their derivatives, as well as the experimental paradigm convolved with the canonical hemodynamic response function, were regressed out of the smoothed fMRI time series. The residual was used for the task-state FC analysis to exclude the artificial correlation between networks induced by shared activations (Poldrack et al., 2011).

Functional Network Extraction

Functional brain network data were extracted using the group ICA method implemented by the GIFT toolbox[2]. Spatially independent component maps and their respective time series were extracted from the data using the following steps. For each subject, preprocessed data were first reduced to 27 components using principle component analysis. Next, individual data were appended along the time dimension and another principle component analysis was performed for group level dimension reduction, from which 18 components were retained. The number of components was estimated based on the minimum description length criterion. Once this had been performed, the infomax algorithm was applied for ICA; here, the algorithm was run 10 times to reduce the effect of subject order. The results were clustered via ICASSO[3] and the most central solution was used to ensure stability. For all components, the stability index of ICA estimate-clusters was around 1, demonstrating that the result was stable even though the subject order was adjusted. Using the back-reconstruction approach, the spatial maps and time courses for each subject were extracted. After visually checking all components, those with a peak in white matter, ventricles, brain stem, or cerebellum, or those with a spatial map and time course dominated by high frequency fluctuations (likely due to motion or physiological effects) were removed. Fourteen components remained based on brain activation maps of the two tasks (**Figure 1**) in which significant activations were observed in auditory regions, visual regions, sensorimotor regions, cerebellar regions, frontoparietal regions, DMN regions and frontoinsular regions in salience network (SN; $p < 0.001$, corrected for false discovery rate). Previous studies about the brain networks involved into auditory cognitive tasks (Schneiders et al., 2012; Huang et al., 2013; Kumar et al., 2016) were also referred to. Finally, the 14 components were grouped into 7 functional brain networks.

Dynamic FNC Calculation

The whole data processing steps were illustrated in **Figure 2**. For time series data of the 14 selected components, we first performed linear detrending and low pass filtering (0.1 Hz). Next, a sliding-window approach was used to calculate dynamic FNC. A window size of 30s was selected according to previous studies, which suggested that 30s–60s of data can effectively capture dynamic information (Hutchison et al., 2013; Allen et al., 2014; Faghiri et al., 2018). A tapered window was created by convolving a Gaussian with a rectangular function. For each window, a full correlation matrix was calculated. The sliding step was 1 sample, resulting in a total of 178 dynamic FNC matrices.

Brain State Clustering and State Analysis

To examine the reoccurring FNC patterns in the temporal process, we used k-means clustering on all sliding-window FNC matrixes of all subjects by Manhattan distance because L^1 distance is more suitable for calculating similarity of high-dimensional data (Charu et al., 2001). A maximum iteration of 150 was used on the time-varying FNC matrices to cluster brain states. Different number of clusters was calculated from 2 to 10. Through dividing within- by between-cluster distances, four clusters was determined by the elbow criterion of the cluster validity index.

After obtained the four brain states, state transition was defined as the number of times a subject transitioned from one state to another. The time proportion of each subject stayed in each state within the whole task duration was defined as fractional dwell time in that state. Due to the non-normality of the two measures after Kolmogorov-Smirnov test (the number of state transitions: $p = 0.026$, and the fractional dwell time in four states: $p = 0.034$, 0.021, 0.117, and 0.200) with SPSS 22.0 software[4], the number of state transitions and the fractional dwell time of each state were separately compared between the two tasks using permutation test. The permutation test was performed as follows. Mean inter-group difference of each measure was calculated firstly, and then all the values of this measure were randomly reassigned into the two groups for 10,000 times. If less than 5% of mean values of randomized inter-group differences were equal or larger than the mean value of original inter-group differences, the result was seemed as significant ($p < 0.05$). In addition, the spatial strength of each state was also compared between tasks using paired t-tests ($p < 0.05$, corrected for false discovery rate).

[1] http://www.rfmri.org/dpabi

[2] http://mialab.mrn.org/software/gift

[3] http://research.ics.aalto.fi/ica/icasso/

[4] https://www.ibm.com/analytics/spss-statistics-software

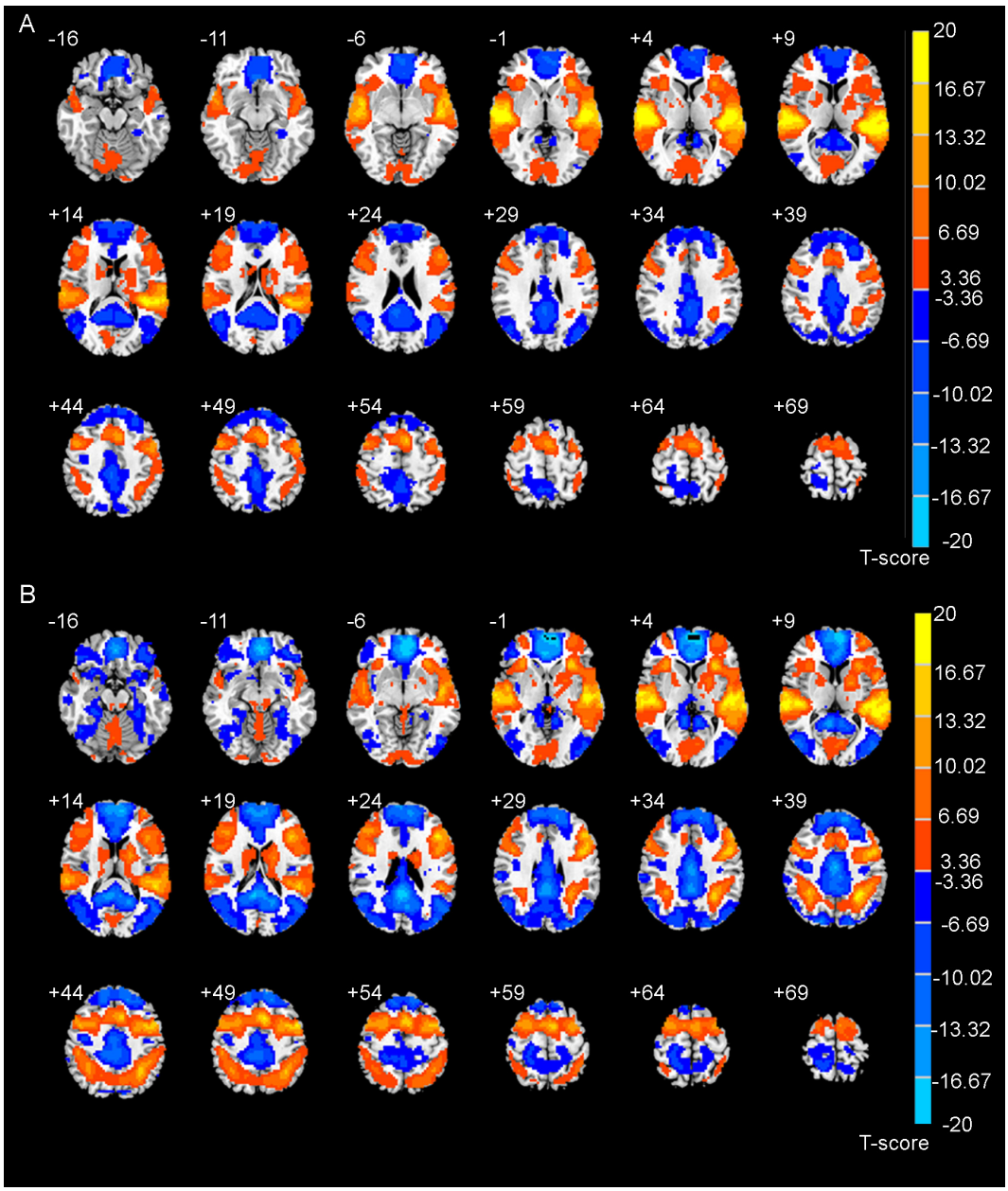

FIGURE 1 | The brain activation maps during auditory discrimination task **(A)** and auditory working memory task **(B)** calculated by general linear model.

Correlation Analysis of Brain State Measure With Behavior

For the auditory discrimination and auditory working memory tasks, dprime scores (Haatveit et al., 2010) were calculated separately to evaluate behavioral performance. Pearson correlations of dprime score with number of state transition and fractional dwell times were conducted to examine whether the dynamic brain network states were related to behavior. The framewise displacement of each subject was used as a covariate in the partial correlation analysis. The significance of the results was tested using fisher t-test ($p < 0.05$).

RESULTS

Spatial Maps of Task-Related Functional Brain Networks

After removing the components related to artifacts, we selected 14 task-related brain network components based on the spatial maps and frequency distribution as mentioned in the method section. The extracted 14 independent components were distributed in 7 functional networks, including the AN, the visual network (VN), the sensorimotor network (SMN), the cerebellar network (CER), the frontoparietal network (FPN), the DMN,

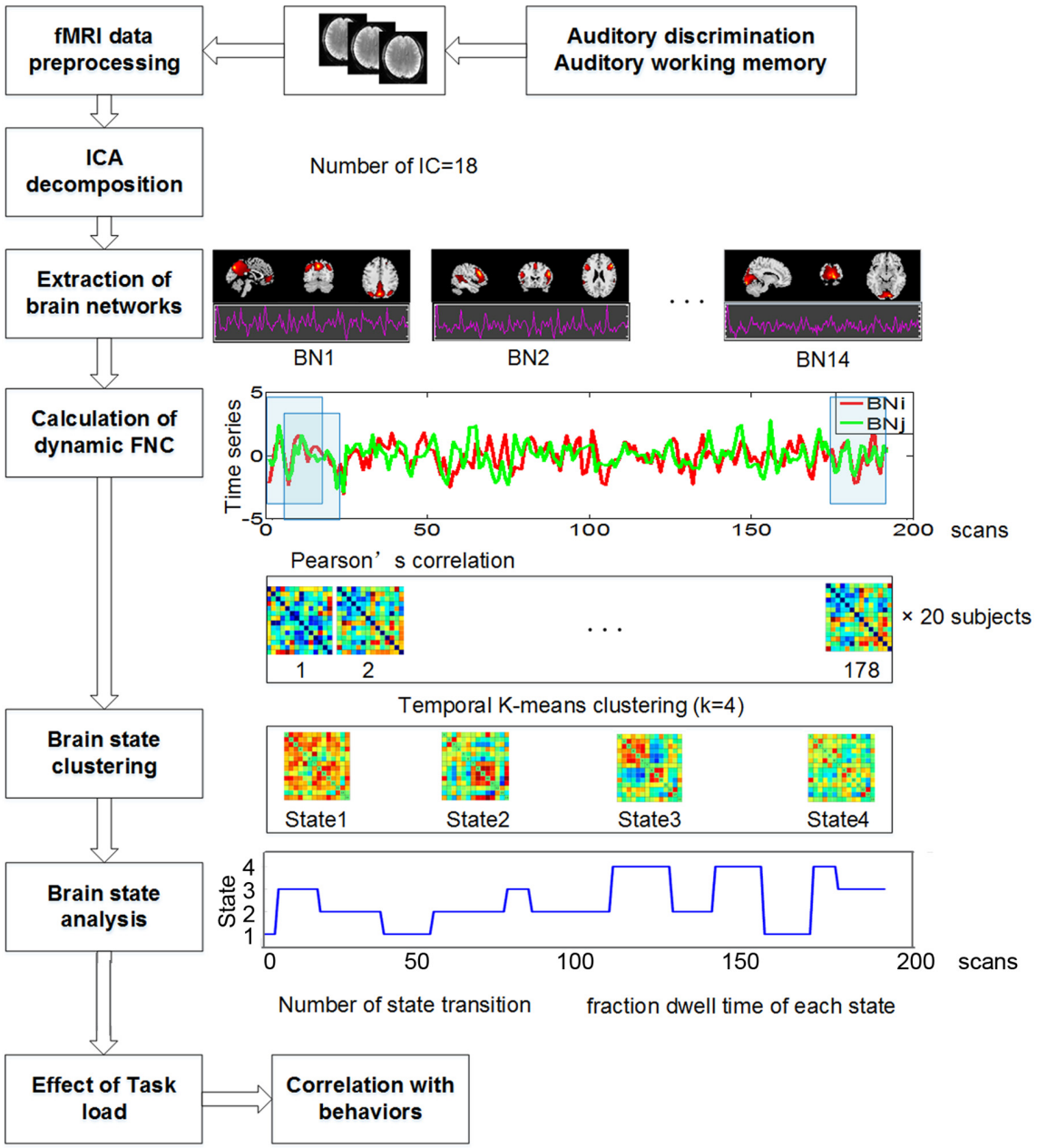

FIGURE 2 | The whole data processing steps.

and the SN (**Figure 3**). The spatial maps of functional networks were displayed using a threshold of z-score > 2.0 and multiple components within one functional network were displayed in a composite plot.

Dynamic Functional Network Connectivity Patterns

The dynamic interactions among the seven functional networks were evaluated using a sliding-window correlation analysis method on the corresponding time series. The 178 dynamic FNC matrices were clustered into 4 brain states. For better visualization, each state was represented by its centroid and is shown in **Figure 4** using a threshold of absolute correlation value r > 0.5 (The original connectivity matrix is shown in **Supplementary Figure S1**).

As shown in **Figure 4**, State 1 consisted of all seven brain networks. In State2, strong connectivity was observed in three higher cognitive networks (FPN, DMN, SN) and two primary networks (AN and VN). In this state, we can see strong inter-network interactions among the three higher cognitive networks as well as between a higher cognitive network (FPN) and

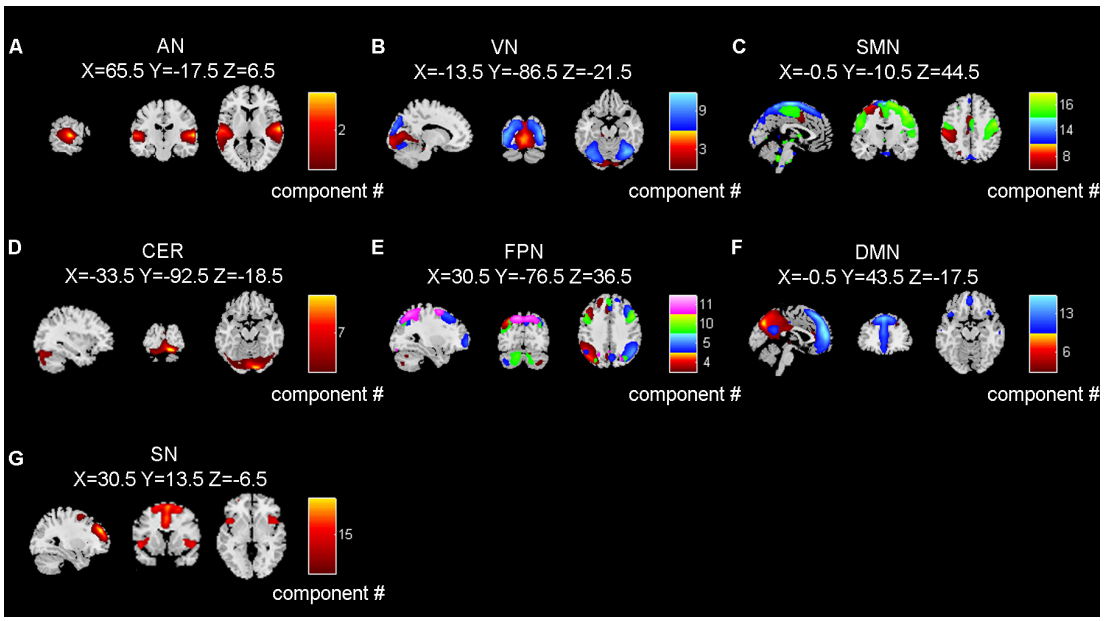

FIGURE 3 | The seven brain networks extracted by independent component analysis. **(A–G)** refer to AN (auditory network), VN (visual network), SMN (sensorimotor network), CER (cerebellum network), FPN (frontoparietal network), DMN (default mode network), and SN (salience network), respectively. Number in color bar means component number. Brain networks in panels **(B,C,E,F)** were composed by multiple independent components.

a primary network (AN). In contrast, strong FC in State 3 and State 4 involved two cognitive networks (DMN, FPN) and two primary networks (SMN, VN or AN). In State 3, we observed strong cross-network interactions between FPN and SMN as well as between SMN and VN. With respect to State 4, only one strong inter-network interaction was found between FPN and AN.

Brain State Analysis Results

The average state transition times and fractional dwell time of each state for the two tasks are listed in **Table 1**. The number of state transitions were similar for both tasks (permutation test, $p = 0.17$). For fractional dwell time, it can be seen that nearly the same percentage of time was spent in the two tasks for States 3 and 4, while the fractional dwell times of State 1 and State 2 showed opposite trends in the auditory discrimination and auditory 2-back tasks. When comparing the fractional dwell times of each state between the two tasks, there was no significant differences (permutation test, $p = 0.33$, 0.20, 0.98, and 0.96, respectively for the four states).

The spatial pattern of each state was also compared between the auditory discrimination task and the auditory working memory task. Significant differences were observed for States 1, 3, and 4, but only differences in State 4 were retained after correcting for multiple comparisons (see **Figure 5**). In State 4, stronger negative connectivity between FPN and DMN and stronger positive connectivity between DMN and SMN were found for the working memory task, while stronger positive connectivity within FPN was found for the auditory discrimination task ($p < 0.05$, FDR corrected).

Correlation Results of Brain State Measure With Behavior

Significant correlation of dprime score with the number of state transition and with fractional dwell time in different states for the two tasks were separately reported in **Figure 6**. It can be seen clearly that a negative correlation ($r = -0.590$, $p = 0.004$, **Figure 6A**) between dprime scores and the number of state transitions in the auditory discrimination task. For the auditory working memory task, there was a positive correlation ($r = 0.577$, $p = 0.005$, **Figure 6B**) between dprime scores and the fractional dwell time in State 2. No significant correlation was observed between dprime scores and the fractional dwell time in other States.

DISCUSSION

In this study, we used data-driven ICA method to extract functional brain networks, and then a temporal clustering analysis on the sliding-window FNC to reveal the time-variable FNC pattern (i.e., brain state) during two goal-directed auditory tasks. This approach allowed us to isolate moment-to-moment fluctuations in inter-network synchrony, which were related to behavioral variability during auditory discrimination and working memory tasks. The findings in this study also reveal the modulation of cognitive demands on the connectivity of time-variable functional networks. Altogether, this study provides a new perspective on time-sensitive shifts in brain network interactions, and this may help us understand the key neural mechanisms underlying goal-directed auditory tasks.

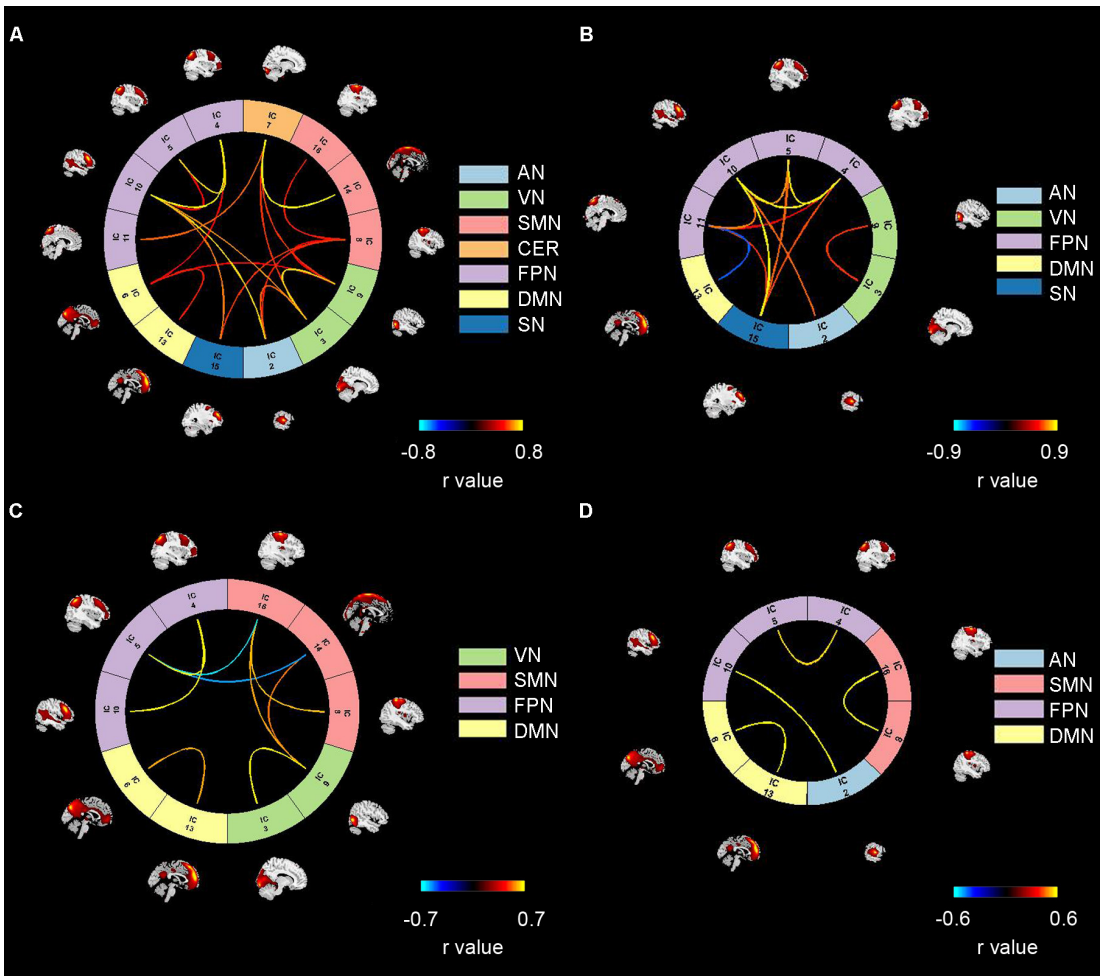

FIGURE 4 | Four reoccurring brain states during auditory processing revealed by clustering analysis. Panels **(A–D)** refer to State 1, 2, 3, 4. Color bars refer to the connectivity strength. AN: auditory network; VN: visual network; SMN: sensorimotor network; CER: cerebellum network; FPN: frontoparietal network; DMN: default mode network; SN: salience network.

TABLE 1 | The average state transition times and fractional time in each state for the two tasks.

	No. of state transition	Fractional dwell time in each state			
		State 1	State 2	State 3	State 4
Auditory discrimination	10.15 ± 0.93	0.26 ± 0.05	0.15 ± 0.03	0.22 ± 0.05	0.37 ± 0.05
Auditory 2-back	11.7 ± 0.64	0.19 ± 0.04	0.21 ± 0.04	0.22 ± 0.03	0.37 ± 0.04

In this study, seven brain networks were found to be configured into four states. State 1 included all networks. States 2, 3, and 4 mainly included strong connectivity in FPN, DMN, AN, VN, and SMN, but the interaction patterns of these networks differed in each state, demonstrating that these states may contribute to different cognitive processing. The network configuration is consistent with previous auditory cognitive studies (Schneiders et al., 2012; Huang et al., 2013; Kumar et al., 2016). For example, using auditory near perception threshold (NT) paradigms, researchers observed that a stronger integration of the auditory network with the frontoparietal and other high-order cognitive networks was key for subsequent auditory performance (Leske et al., 2015). In another study, researchers investigated the brain system for actively maintaining sound memory over short periods of time (Kumar et al., 2016). Their results supported the hypothesis that a system maintained sound-specific representations in the auditory cortex by projecting from higher-order areas, including the hippocampus and frontal cortex. Another recent study documented that the activation of the auditory cortex and adjacent regions in the IPL were strongly modulated during active listening and depended on task requirements (Wikman and Rinne, 2018). In contrast

to these studies, we investigated network interactions from a dynamic perspective and found that brain States 2 and 4 mainly showed a strong interaction between FPN and AN, suggesting that these two states may contribute to sound maintenance and active listening. State 3 contained strong interactions in both FPN-SMN and SMN-VN connectivity. The involvement of SMN, which is important in motor output (De Luca et al., 2005), implies that these interactions may contribute to cognitive decision and button-press. Further studies with high temporal resolution technique are needed to verify the inferences.

Moreover, a comparison between the auditory working memory and discrimination task datasets showed significant differences in State 4, demonstrating that State 4 is an indicator of cognitive load. The load-related increases in connectivity among cognitive (FPN and DMN) and SMNs are coincident with the finding of increased task-driven connectivity between the frontoparietal, dorsal attention, and sensory networks by a previous study (Shine et al., 2016). Moreover, the increased negative correlation of FPN-DMN connectivity in working memory task is consistent with a previous finding (Schneiders et al., 2012), and further reveal the cognitive resources demanded for sound maintenance in this state. These results also suggest that global integration may have facilitated communication during the more challenging working memory task.

Interestingly, by using dynamic network analysis in this study, shifts among brain networks can be measured and the brain-behavior relationship showed that the number of brain state transitions was negatively correlated with auditory discrimination performance, meaning that fewer state transitions contribute to better behavioral performance, but this is not the case for the auditory working memory task. In this study, we also found that the fractional dwell time in State 2 was positively correlated with auditory working memory behavior. In State 2, the triple networks (i.e., FPN, DMN, and SN) and the typical FPN-DMN anticorrelation were most prominent. The triple networks have been suggested as the most crucial components of a unified network model and are thought to be extensively involved in diverse cognitive functions (Menon, 2011; Li et al., 2018). A strong competitive relationship between FPN and DMN was previously reported to be significantly correlated with working memory behavior (Hampson et al., 2010) and may represent a cerebral mechanism that switches mental focus between internal channels (supported by DMN) and external, attention-demanding events (Hampson et al., 2010). Moreover,

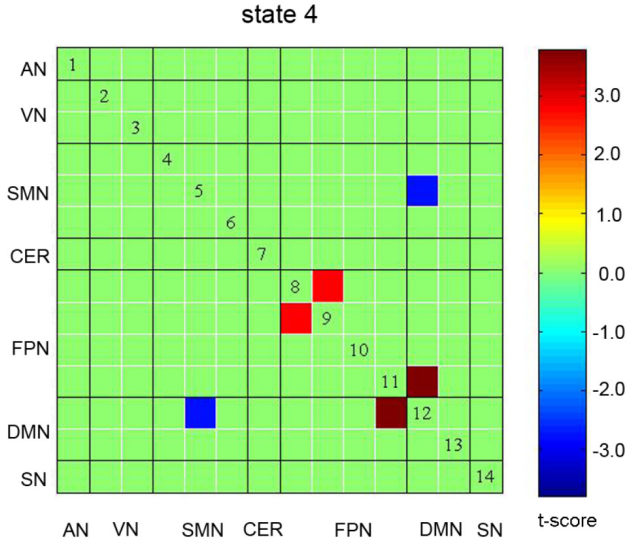

FIGURE 5 | Significant differences in spatial pattern of State 4 in the comparison of auditory discrimination with auditory working memory tasks. No significant differences were observed in other States. AN: auditory network; VN: visual network; SMN: sensorimotor network; CER: cerebellum network; FPN: frontoparietal network; DMN: default mode network; SN: salience network.

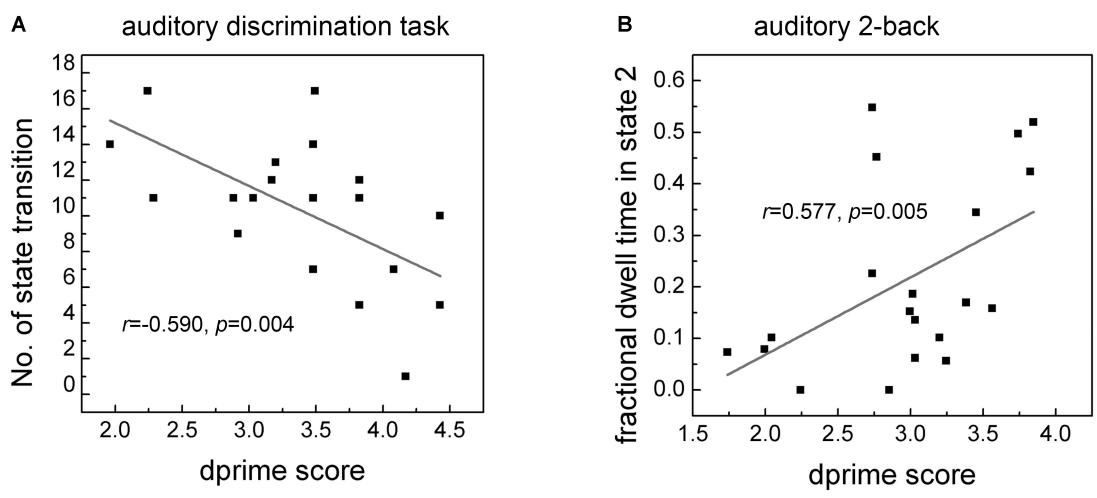

FIGURE 6 | Correlation of dprime score with the number of state transition **(A)** and fractional dwell time in corresponding state 2 **(B)** in the two tasks. Only fractional dwell time in State 2 showed significant correlation with dprime score of auditory 2-back task.

the connectivity between the high-order cognitive network (FPN) and the primary network (AN) was also found to be significant in this state. Taken in the light of these previous studies, our findings suggest that State 2 probably contributes to top-down attention switching, cognitive processing, and behavioral modulation processing (Zhang et al., 2015; Le Merre et al., 2018), all of which are necessary for good performance. Therefore, more time spent in this state should correspond to better auditory memory performance.

CONCLUSION

The human brain network traverses segregated and integrated states over time (Shine et al., 2016). The dynamic FNC analysis used in this study can help identify network configurations of each state, as well as the specific states that enable fast, effective performance on goal-directed auditory tasks. Other approaches are also used in investigating dynamic brain connectivity analysis (Calhoun and Adali, 2016) and a dynamic approach can also be found with a stable non-dynamic model. Future studies can further explore the effectiveness of different methods. In summary, building on the results of previous auditory cognitive studies, the dynamic functional network analysis in this study enrich our understanding of the neural mechanisms underlying auditory discrimination and working memory.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

This study was approved by the Research Ethics Committee of Tianjin Key Lab of Cognitive Computing and Application, Tianjin University. All procedures followed were in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, and the applicable revisions at the time of the investigation. Informed consent was obtained from all subjects for being included in the study.

AUTHOR CONTRIBUTIONS

GZ designed the experiment, and wrote and revised the manuscript. JZ collected the data. GZ and YL contributed to the data analysis.

FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 61876126 and 61503278), Peiyang Scholar Program of Tianjin University (No. 2018XRG-0037), and China Scholarship Council (No. 201706255035).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2019.01220/full#supplementary-material

REFERENCES

Alavash, M., Daube, C., Wostmann, M., Brandmeyer, A., and Obleser, J. (2017). Large-scale network dynamics of beta-band oscillations underlie auditory perceptual decision-making. *Netw. Neurosci.* 1, 166–191. doi: 10.1162/NETN_a_00009

Allen, E. A., Damaraju, E., Plis, S. M., Erhardt, E. B., Eichele, T., and Calhoun, V. D. (2014). Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex* 24, 663–676. doi: 10.1093/cercor/bhs352

Calhoun, V. D., and Adali, T. (2016). Time-varying brain connectivity in fMRI data whole-brain data-driven approaches for capturing and characterizing dynamic states. *IEEE Signal Process. Mag.* 33, 52–66. doi: 10.1109/msp.2015.2478915

Charu, C. C., Hinneburg, A., and Keim, D. A. (2001). "On the surprising behavior of distance metrics in high dimensional space," in *Proceedings of the 8th International Conference of Database Theory, (ICDT 2001)*, (London: Springer).

Damoiseaux, J., Rombouts, S., Barkhof, F., Scheltens, P., Stam, C., Smith, S. M., et al. (2006). Consistent resting-state networks across healthy subjects. *Proc. Natl. Acad. Sci. U.S.A.* 103, 13848–13853. doi: 10.1073/pnas.0601417103

De Luca, M., Smith, S., De Stefano, N., Federico, A., and Matthews, P. M. (2005). Blood oxygenation level dependent contrast resting state networks are relevant to functional activity in the neocortical sensorimotor system. *Exp. Brain Res.* 167, 587–594. doi: 10.1007/s00221-005-0059-1

Engel, L. R., Frum, C., Puce, A., Walker, N. A., and Lewis, J. W. (2009). Different categories of living and non-living sound-sources activate distinct cortical networks. *Neuroimage* 47, 1778–1791. doi: 10.1016/j.neuroimage.2009.05.041

Faghiri, A., Stephen, J. M., Wang, Y. P., Wilson, T. W., and Calhoun, V. D. (2018). Changing brain connectivity dynamics: from early childhood to adulthood. *Hum. Brain Mapp.* 39, 1108–1117. doi: 10.1002/hbm.23896

Gonzalez-Castillo, J., and Bandettini, P. A. (2018). Task-based dynamic functional connectivity: recent findings, and open questions. *Neuroimage* 180(Pt B), 526–533. doi: 10.1016/j.neuroimage.2017.08.006

Gonzalez-Castillo, J., Hoy, C. W., Handwerker, D. A., Robinson, M. E., Buchanan, L. C., Saad, Z. S., et al. (2015). Tracking ongoing cognition in individuals using brief, whole-brain functional connectivity patterns. *Proc. Natl. Acad. Sci. U.S.A.* 112, 8762–8767. doi: 10.1073/pnas.1501242112

Haatveit, B. C., Sundet, K., Hugdahl, K., Ueland, T., Melle, I., and Andreassen, O. A. (2010). The validity of d prime as a working memory index: results from the "Bergen n-back task". *J. Clin. Exp. Neuropsychol.* 32, 871–880. doi: 10.1080/13803391003596421

Hakkinen, S., and Rinne, T. (2018). Intrinsic, stimulus-driven and task-dependent connectivity in human auditory cortex. *Brain Struct. Funct.* 223, 2113–2127. doi: 10.1007/s00429-018-1612-6

Hampson, M., Driesen, N., Roth, J. K., Gore, J. C., and Constable, R. T. (2010). Functional connectivity between task-positive and task-negative brain areas and its relation to working memory performance. *Magn. Reson. Imaging* 28, 1051–1057. doi: 10.1016/j.mri.2010.03.021

Huang, S., Seidman, L. J., Rossi, S., and Ahveninen, J. (2013). Distinct cortical networks activated by auditory attention, and working memory load. *Neuroimage* 83, 1098–1108. doi: 10.1016/j.neuroimage.2013.07.074

Hutchison, R. M., Womelsdorf, T., Allen, E. A., Bandettini, P. A., Calhoun, V. D., Corbetta, M., et al. (2013). Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage* 80, 360–378. doi: 10.1016/j.neuroimage.2013.05.079

Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration, and motion correction of brain images. *NeuroImage* 17, 825–841. doi: 10.1016/s1053-8119(02)91132-8

Jung, J., Nam, K., Cho, H., and Kim, S. (2018). Mapping the neural dynamics of korean-english bilinguals with medium proficiency during auditory word processing. *Front. Psychol.* 9:983. doi: 10.3389/fpsyg.2018.00983

Kopell, N. J., Gritton, H. J., Whittington, M. A., and Kramer, M. A. (2014). Beyond the connectome: the dynome. *Neuron* 83, 1319–1328. doi: 10.1016/j.neuron.2014.08.016

Kumar, S., Joseph, S., Gander, P. E., Barascud, N., Halpern, A. R., and Griffiths, T. D. (2016). A brain system for auditory working memory. *J. Neurosci.* 36, 4492–4505. doi: 10.1523/JNEUROSCI.4341-14.2016

Le Merre, P., Esmaeili, V., Charriere, E., Galan, K., Salin, P. A., Petersen, C. C. H., et al. (2018). Reward-based learning drives rapid sensory signals in medial prefrontal cortex and dorsal hippocampus necessary for goal-directed behavior. *Neuron* 97, 83.e5–91.e5. doi: 10.1016/j.neuron.2017.11.031

Leske, S., Ruhnau, P., Frey, J., Lithari, C., Muller, N., Hartmann, T., et al. (2015). Prestimulus network integration of auditory cortex predisposes near-threshold perception independently of local excitability. *Cereb. Cortex* 25, 4898–4907. doi: 10.1093/cercor/bhv212

Li, R., Zhang, S., Yin, S., Ren, W., He, R., and Li, J. (2018). The fronto-insular cortex causally mediates the default-mode and central-executive networks to contribute to individual cognitive performance in healthy elderly. *Hum. Brain Mapp.* 39, 4302–4311. doi: 10.1002/hbm.24247

Menon, V. (2011). Large-scale brain networks and psychopathology: a unifying triple network model. *Trends Cogn. Sci.* 15, 483–506. doi: 10.1016/j.tics.2011.08.003

Poldrack, R. A., Mumford, J. A., and Nichols, T. E. (2011). *Handbook of Functional MRI Data Analysis*. Cambridge: Cambridge University Press.

Schneiders, J. A., Opitz, B., Tang, H., Deng, Y., Xie, C., Li, H., et al. (2012). The impact of auditory working memory training on the fronto-parietal working memory network. *Front. Hum. Neurosci.* 6:173. doi: 10.3389/fnhum.2012.00173

Shine, J. M., Bissett, P. G., Bell, P. T., Koyejo, O., Balsters, J. H., Gorgolewski, K. J., et al. (2016). The dynamics of functional brain networks: integrated network states during cognitive task performance. *Neuron* 92, 544–554. doi: 10.1016/j.neuron.2016.09.018

Shirer, W. R., Ryali, S., Rykhlevskaia, E., Menon, V., and Greicius, M. D. (2012). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cereb. Cortex* 22, 158–165. doi: 10.1093/cercor/bhr099

Wikman, P., and Rinne, T. (2018). Interaction of the effects associated with auditory-motor integration and attention-engaging listening tasks. *Neuropsychologia* 124, 322–336. doi: 10.1016/j.neuropsychologia.2018.11.006

Zhang, G., Yao, L., Shen, J., Yang, Y., and Zhao, X. (2015). Reorganization of functional brain networks mediates the improvement of cognitive performance following real-time neurofeedback training of working memory. *Hum. Brain Mapp.* 36, 1705–1715. doi: 10.1002/hbm.22731

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.

Copyright © 2019 Zhang, Li and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 19 November 2019
doi: 10.3389/fnins.2019.01241

Temporal Discrimination Thresholds and Proprioceptive Performance: Impact of Age and Nerve Conduction

*Thorsten M. Odorfer, Teresa Wind and Daniel Zeller**

Department of Neurology, University of Würzburg, Würzburg, Germany

OPEN ACCESS

Edited by:
Daya Shankar Gupta,
Camden County College,
United States

Reviewed by:
Lorenzo Rocchi,
University College London,
United Kingdom
Roger David Adams,
University of Canberra, Australia
Bülent Cengiz,
Gazi University, Turkey

***Correspondence:**
Daniel Zeller
zeller_d@ukw.de

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *19 August 2019*
Accepted: *04 November 2019*
Published: *19 November 2019*

Citation:
Odorfer TM, Wind T and Zeller D (2019) Temporal Discrimination Thresholds and Proprioceptive Performance: Impact of Age and Nerve Conduction. Front. Neurosci. 13:1241. doi: 10.3389/fnins.2019.01241

Background: Increasing attention is payed to the contribution of somatosensory processing in motor control. In particular, temporal somatosensory discrimination has been found to be altered differentially in common movement disorders. To date, there have only been speculations as to how impaired temporal discrimination and clinical motor signs may relate to each other. Prior to disentangling this relationship, potential confounders of temporal discrimination, in particular age and peripheral nerve conduction, should be assessed, and a quantifiable measure of proprioceptive performance should be established.

Objective: To assess the influence of age and polyneuropathy (PNP) on somatosensory temporal discrimination threshold (STDT), temporal discrimination movement threshold (TDMT), and behavioral measures of proprioception of upper and lower limbs.

Methods: STDT and TDMT were assessed in 79 subjects (54 healthy, 25 with PNP; age 30–79 years). STDT was tested with surface electrodes over the thenar or dorsal foot region. TDMT was probed with needle electrodes in flexor carpi radialis (FCR) and tibialis anterior (TA) muscle. Goniometer-based devices were used to assess limb proprioception during (i) active pointing to LED markers, (ii) active movements in response to variable visual cues, and (iii) estimation of limb position following passive movements. Pointing (or estimation) error was taken as a measure of proprioceptive performance.

Results: In healthy subjects, higher age was associated with higher STDT and TDMT at upper and lower extremities, while age did not correlate with proprioceptive performance. Patients with PNP showed higher STDT and TDMT values and decreased proprioceptive performance in active pointing tasks compared to matched healthy subjects. As an additional finding, there was a significant correlation between performance in active pointing tasks and temporal discrimination thresholds.

Conclusion: Given their notable impact on measures of temporal discrimination, age and peripheral nerve conduction need to be accounted for if STDT and TDMT are applied in patients with movement disorders. As a side observation, the correlation between measures of proprioception and temporal discrimination may prompt further studies on the presumptive link between these two domains.

Keywords: pointing task, position estimation, aging, kinesthesia, somatosensory temporal discrimination, TDMT, temporal discrimination threshold

BACKGROUND

Temporal aspects of somatosensory processing have drawn increasing interest as potential markers in the differential diagnostic workup of movement disorders. In particular, STDT and TDMT have been shown to be differentially involved. STDT is a neurophysiological paradigm testing the shortest ISI at which a subject can perceive successive electrical stimuli applied to the skin as separate. Higher STDTs have been consistently described in several types of dystonia (Tinazzi et al., 1999, 2004; Bara-Jimenez et al., 2000; Aglioti et al., 2003; Fiorio et al., 2003, 2008) as well as in PD (Artieda et al., 1992; Conte et al., 2010, 2016) and multiple system atrophy (Rocchi et al., 2013). TDMT is defined as the shortest interval at which a subject perceives two externally induced passive movements as separate (Tinazzi et al., 2005). Compared to healthy controls, TDMT has been shown to be increased in PD patients (Fiorio et al., 2007) and patients with essential or functional tremor (Tinazzi et al., 2013a, 2014), whereas it was found normal in patients with writer's cramp (Tinazzi et al., 2006) and dystonia with tremor (Tinazzi et al., 2013a).

To date, there have only been speculations as to how an impairment of temporal discrimination performance and clinical motor signs in these movement disorders may relate to each other (Riemann and Lephart, 2002a,b; Lee et al., 2016), warranting the need for further research. However, prior to disentangling this relationship on the CNS level, we consider it reasonable to examine how temporal discrimination is influenced by age and peripheral nerve conduction – factors which are likely to confound analyses in groups of movement disorder patients with high inter-group heterogeneity. Moreover, a quantifiable measure of proprioceptive performance should be established in order to assess potential associations between temporal discrimination and kinesthesia later on.

To this end, we assessed STDT and TDMT of upper and lower extremities in healthy subjects of different age and in patients with PNP, combined with a set of three proprioception tasks of the corresponding limbs. We hypothesized that higher age and PNP are associated with (i) increased discrimination thresholds and (ii) decreased proprioceptive performance.

Abbreviations: CNS, central nervous system; EMG, electromyography; FCR, flexor carpi radialis muscle; fMRI, functional magnetic resonance imaging; GABA, gamma-aminobutyric acid; ISI, interstimulus interval; LED, light-emitting diode; PD, Parkinson's disease; PNP, polyneuropathy; SSEP, somatosensory evoked potential; STDT, somatosensory temporal discrimination threshold; TA, tibialis anterior muscle; TDMT, temporal discrimination movement threshold.

MATERIALS AND METHODS

The study conformed to the principles of the Declaration of Helsinki. It was approved by the Ethics committee of the Medical Faculty at the University of Würzburg.

Subjects

A total of 54 volunteers without a history of neurological or psychiatric disease and without clinical symptoms or signs of such disease were included. Additional exclusion criteria were a medical history of diabetes or coagulation disorders, and ongoing medication with oral anticoagulant drugs. In addition, 25 patients with a diagnosis of chronic PNP were included. Nerve conduction studies and SSEPs were collected in order to characterize PNP patients, and to exclude impairment of peripheral nerve or posterior column conduction in healthy subjects older than 60 years. Conduction studies were performed at our Clinical Neurophysiology Laboratory (Schwarzer Topas EMG System, Natus Europe, Planegg, Germany) according to the clinical standard.

All participants gave their written informed consent for research.

STDT and TDMT

Somatosensory temporal discrimination threshold was tested with surface electrodes (anode and cathode with 1 mm diameter and 1.5 cm distance in between) placed over the thenar or dorsal foot region. Pairs of square wave electric stimuli with a duration of 0.2 ms were provided by a constant current stimulator (Digitimer, Welwyn Garden City, United Kingdom). Stimulation intensity was determined individually by providing stimuli with stepwise increasing current until participants were able to perceive stimuli clearly (i.e., 10 out of 10 attempts). ISIs were presented in an ascending sequence, starting from 0 ms, in steps of 5 ms. STDT was defined as the shortest ISI when participants perceived two separated pulses in three successive intervals (Tinazzi et al., 1999, 2013a). The mean value of three runs was taken for further analysis.

TDMT was measured following the procedure described by Tinazzi et al. (2005). An insulated tungsten needle microelectrode was inserted at the motor point of the FCR or the TA muscle. The motor point was determined as the cathode position with maximum muscle contraction at stimulation by a surface electrode. The anode was a surface electrode placed 3–4 cm distally to the cathode. Pairs of subsequent electric stimuli (0.2 ms duration, 1–2 mA intensity, below individual stimulus intensity

of STDT testing in all cases) with increasing and decreasing ISI (2 runs each) were provided. TDMT was defined as the shortest ISI at which subjects were able to clearly (i.e., three times in a row) identify two separate movements of wrist flexion or foot dorsal extension. A movement consisted of a distinct perceptible muscle contraction without feeling pain or discomfort, along with an observable slight wrist flexion or foot dorsal extension, respectively (Tinazzi et al., 2005). To minimize possible distraction by external stimuli participants wore earplugs and sleep masks.

The mean value of four runs was taken for further analysis.

Proprioceptive Testing

Proprioception of limbs was assessed by custom-made goniometers without visual feedback of the respective extremity throughout the testing procedure (**Figure 1**). The device for the upper extremity was built to assess wrist flexion in a range of 0°–75° (**Figure 1A**), the one for the lower extremity measured foot dorsiflexion in a range of 0°–60° (**Figure 1B**). Position 0° marks the starting point of motion in our experiment, which corresponds to 15° hand extension and 30° plantar flexion relating to neutral zero method. The goniometers allowed the investigator to monitor the movements directly and to quantify their extent on a scale placed outside the box, invisible for the participant. The respective extremity was fixed with splints and tapes in order to exclude other joint movements.

Subjects were instructed to perform three different proprioceptive tasks as precisely as possible, and to initiate limb movements only on explicit request by the investigator. The first task (LED) comprised pointing to LED markers attached to the outside of the device at 15°, 30°, 45°, 60°, and 75° for the upper limbs (**Figure 1A**), and 15°, 30°, 45°, and 60° for the right foot (**Figure 1B**). The participants were instructed to point to the one lighted LED (e.g., 45° in **Figure 1A**) by one single and quick movement. The examiner documented the reached position, and the subject was asked to return to the starting position. Thereafter, another LED was activated, and the procedure started from its beginning. Each LED lighted up three times during the experiment in a randomized order. In the second task (ARROW), participants were asked to move the respective limb in proportion to curved arrows of different angular dimensions (15°, 30°, or 45°), which were shown on a computer screen (**Figure 1C**). Thereby, subjects were instructed to imitate the range of movement symbolized by the arrow length. Each length was presented three times in a randomized order. The basic testing procedure was otherwise similar to LED task. In the third task (PASSIVE), which was only performed by the upper limbs, subjects were asked to estimate the position of their limb after passive movements by the experimenter (right side 9°, 21°, 31°, 49°, 63° and left side 13°, 25°, 43°, 51°, 65°, each angle presented twice, one after another). Only for this task, an additional scale was installed at the front of the device so that the participants were able to indicate the felt position of the index finger by telling the corresponding number on the scale. Pointing (LED, ARROW) or estimation (PASSIVE) errors (in degree) were taken as a measure of proprioceptive performance. Conduction of the entire assessment took an average of 30 min.

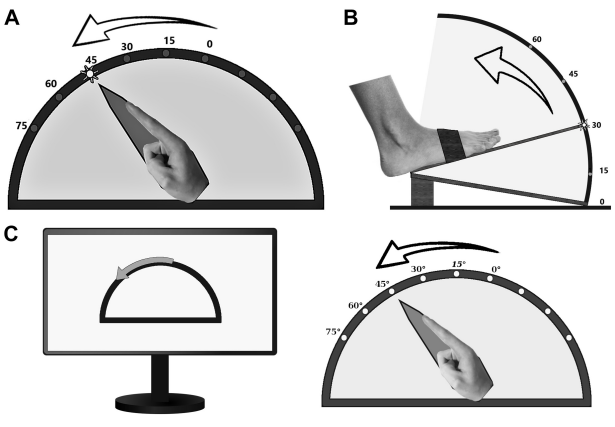

FIGURE 1 | Custom-made goniometers to quantify pointing or estimation errors of **(A)** the upper and **(B)** the lower limb. **(C)** ARROW task: example of a computer screen instruction (left part) and the corresponding active movement of the upper limb.

Statistical Analyses

SPSS software (IBM) was used for statistical analyses. We tested for normality by using the Shapiro–Wilk test. As data were not normally distributed, we applied the Mann–Whitney *U* test for group comparisons and the Spearman test for correlations. Statistical significance was set at a level of $p < 0.05$. The Benjamini–Hochberg procedure was used to correct for multiple comparisons.

RESULTS

Demographic and Clinical Data

A total of 54 healthy subjects (37 females) with a median age of 54 (range 30–76) years were included into this study. In addition, 25 patients (10 females) diagnosed with PNP with a median age of 61 (range 46–79) years were included. Demographic and clinical data of PNP patients and a subgroup of age- and sex-matched healthy controls are summarized in **Supplementary Table 1**.

STDT and TDMT: Association With Age and PNP

The results of STDT and TDMT assessment are presented in the upper part of **Table 1**.

In the group of healthy controls, higher age was associated with higher discrimination threshold levels for STDT of the upper extremities ($r = 0.348$; $p < 0.001$; **Figure 2A**) and the foot ($r = 0.581$; $p < 0.001$; **Figure 2B**). Moreover, higher age was associated with higher TDMT levels of the FCR ($r = 0.267$; $p = 0.005$; **Figure 3A**) and the TA muscle ($r = 0.465$; $p < 0.001$; **Figure 3B**; all significant after Benjamini–Hochberg adjustment).

Compared to matched controls, patients with PNP showed elevated STDT and TDMT values, with higher thresholds for the upper and lower limbs (**Table 1**).

TABLE 1 | Results of proprioception and temporal discrimination tasks.

Condition	Controls (entire cohort)		Controls (age and sex matched)		PNP patients		Level of significance Mann–Whitney *U* test
	Mean	SD	Mean	SD	Mean	SD	*p*-value
STDT right hand (ms)	81.4	17.8	81.4	20.4	121.8	34.3	<0.001*
STDT left hand (ms)§	79.6	17.7					
STDT right foot (ms)	109.5	20.3	114.9	20.2	157.8	39.6	0.001*
TDMT right hand (ms)	81.3	16.6	83.3	17.1	126.3	33.4	<0.001*
TDMT left hand (ms)§	82.2	15.3					
TDMT right foot (ms)	102.9	17.0	106.8	17.9	155.2	35.1	<0.001*
LED right hand (°)	3.2	1.9	3.0	1.8	5.9	2.8	<0.001*
LED left hand (°)§	3.6	2.1					
LED right foot (°)	2.7	1.6	2.6	1.4	6.5	2.5	<0.001*
ARROW right hand (°)	3.2	2.7	2.9	2.9	5.8	2.7	<0.001*
ARROW left hand (°)§	3.2	1.6					
ARROW right foot (°)	2.6	1.4	2.3	1.2	8.4	4.1	<0.001*
PASSIVE right hand (°)	2.2	1.1	2.5	1.3	2.2	1.6	0.132
PASSIVE left hand (°)§	1.7	1.0					

§ *Task not performed in PNP group. *Significant after Benjamini–Hochberg adjustment. STDT: somatosensory temporal discrimination threshold, TDMT: temporal motor discrimination threshold, PNP: polyneuropathy, SD: standard deviation.*

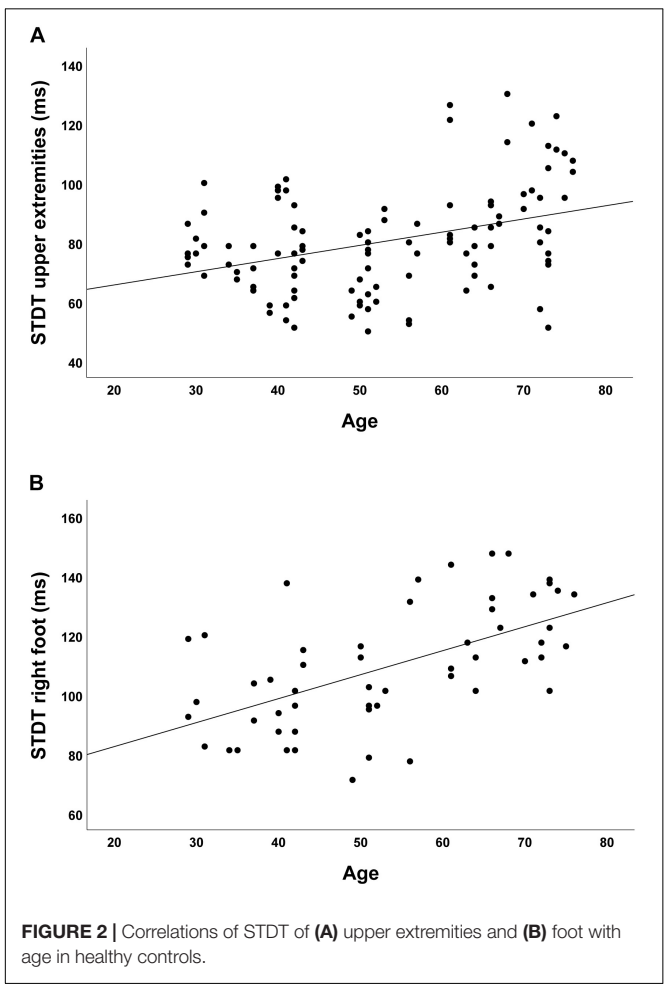

FIGURE 2 | Correlations of STDT of **(A)** upper extremities and **(B)** foot with age in healthy controls.

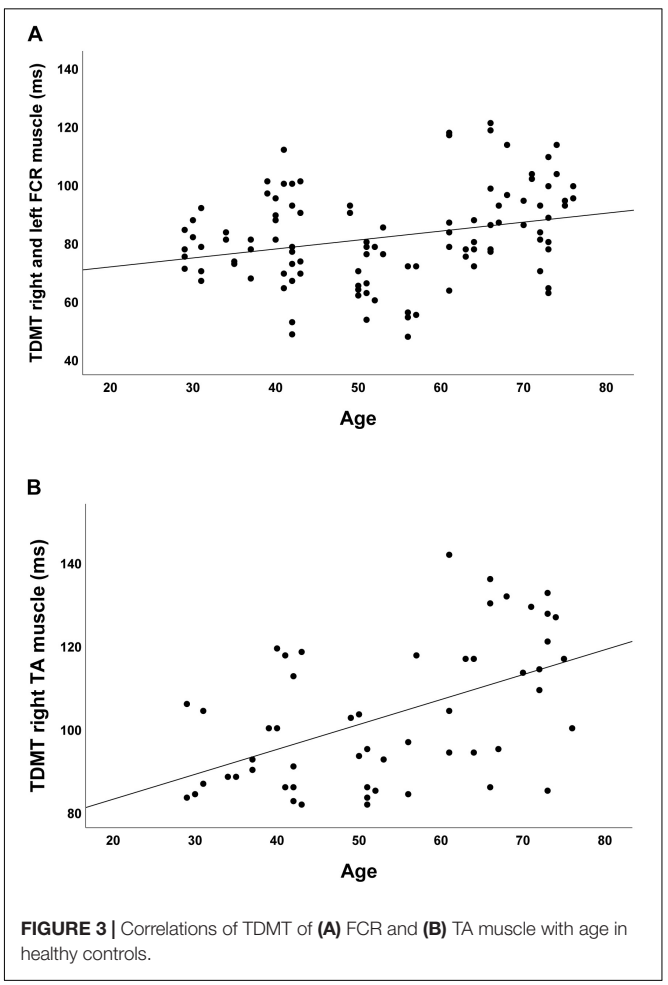

FIGURE 3 | Correlations of TDMT of **(A)** FCR and **(B)** TA muscle with age in healthy controls.

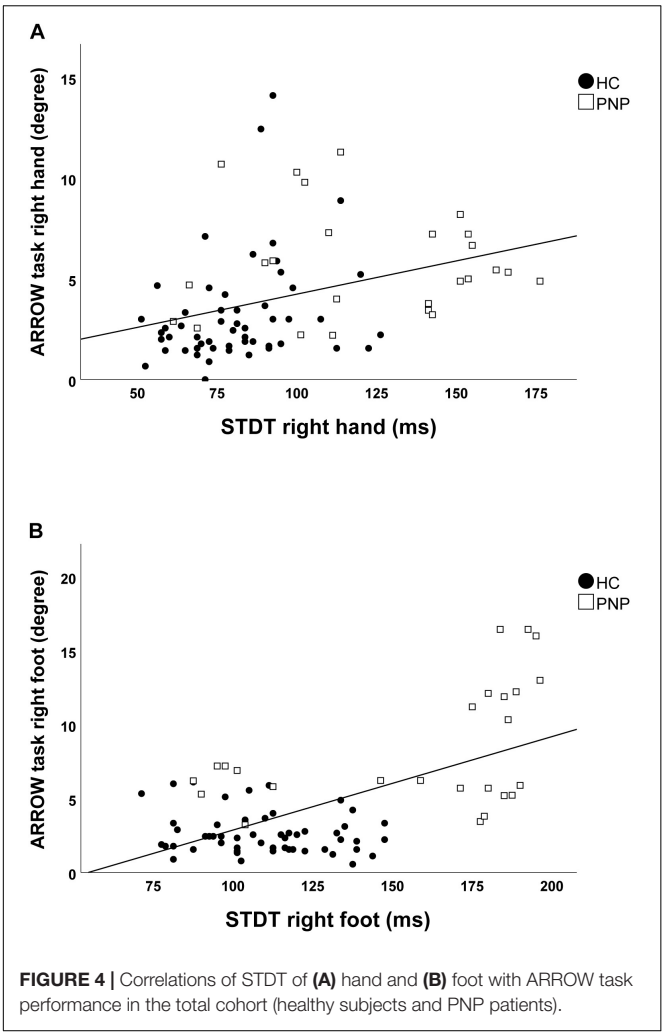

FIGURE 4 | Correlations of STDT of **(A)** hand and **(B)** foot with ARROW task performance in the total cohort (healthy subjects and PNP patients).

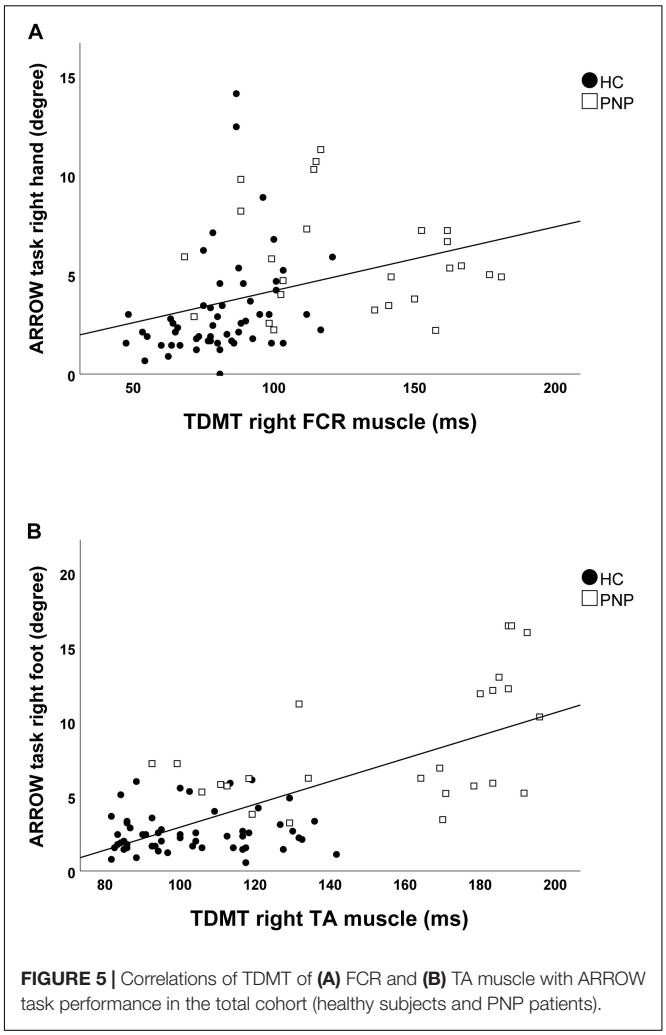

FIGURE 5 | Correlations of TDMT of **(A)** FCR and **(B)** TA muscle with ARROW task performance in the total cohort (healthy subjects and PNP patients).

Kinesthesia: Association With Age and PNP

The results of the proprioceptive testing procedures (LED, ARROW, PASSIVE) are presented in the lower part of **Table 1**.

In the whole group of healthy controls there was no significant correlation between age and higher pointing errors in all performed tasks.

Compared to the group of matched healthy participants, patients with PNP performed worse in the pointing tasks (LED, ARROW), with higher pointing errors for the upper and lower limbs. In the PASSIVE condition, performance was comparable between patients and controls (**Table 1**).

STDT, TDMT, and Kinesthesia

Screening the entire data for a potential correlation between the two domains, we found higher temporal discrimination thresholds to be associated with higher pointing errors in the ARROW task at upper (STDT right hand: $r = 0.477$; $p < 0.001$; **Figure 4A**/TDMT right FCR muscle: $r = 0.546$; $p < 0.001$; **Figure 5A**) and lower limbs (STDT right foot: $r = 0.336$; $p = 0.002$; **Figure 4B**/TDMT right TA muscle: $r = 0.523$; $p < 0.001$; **Figure 5B**). Comparable correlations were found for the LED pointing task: STDT and TDMT correlated significantly with pointing errors (STDT right hand: $r = 0.369$; $p = 0.001$; **Figure 6A**/TDMT right FCR muscle: $r = 0.435$; $p < 0.001$; **Figure 7A**/STDT right foot: $r = 0.372$; $p = 0.001$; **Figure 6B**/TDMT right TA muscle: $r = 0.460$; $p < 0.001$; **Figure 7B**). All these correlations were significant after Benjamini–Hochberg procedure.

In contrast, there was no significant correlation of temporal discrimination thresholds with performance in the PASSIVE estimation task.

DISCUSSION

In the present study we assessed temporal discrimination thresholds (STDT and TDMT) and proprioceptive performance in a group of healthy controls and PNP patients. Our main goal was to determine the influence of age and peripheral nerve conduction on both measures as groundwork for their use in patients with movement disorders. In line with our first hypothesis, higher age and PNP were associated with increased

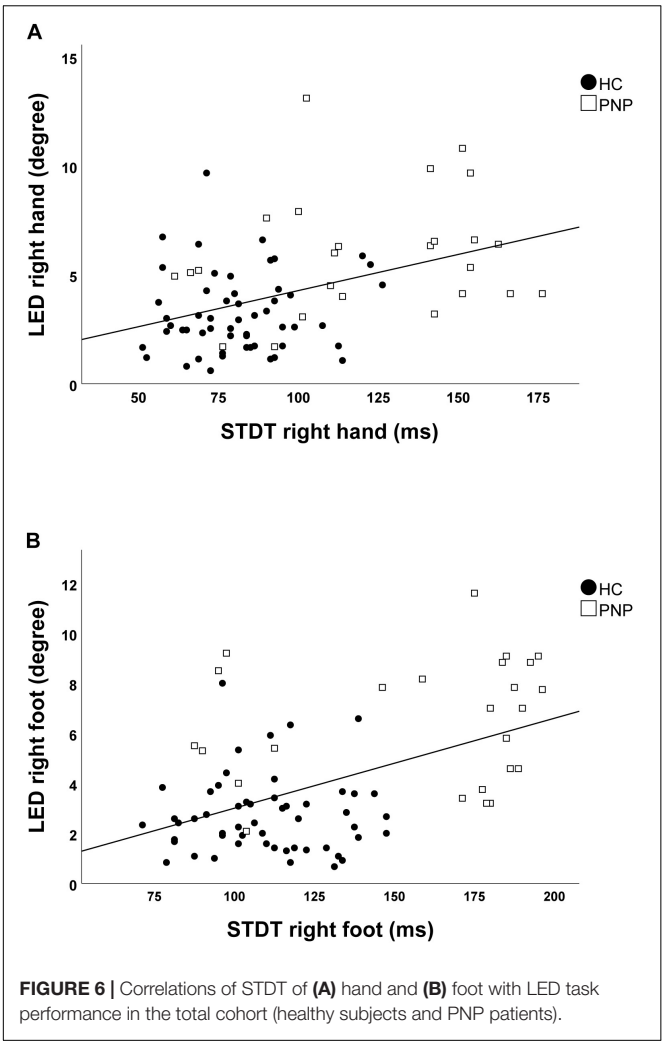

FIGURE 6 | Correlations of STDT of **(A)** hand and **(B)** foot with LED task performance in the total cohort (healthy subjects and PNP patients).

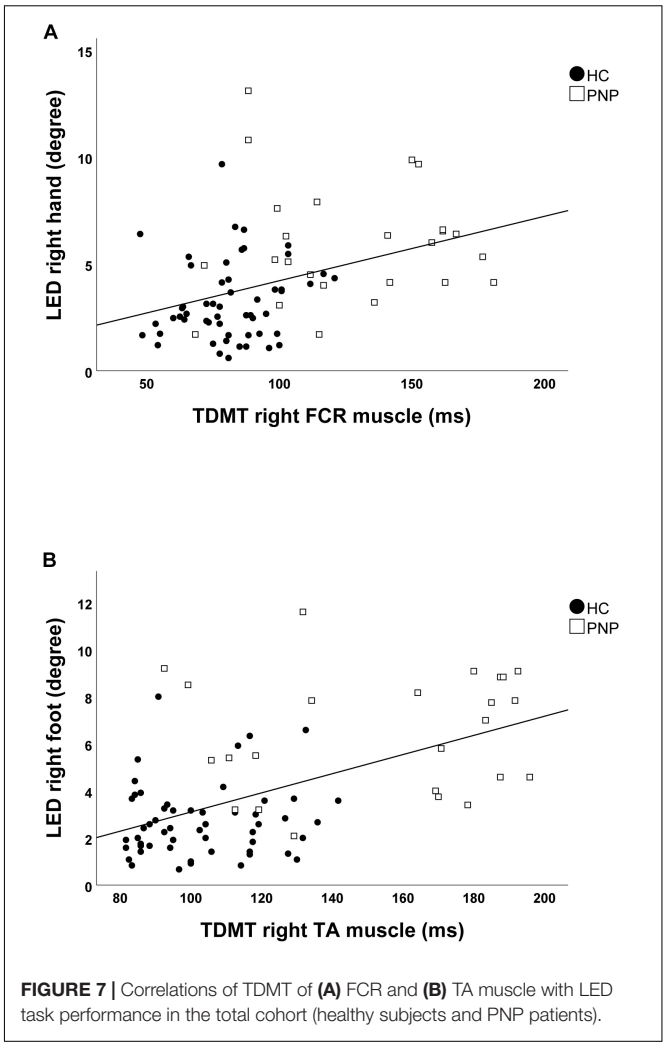

FIGURE 7 | Correlations of TDMT of **(A)** FCR and **(B)** TA muscle with LED task performance in the total cohort (healthy subjects and PNP patients).

discrimination thresholds. Our second hypothesis, in contrast, was only partially confirmed: PNP, but not age, was associated with decreased proprioceptive performance.

Across healthy subjects of different age groups, we found higher age to be correlated with increased temporal discrimination values. In line with a previous study (Ramos et al., 2016), we found positive correlations of age and STDT for the upper limbs, and extended this finding to the lower limbs. Beyond STDT, we found a significant association of higher age with higher TDMT levels of the FCR and an even stronger correlation with TDMT of the TA muscle. Indeed, to the best of our knowledge, this is the first study to demonstrate feasibility of the TDMT paradigm at lower limbs. Together, our data strongly supports the hypothesis that both STDT and TDMT performances decrease during aging. As PNP had been excluded clinically and neurophysiologically in all elderly controls, age-dependent decline in TDT and TDMT may most likely be attributed to changes within central circuits, with several potential reasons (Ramos et al., 2016): Beside a decrease of GABA-ergic neurotransmission in aging (Mora et al., 2008; Lehmann et al., 2012), a reduction of neuronal plasticity within the somatosensory cortex is supposed to play a crucial role in this process (Pellicciari et al., 2009). The latter may be supported by recent findings of an STDT improvement after high frequency repetitive sensory stimulation in healthy subjects which was less prominent in older participants (Erro et al., 2016). The contribution of (subclinical) neurocognitive deficits, which are more frequent in higher age, in temporal perception remains uncertain (Gibbon et al., 1984; Allman and Meck, 2012).

Our finding of significantly increased STDT and TDMT at the upper and lower limbs in patients with PNP points to a considerable impact of peripheral nerve conduction on temporal discrimination. Indeed, blurred signal conduction due to temporal dispersion within the peripheral nerve might well explain our observation of poor temporal discrimination in patients with PNP. As a consequence, peripheral conditions with high general prevalence, like PNP, carpal tunnel syndrome, or radicular compression, might inevitably limit the individual validity of temporal discrimination thresholds in subjects with movement disorders and additional peripheral conditions.

As a potential limitation, the group of PNP patients was rather heterogenous with respect to type and severity of neuropathy.

While an increase of STDT and TDMT in demyelinating PNP might be more suggestive than in axonal PNP or small fiber neuropathy, the respective impact of different types of nerve damage remains speculative. Moreover, nerve conduction studies predominantly reveal information about the distal part of the nerve, whereas proximal demyelination may be missed easily. Consequently, the attempt to disentangle the causal relationship between a particular finding in nerve conduction studies and temporal discrimination performance would be an overinterpretation of our data. This is even more so since the CNS would not remain unaffected by peripheral neuropathy – plastic processes like cortical rearrangement following PNP-related partial denervation might also affect temporal discrimination on the CNS level.

Notably, we did not find a *differential* association of STDT and TDMT with age or PNP: both parameters changed concordantly, i.e., they were positively correlated with age, and both were similarly increased in PNP patients as compared to healthy subjects. Thus, physiological aging as well as PNP influence STDT and TDMT similarly. This is in line with the observation that specific patterns of altered temporal discrimination in disorders like essential tremor or dystonia can be attributed to specific changes at the level of the CNS. Several brain areas have been suggested by fMRI to be involved in temporal discrimination, first of all the basal ganglia (Pastor et al., 2004), in particular the putamen (Kimmich et al., 2014), and the superior colliculus (Hutchinson et al., 2014; Mc Govern et al., 2017). In addition, involvement of the prefrontal cortex, anterior cingulate, pre-supplementary motor area, precentral gyrus, sensorimotor cortex, inferior parietal lobule, and cerebellum has been observed (Pastor et al., 2004; Di Biasio et al., 2015; Rocchi et al., 2017; Erro et al., 2018).

We did not find a significant correlation between age and pointing or estimation errors, respectively. Earlier studies have demonstrated an age-related decline of passive finger (Ko et al., 2015; Zhang et al., 2015; Ingemanson et al., 2016; Rinderknecht et al., 2017) and ankle (Ko et al., 2015) proprioception, while age had no effect on the accuracy rate in a pointing task under restricted visual feedback conditions (Zhang et al., 2015). In the light of the different aspects of proprioception assessed at different joints in these studies, our findings would be compatible with an age-dependent decline in passive, but not active proprioceptive performance, which might be restricted to distal joints and therefore, not appear at the wrist. In general, it is still under debate whether passive and active proprioception are different neuronal concepts and potentially involve different central circuits (Gritsenko et al., 2007; Capaday et al., 2013). In addition, we cannot exclude the possibility that our PASSIVE task at the wrist might have lacked the precision to detect very small angular differences, for example due to the fact that the movement was performed by the experimenter rather than an electric device or to limited control for cues from the skin touching the goniometer splint.

Patients with PNP showed an impaired kinesthetic performance as evidenced by higher pointing errors compared to controls. In the light of reduced peripheral afferent input due to PNP, this finding is not surprising (Rothwell et al., 1982; Sainburg et al., 1995). As for the PASSIVE task, we consider it likely that the low sensitivity of an assessment at the wrist (compared to finger joints) in addition to the methodological limitations discussed above might explain the lack of significant differences between PNP patients and matched controls. As our foot goniometer did not allow movements with gravity eliminated, the PASSIVE condition was only performed at the upper extremities, where, from clinical experience, proprioceptive dysfunction due to PNP is less prominent.

As an additional observation, we found correlations between STDT and TDMT and the performance in two different active pointing tasks: Higher temporal discrimination thresholds were associated with higher pointing errors in the ARROW and LED task in upper and lower extremities. Though postulated by earlier studies, which had found differential alterations of STDT and TDMT in patients with movement disorders (Artieda et al., 1992; Bara-Jimenez et al., 2000; Tinazzi et al., 2006; Fiorio et al., 2007), the association of both parameters with behavioral measures of proprioception had never been systematically assessed.

Active pointing movements rely on permanent "on-line" adjustments of their temporospatial properties in order to achieve a high level of precision (Georgopoulos, 2002). Temporal discrimination is supposed to involve a complex network process within the CNS subserving conscious evaluation of double stimuli, i.e., to answer the question whether a subject feels a single or a double stimulus. It therefore, seems worth speculating on whether temporal discrimination thresholds and pointing precision may be considered different measures of one and the same network. In this case, their correlation would support the notion that kinesthesia is indeed the link between impaired temporal discrimination and neurological conditions in several movement disorders (Tinazzi et al., 2005, 2013b; Fiorio et al., 2007).

As correlations do not allow directional or causal inference, an alternative explanation might be that both temporal discrimination and pointing performance are modulated comparably by a common third parameter. This could either be an unidentified parameter or simply peripheral nerve conduction with PNP-related variability. In the latter case we would expect a lack of correlation when PNP patients are excluded. However, there remained to be a significant correlation between STDT and TDMT and ARROW task in upper limbs in the group of healthy subjects, which supports an association of the two parameters above their common modulation by peripheral conduction.

Functional imaging studies on proprioception during either vibration induced illusory motion or passive extremity movements revealed involvement of a number of brain regions partially overlapping with those discussed in the context of temporal discrimination (Weiller et al., 1996; Gelnar et al., 1998; Mima et al., 1999; Francis et al., 2000; Romaiguere et al., 2003; Naito et al., 2007; Kavounoudias et al., 2008). It is beyond the scope of our study to speculate about the neural underpinnings of both phenomena in the CNS. However, future research should specifically assess proprioceptive performance by means of a "top-down based" task in the particular movement disorders to probe for behavioral correlates of differential STDT and TDMT changes. Functional imaging and non-invasive brain

stimulation might complement such studies in order to identify brain structures involved in both processes.

CONCLUSION

Age and PNP have significant impact on measures of temporal discrimination and/or proprioceptive capacity. STDT and TDMT increase with age, and PNP is associated with higher STDT/TDMT values and reduced precision in pointing tasks. If applied in studies on movement disorder, where STDT/TDMT may be used in order to define corresponding endophenotypes, it is important to account for these factors to increase validity of the measurements. This is particularly important in view of a significantly higher prevalence of PNP in patients with PD as compared to controls (Conradt et al., 2018). As an additional observation of high interest, higher error rates in pointing tasks correlate with elevated discrimination thresholds. This may prompt further studies on the presumptive link between these two domains and their potential use as endophenotypic markers in neurological conditions.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by Ethik-Kommission der Universität Würzburg. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

TO and DZ designed the study concept and wrote the manuscript with the help of TW. TW and TO carried out the experiments and performed the statistical data analysis.

FUNDING

This study was supported by Departmental Research Funds. The publication was supported by the German Research Foundation (DFG) and the University of Würzburg in the funding programme Open Access Publishing.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2019.01241/full#supplementary-material

REFERENCES

Aglioti, S. M., Fiorio, M., Forster, B., and Tinazzi, M. (2003). Temporal discrimination of cross-modal and unimodal stimuli in generalized dystonia. *Neurology* 60, 782–785. doi: 10.1212/01.wnl.0000046528.24693.5b

Allman, M. J., and Meck, W. H. (2012). Pathophysiological distortions in time perception and timed performance. *Brain* 135, 656–677. doi: 10.1093/brain/awr210

Artieda, J., Pastor, M. A., Lacruz, F., and Obeso, J. A. (1992). Temporal discrimination is abnormal in Parkinson's disease. *Brain* 115(Pt 1), 199–210. doi: 10.1093/brain/115.1.199

Bara-Jimenez, W., Shelton, P., Sanger, T. D., and Hallett, M. (2000). Sensory discrimination capabilities in patients with focal hand dystonia. *Ann. Neurol.* 47, 377–380. doi: 10.1002/1531-8249(200003)47:3<377::aid-ana16>3.3.co;2-u

Capaday, C., Darling, W. G., Stanek, K., and Van Vreeswijk, C. (2013). Pointing to oneself: active versus passive proprioception revisited and implications for internal models of motor system function. *Exp. Brain Res.* 229, 171–180. doi: 10.1007/s00221-013-3603-4

Conradt, C., Guo, D., Miclea, A., Nisslein, T., Ismail, C., Chatamra, K., et al. (2018). Increased prevalence of polyneuropathy in parkinson's disease patients: an observational study. *J. Parkinsons Dis.* 8, 141–144. doi: 10.3233/jpd-161057

Conte, A., Leodori, G., Ferrazzano, G., De Bartolo, M. I., Manzo, N., Fabbrini, G., et al. (2016). Somatosensory temporal discrimination threshold in Parkinson's disease parallels disease severity and duration. *Clin. Neurophysiol.* 127, 2985–2989. doi: 10.1016/j.clinph.2016.06.026

Conte, A., Modugno, N., Lena, F., Dispenza, S., Gandolfi, B., Iezzi, E., et al. (2010). Subthalamic nucleus stimulation and somatosensory temporal discrimination in Parkinson's disease. *Brain* 133, 2656–2663. doi: 10.1093/brain/awq191

Di Biasio, F., Conte, A., Bologna, M., Iezzi, E., Rocchi, L., Modugno, N., et al. (2015). Does the cerebellum intervene in the abnormal somatosensory temporal discrimination in Parkinson's disease? *Parkinsonism Relat. Disord.* 21, 789–792. doi: 10.1016/j.parkreldis.2015.04.004

Erro, R., Rocchi, L., Antelmi, E., Liguori, R., Tinazzi, M., Berardelli, A., et al. (2018). High frequency somatosensory stimulation in dystonia: evidence fordefective inhibitory plasticity. *Mov. Disord.* 33, 1902–1909. doi: 10.1002/mds.27470

Erro, R., Rocchi, L., Antelmi, E., Palladino, R., Tinazzi, M., Rothwell, J., et al. (2016). High frequency repetitive sensory stimulation improves temporal discrimination in healthy subjects. *Clin. Neurophysiol.* 127, 817–820. doi: 10.1016/j.clinph.2015.06.023

Fiorio, M., Stanzani, C., Rothwell, J. C., Bhatia, K. P., Moretto, G., Fiaschi, A., et al. (2007). Defective temporal discrimination of passive movements in Parkinson's disease. *Neurosci. Lett.* 417, 312–315. doi: 10.1016/j.neulet.2007.02.050

Fiorio, M., Tinazzi, M., Bertolasi, L., and Aglioti, S. M. (2003). Temporal processing of visuotactile and tactile stimuli in writer's cramp. *Ann. Neurol.* 53, 630–635. doi: 10.1002/ana.10525

Fiorio, M., Tinazzi, M., Scontrini, A., Stanzani, C., Gambarin, M., Fiaschi, A., et al. (2008). Tactile temporal discrimination in patients with blepharospasm. *J. Neurol. Neurosurg. Psychiatry* 79, 796–798. doi: 10.1136/jnnp.2007.131524

Francis, S. T., Kelly, E. F., Bowtell, R., Dunseath, W. J., Folger, S. E., and Mcglone, F. (2000). fMRI of the responses to vibratory stimulation of digit tips. *Neuroimage* 11, 188–202. doi: 10.1006/nimg.2000.0541

Gelnar, P. A., Krauss, B. R., Szeverenyi, N. M., and Apkarian, A. V. (1998). Fingertip representation in the human somatosensory cortex: an fMRI study. *Neuroimage* 7, 261–283. doi: 10.1006/nimg.1998.0341

Georgopoulos, A. P. (2002). Cognitive motor control: spatial and temporal aspects. *Curr. Opin. Neurobiol.* 12, 678–683. doi: 10.1016/s0959-4388(02)00382-3

Gibbon, J., Church, R. M., and Meck, W. H. (1984). Scalar timing in memory. *Ann. N. Y. Acad. Sci.* 423, 52–77. doi: 10.1111/j.1749-6632.1984.tb23417.x

Gritsenko, V., Krouchev, N. I., and Kalaska, J. F. (2007). Afferent input, efference copy, signal noise, and biases in perception of joint angle during active versus passive elbow movements. *J. Neurophysiol.* 98, 1140–1154. doi: 10.1152/jn.00162.2007

Hutchinson, M., Isa, T., Molloy, A., Kimmich, O., Williams, L., Molloy, F., et al. (2014). Cervical dystonia: a disorder of the midbrain network for covert attentional orienting. *Front. Neurol.* 5:54. doi: 10.3389/fneur.2014.00054

Ingemanson, M. L., Rowe, J. B., Chan, V., Wolbrecht, E. T., Cramer, S. C., and Reinkensmeyer, D. J. (2016). Use of a robotic device to measure age-related decline in finger proprioception. *Exp. Brain Res.* 234, 83–93. doi: 10.1007/s00221-015-4440-4

Kavounoudias, A., Roll, J. P., Anton, J. L., Nazarian, B., Roth, M., and Roll, R. (2008). Proprio-tactile integration for kinesthetic perception: an fMRI study. *Neuropsychologia* 46, 567–575. doi: 10.1016/j.neuropsychologia.2007.10.002

Kimmich, O., Molloy, A., Whelan, R., Williams, L., Bradley, D., Balsters, J., et al. (2014). Temporal discrimination, a cervical dystonia endophenotype: penetrance and functional correlates. *Mov. Disord.* 29, 804–811. doi: 10.1002/mds.25822

Ko, S. U., Simonsick, E., Deshpande, N., and Ferrucci, L. (2015). Sex-specific age associations of ankle proprioception test performance in older adults: results from the Baltimore Longitudinal Study of Aging. *Age Ageing* 44, 485–490. doi: 10.1093/ageing/afv005

Lee, M. J., Son, J. S., Lee, J. H., Kim, S. J., Lyoo, C. H., and Lee, M. S. (2016). Impact of prolonged temporal discrimination threshold on finger movements of Parkinson's disease. *PLoS One* 11:e0167034. doi: 10.1371/journal.pone.0167034

Lehmann, K., Steinecke, A., and Bolz, J. (2012). GABA through the ages: regulation of cortical function and plasticity by inhibitory interneurons. *Neural Plast.* 2012:892784. doi: 10.1155/2012/892784

Mc Govern, E. M., Killian, O., Narasimham, S., Quinlivan, B., Butler, J. B., Beck, R., et al. (2017). Disrupted superior collicular activity may reveal cervical dystonia disease pathomechanisms. *Sci. Rep.* 7:16753. doi: 10.1038/s41598-017-17074-x

Mima, T., Sadato, N., Yazawa, S., Hanakawa, T., Fukuyama, H., Yonekura, Y., et al. (1999). Brain structures related to active and passive finger movements in man. *Brain* 122(Pt 10), 1989–1997. doi: 10.1093/brain/122.10.1989

Mora, F., Segovia, G., and Del Arco, A. (2008). Glutamate-dopamine-GABA interactions in the aging basal ganglia. *Brain Res. Rev.* 58, 340–353. doi: 10.1016/j.brainresrev.2007.10.006

Naito, E., Nakashima, T., Kito, T., Aramaki, Y., Okada, T., and Sadato, N. (2007). Human limb-specific and non-limb-specific brain representations during kinesthetic illusory movements of the upper and lower extremities. *Eur. J. Neurosci.* 25, 3476–3487. doi: 10.1111/j.1460-9568.2007.05587.x

Pastor, M. A., Day, B. L., Macaluso, E., Friston, K. J., and Frackowiak, R. S. (2004). The functional neuroanatomy of temporal discrimination. *J. Neurosci.* 24, 2585–2591. doi: 10.1523/jneurosci.4210-03.2004

Pellicciari, M. C., Miniussi, C., Rossini, P. M., and De Gennaro, L. (2009). Increased cortical plasticity in the elderly: changes in the somatosensory cortex after paired associative stimulation. *Neuroscience* 163, 266–276. doi: 10.1016/j.neuroscience.2009.06.013

Ramos, V. F., Esquenazi, A., Villegas, M. A., Wu, T., and Hallett, M. (2016). Temporal discrimination threshold with healthy aging. *Neurobiol. Aging* 43, 174–179. doi: 10.1016/j.neurobiolaging.2016.04.009

Riemann, B. L., and Lephart, S. M. (2002a). The sensorimotor system, part I: the physiologic basis of functional joint stability. *J. Athl. Train* 37, 71–79.

Riemann, B. L., and Lephart, S. M. (2002b). The sensorimotor system, Part II: the role of proprioception in motor control and functional joint stability. *J. Athl. Train* 37, 80–84.

Rinderknecht, M. D., Lambercy, O., Raible, V., Liepert, J., and Gassert, R. (2017). Age-based model for metacarpophalangeal joint proprioception in elderly. *Clin. Interv. Aging* 12, 635–643. doi: 10.2147/CIA.S129601

Rocchi, L., Conte, A., Nardella, A., Li Voti, P., Di Biasio, F., Leodori, G., et al. (2013). Somatosensory temporal discrimination threshold may help to differentiate patients with multiple system atrophy from patients with Parkinson's disease. *Eur. J. Neurol.* 20, 714–719. doi: 10.1111/ene.12059

Rocchi, L., Erro, R., Antelmi, E., Berardelli, A., Tinazzi, M., Liguori, R., et al. (2017). High frequency somatosensory stimulation increases sensori-motor inhibition and leads to perceptual improvement in healthy subjects. *Clin. Neurophysiol.* 128, 1015–1025. doi: 10.1016/j.clinph.2017.03.046

Romaiguere, P., Anton, J. L., Roth, M., Casini, L., and Roll, J. P. (2003). Motor and parietal cortical areas both underlie kinaesthesia. *Brain Res. Cogn. Brain Res.* 16, 74–82. doi: 10.1016/s0926-6410(02)00221-5

Rothwell, J. C., Traub, M. M., Day, B. L., Obeso, J. A., Thomas, P. K., and Marsden, C. D. (1982). Manual motor performance in a deafferented man. *Brain* 105(Pt 3), 515–542. doi: 10.1093/brain/105.3.515

Sainburg, R. L., Ghilardi, M. F., Poizner, H., and Ghez, C. (1995). Control of limb dynamics in normal subjects and patients without proprioception. *J. Neurophysiol.* 73, 820–835. doi: 10.1152/jn.1995.73.2.820

Tinazzi, M., Fasano, A., Di Matteo, A., Conte, A., Bove, F., Bovi, T., et al. (2013a). Temporal discrimination in patients with dystonia and tremor and patients with essential tremor. *Neurology* 80, 76–84. doi: 10.1212/WNL.0b013e31827b1a54

Tinazzi, M., Morgante, F., Peretti, A., Mariotti, C., Panzeri, M., Fiorio, M., et al. (2013b). Impaired temporal processing of tactile and proprioceptive stimuli in cerebellar degeneration. *PLoS One* 8:e78628. doi: 10.1371/journal.pone.0078628

Tinazzi, M., Fasano, A., Peretti, A., Bove, F., Conte, A., Dallocchio, C., et al. (2014). Tactile and proprioceptive temporal discrimination are impaired in functional tremor. *PLoS One* 9:e102328. doi: 10.1371/journal.pone.0102328

Tinazzi, M., Fiorio, M., Bertolasi, L., and Aglioti, S. M. (2004). Timing of tactile and visuo-tactile events is impaired in patients with cervical dystonia. *J. Neurol.* 251, 85–90. doi: 10.1007/s00415-004-0282-x

Tinazzi, M., Fiorio, M., Stanzani, C., Moretto, G., Smania, N., Fiaschi, A., et al. (2006). Temporal discrimination of two passive movements in writer's cramp. *Mov. Disord.* 21, 1131–1135. doi: 10.1002/mds.20892

Tinazzi, M., Frasson, E., Bertolasi, L., Fiaschi, A., and Aglioti, S. (1999). Temporal discrimination of somesthetic stimuli is impaired in dystonic patients. *Neuroreport* 10, 1547–1550. doi: 10.1097/00001756-199905140-00028

Tinazzi, M., Stanzani, C., Fiorio, M., Smania, N., Moretto, G., Fiaschi, A., et al. (2005). Temporal discrimination of two passive movements in humans: a new psychophysical approach to assessing kinaesthesia. *Exp. Brain Res.* 166, 184–189. doi: 10.1007/s00221-005-2353-3

Weiller, C., Juptner, M., Fellows, S., Rijntjes, M., Leonhardt, G., Kiebel, S., et al. (1996). Brain representation of active and passive movements. *Neuroimage* 4, 105–110. doi: 10.1006/nimg.1996.0034

Zhang, L., Yang, J., Inai, Y., Huang, Q., and Wu, J. (2015). Effects of aging on pointing movements under restricted visual feedback conditions. *Hum. Mov. Sci.* 40, 1–13. doi: 10.1016/j.humov.2014.11.009 doi: 10.1016/j.humov.2014.11.009

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.

Copyright © 2019 Odorfer, Wind and Zeller. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Computational Neuroscience

METHODS
published: 10 December 2019
doi: 10.3389/fncom.2019.00082

A Spike Train Distance Robust to Firing Rate Changes Based on the Earth Mover's Distance

*Duho Sihn and Sung-Phil Kim**

Department of Human Factors Engineering, Ulsan National Institute of Science and Technology (UNIST), Ulsan, South Korea

Neural spike train analysis methods are mainly used for understanding the temporal aspects of neural information processing. One approach is to measure the dissimilarity between the spike trains of a pair of neurons, often referred to as the spike train distance. The spike train distance has been often used to classify neuronal units with similar temporal patterns. Several methods to compute spike train distance have been developed so far. Intuitively, a desirable distance should be the shortest length between two objects. The Earth Mover's Distance (EMD) can compute spike train distance by measuring the shortest length between two spike trains *via* shifting a fraction of spikes from one spike train to another. The EMD could accurately measure spike timing differences, temporal similarity, and spikes time synchrony. It is also robust to firing rate changes. Victor and Purpura (1996) distance measures the minimum cost between two spike trains. Although it also measures the shortest path between spike trains, its output can vary with the time-scale parameter. In contrast, the EMD measures distance in a unique way by calculating the genuine shortest length between spike trains. The EMD also outperforms other existing spike train distance methods in measuring various aspects of the temporal characteristics of spike trains and in robustness to firing rate changes. The EMD can effectively measure the shortest length between spike trains without being considerably affected by the overall firing rate difference between them. Hence, it is suitable for pure temporal coding exclusively, which is a predominant premise underlying the present study.

Keywords: neural spike train, spike train distance, Earth Mover's Distance, temporal coding, synchrony

OPEN ACCESS

Edited by:
Daya Shankar Gupta, Camden County College, United States

Reviewed by:
Michael Denker, Julich Research Centre, Germany
Thomas Kreuz, Italian National Research Council (CNR), Italy

***Correspondence:**
Sung-Phil Kim
spkim@unist.ac.kr

Received: *15 May 2019*
Accepted: *25 November 2019*
Published: *10 December 2019*

Citation:
Sihn D and Kim S-P (2019) A Spike Train Distance Robust to Firing Rate Changes Based on the Earth Mover's Distance. Front. Comput. Neurosci. 13:82. doi: 10.3389/fncom.2019.00082

INTRODUCTION

A spike train is the sequence of neuronal firing timings, where a spike refers to the firing of an action potential. The temporal pattern of a spike train encodes information in various ways. Besides firing rates, the temporal pattern of spike timings also carries important information about brain functions. For instance, it has been shown that temporal patterns encode the information of auditory (Machens et al., 2001; Narayan et al., 2006; Wang et al., 2007; Fukushima et al., 2015; Krause et al., 2017), gustatory (Di Lorenzo and Victor, 2003), motor (Vargas-Irwin et al., 2015), olfactory (MacLeod et al., 1998), somatosensory (Harvey et al., 2013), vestibular (Jamali et al., 2016), and visual (Mechler et al., 1998; Victor and Purpura, 1998; Reich et al., 2001; Carrillo-Reid et al., 2015) systems, as well as behavioral adaptation (Logiaco et al., 2015) and sleep (Tabuchi et al., 2018).

One of the many approaches for studying a temporal firing pattern is to measure the dissimilarity between a pair of spike trains, which is often represented by the spike train distance. The measurement of spike train distances can be designed to represent rate coding, temporal coding, or both. Several methods to measure a spike train distance have been proposed. Victor and Purpura introduced a cost-based distance that assigns a cost to shifting, adding, or deleting a spike (Victor and Purpura, 1996). In this method, the spike train distance is defined as the minimum of all possible sums of costs. The Victor and Purpura distance depends on a time-scale parameter where the smaller value of the time-scale parameter emphasizes temporal coding while the larger value does rate coding. van Rossum (2001) also developed a spike train distance that measures a difference between spike trains convolved with exponential functions. Most spike train distances are rate-sensitive, reflecting an overall rate difference between spike trains to a certain extent even with an extreme time-scale parameter (Satuvuori and Kreuz, 2018). Here, the overall rate denotes the total number of spikes in a spike train divided by the time length of the train. If one aims to measure a distance between a pair of spike trains independent of the overall rate difference, which we call as purely timing-sensitive, the distance should reflect only a difference of spike timing distributions, no matter how different the overall firing rate is between trains.

A purely timing-sensitive spike train distance is important to neuroscience studies on temporal coding, which assumes that neurons code information in spike timing patterns (Tabuchi et al., 2018). If a spike train distance is rate-sensitive, it would be difficult to clarify whether a given result from a neural spike train analysis is based only on the temporal information. It has been suggested that precise spike timing plays a crucial role in neural information processing (Butts et al., 2007; Gollisch and Meister, 2008; Johansson and Flanagan, 2009).

Kreuz et al. (2007) developed the rate-sensitive ISI-distance, a spike train distance based on a ratio between the inter-spike intervals of two spike trains. This was followed by the SPIKE-distance, a complementary distance which is still sensitive to rates but with a heightened sensitivity to spike timing (Kreuz et al., 2013). Finally, by removing rate dependence from the SPIKE-distance, Satuvuori et al. (2017) proposed the RI-SPIKE-distance as a distance purely sensitive to timing. The spike train distances developed so far have been used in a number of studies for the analysis of neural firing patterns (MacLeod et al., 1998; Mechler et al., 1998; Victor and Purpura, 1998; Machens et al., 2001; Reich et al., 2001; Di Lorenzo and Victor, 2003; Narayan et al., 2006; Wang et al., 2007; Harvey et al., 2013; Fukushima et al., 2015; Logiaco et al., 2015; Vargas-Irwin et al., 2015; Jamali et al., 2016; Krause et al., 2017).

Nevertheless, in an intuitive manner, one of the desirable properties of *distance* would be a capability to measure the shortest length between two objects. In this sense, the previous methods to measure spike train distance have not clearly represented the shortest length because they do not minimize the distance value explicitly, except for the Victor and Purpura distance, which explicitly measures the shortest length (Victor and Purpura, 1996). Yet, although this distance represents the minimum cost related to the shortest length, it suffers from the fact that distance output is not unique because this approach employs a parameter (i.e., q in their model) assigned to the cost for spike time shift. Thus, distance output depends on how q is determined. This property can be advantageous for some spike train analyses, but not in other cases that need a unique value (Chicharro et al., 2011). In the present study, therefore, we adopt the Earth Mover's Distance (EMD) to measure spike train distance with a unique shortest length.

The EMD is also called the Wasserstein metric, which defines the distance between a pair of probability distributions. Here, a metric refers to a distance satisfying non-negativity, symmetry and the triangle inequality. It measures the minimal cost based on an underlying distance taken to transfer from a probability distribution to another. It initially dealt with transportation problems (Kantorovich, 1940) and later modified toward today's form (Vaserstein, 1969). The EMD also has been implemented as an algorithm in the field of computer science for the comparison between two images (Rubner et al., 2000). The main idea underlying the EMD is that the shortest distance between two objects is equal to the length of the shortest delivery path from one object to the other. For neural spike data, *delivery* in a spike train operates by moving a part of the spike train from one location to another, with a goal to match one spike train with the other. A delivery path length is then calculated by summing the delivery distance between two locations multiplied by the amount of a delivered part. If we deal with a spike train as a distribution with a sum of 1, then the EMD measures a unique shortest distance between a pair of spike trains in a non-parametric way. A notable difference of the EMD from that of Victor and Purpura (1996) is that delivery in the method of Victor and Purpura moves an entire spike at once while delivery in the EMD can move a part of a spike.

Spike train distance can be used for both rate coding and temporal coding (Satuvuori and Kreuz, 2018). Rate coding accounts for the firing rate profile of neuronal spike trains while temporal coding relates to temporal patterns. In our development of spike train distance, we focus on a particular aspect of the firing rate profile, an overall firing rate difference between spike trains, whereas we refer a temporal pattern of spike train to the distribution of spike timings in time within a spike train. Specifically, the temporal pattern focuses on the pattern of a spike timing distribution as a function of time, not on how many spikes occur in any particular time window. For example, a spike train with spike timings at (0, 1, 10) has a similar temporal pattern to another spike train (0, 0.1, 0.9, 1, 10, 10.1) while their overall firing rates are different. In the case of temporal coding, spike train distance is often used to measure the dissimilarity of temporal patterns of neuronal spike trains, which may not be explained by rate coding alone. For such a case, measurements of spike train distance need to be independent of firing rate changes. On the other hand, if a spike train distance is sensitive to firing rate changes, it may be ambiguous whether the analysis results reflect changes of mere temporal patterns or a mixed effect of firing rates. Hence, robustness against firing rate changes should be a desired property of spike train distances in pure temporal coding studies. Previous methods for spike train distance have

not focused much on this robustness except the one by RI-SPIKE-distance (Satuvuori et al., 2017). As such, the EMD is chosen here to ensure robustness to firing rate changes. The EMD can measure spike distance robust to the overall firing rate difference between spike trains because the EMD normalizes the total amount of spikes in a spike train to 1, making the overall firing rate of each spike train equal.

In this study, we employ the EMD as a spike train distance for pure temporal coding research. Then, we compare the EMD with several other spike train distances using neural spike data generated from a set of simulations. The simulations are designed to evaluate the performance of the spike train distances with respect to essential aspects of temporal patterns, including spike timing differences, temporal similarity, and spike time synchrony, as well as the robustness against firing rate changes in spike trains to deal with pure temporal coding. In this study, we refer temporal coding to a scheme to represent a spiking probability as a function of time. It is different from a time-varying firing rate as it does not reflect actual firing rates over time. In several simulation tests, we evaluate how various spike train distance methods, including the proposed one, represent pure temporal coding using a spike generation probabilistic model, in which a spiking probability varies with time independent of the number of spikes. This includes the test of the robustness of each method against firing rate changes by alternating the total number of spikes while maintaining temporal coding unchanged. The advantages of the EMD in pure temporal coding research are demonstrated by the simulation results. However, it should be noted that these advantages are not directly transferable to rate coding.

MATERIALS AND METHODS

The Earth Mover's Distance as a Spike Train Distance

Two different spike trains may contain a different number of spikes. However, the total number of spikes of each spike train should be equalized to measure the distance between them based only on shifting spikes in time. Victor and Purpura (1996) solved this problem by assigning a cost to adding/deleting a spike and to shifting a spike in time. However, this solution cannot produce a unique distance because it varies with the ratio of two different costs. To address this shortcoming, in the proposed method, we first define a spike train in which each spike is assigned a fixed quantity of 1. Then, we normalized individual spikes by the total number of spikes, N, so that each spike's quantity becomes 1/N after normalization. For the normalization, we consider a spike train as a function f of time t such that

$$f(x) = \begin{cases} \frac{1}{N}, & \text{if a spike occurs at time } t \\ 0, & \text{otherwise} \end{cases} \quad (1)$$

Where N is the number of spikes in the spike train. The overall summation of f must be one except the case of $N = 0$. Hereafter, a spike train will be expressed as functions f or g.

In our method, the EMD between f and g proposed in Rubner et al. was adjusted for one-dimensional data (i.e., a spike train) with a constraint that the sum of f or g should be equal to 1 (Rubner et al., 2000). The EMD is described as follows. We first rewrite the spike trains, $f = \{(x_1,\ 1/N),\ (x_2,\ 1/N),\ \ldots,\ (x_{\mathrm{N}},\ 1/N)\}$ and $g = \{(y_1,\ 1/M),\ (y_2,\ 1/M),\ \ldots,\ (y_M,\ 1/M)\}$ from Eq. 1 where x_{i} and y_{j} are a sequence of spike timings. Let $d(x_{\mathrm{i}},\ y_{\mathrm{j}})$ be an absolute difference between two spike timings x_{i} and y_{j}. Let ξ_{ij} be a flow (amount of delivery) from x_{i} to y_{j} and let $\Xi = [\xi_{\mathrm{ij}}]$ be a matrix of these flows (amount of deliveries) such that it transports f to g satisfying the following conditions: (1) ξ_{ij} is non-negative; (2) $\sum_{\mathrm{i}=1}^{N} \xi_{\mathrm{ij}} \le 1/M$, $\sum_{\mathrm{j}=1}^{M} \xi_{\mathrm{ij}} \le 1/N$; and (3) $\sum_{\mathrm{i}=1}^{N}\sum_{\mathrm{j}=1}^{M} \xi_{\mathrm{ij}} = 1$. Condition 1 fixes the direction of the delivery from i to j. Condition 2 indicates an effective delivery in the sense that it does not take back what has been delivered. Condition 3 indicates that it delivers the entire spike train. The transportation here means that it makes f equal to g by moving parts of f. Then, the EMD between f and g is given by

$$\mathrm{EMD}(f, g) = \min\left\{ \sum_{i=1}^{N}\sum_{j=1}^{M} d(x_i,\ y_j)\,\xi_{ij} \ :\ \Xi = [\xi_{\mathrm{ij}}] \text{ satisfies conditions above} \right\} \quad (2)$$

This concept of spike train distance is illustrated in **Figure 1**.

When $N = 0$ (i.e., no spike in the train), the proposed method cannot calculate the distance directly. However, it can deal with such a case indirectly if we consider a spike train with no spike similar to a spike train with spikes everywhere so that the distance of it to any other trains becomes irrelevant to a certain spike timing pattern. Let f_0 be a spike train with no spike and let g be another spike train to be compared. To calculate $d(f_0,\ g)$, let f_{n} be a spike train with n spikes generated from a uniform probability distribution defined on a certain bounded analysis domain. The bounded analysis domain prevents the distance from increasing to infinity, although the distance measurement depends on how the analysis domain is determined. Then, the EMD calculates $d(f_0,\ g) = \lim_{\mathrm{n}\to\infty} \mathrm{E}(d(f_{\mathrm{n}}, g))$ where $\mathrm{E}(\cdot)$ indicates an expected value. To deal with an empty spike train in the EMD, we attended to an idea that there was also no information about spike timing if spikes are everywhere, uniformly distributed. It means that a spike train with one spike at a specific location holds more information about spike timing than a spike train with uniformly distributed spikes. In this regard, an empty spike train would be more similar to a spike train with uniformly distributed spikes at every location than a spike train with one spike.

The EMD is a mathematical metric, that is, it satisfies the three conditions: non-negativity, symmetry and the triangle inequality (Rubner et al., 2000). This property shows that the EMD conforms to our intuition about distance. Moreover, from the fact that the EMD is calculated solely based on spike timing data, it can be seen that the EMD is the shortest length based on

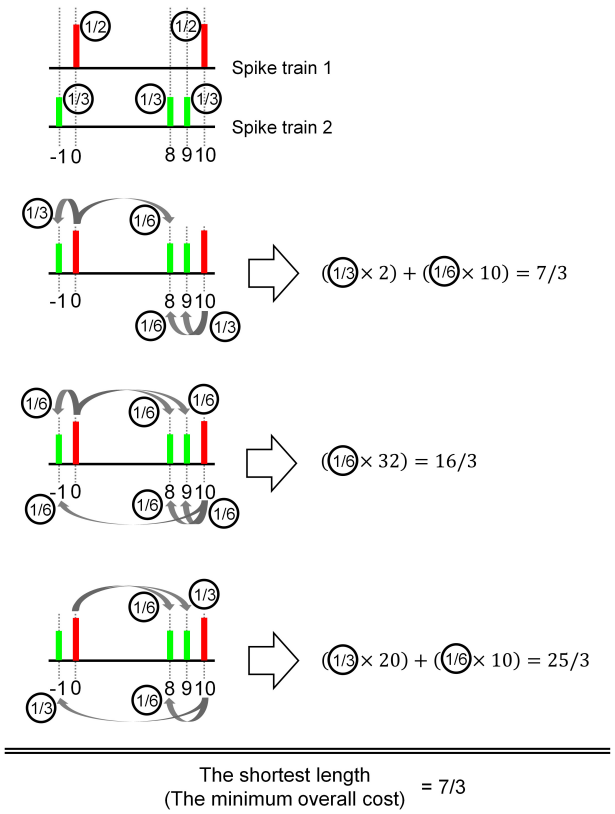

FIGURE 1 | Illustration of the basic concept of the Earth Mover's Distance (EMD) to measure a distance between spike trains. The objective is to take the smallest value among all possible delivery (flow, terminology in EMD) path lengths between two objects (red and green). In this illustrative example, the red object (e.g., spikes) is delivered to the green object in three possible paths. Assuming the size of the red object is 1, the path length is calculated by delivery distance (in time) multiplied by the amount of delivery (i.e., size of the object). It is also possible to deliver only a fraction of the object, as shown in the third case. If there are multiple deliveries toward the target object, the final delivery path length is the sum of individual delivery path lengths.

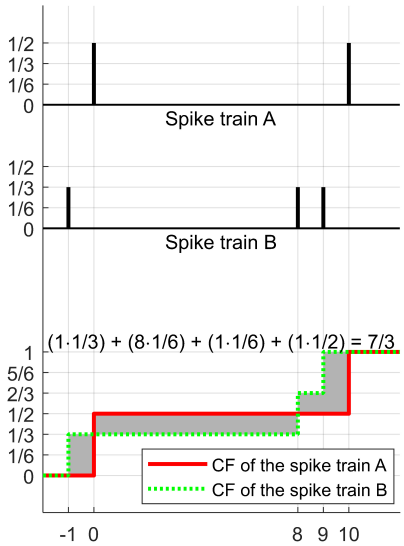

FIGURE 2 | Illustration of the calculation procedure of the EMD described in Cohen (1999). The distance between two spike trains, A and B, is calculated. Initially, the non-negative values are assigned to every spike such that the sum of the values in each train is equal to 1 (e.g., 1/2 for each spike in A or 1/3 for each in B). The next step is to produce the cumulative functions (CF) for each spike train (red bold line indicates the CF of spike train A and green dotted line indicates the CF of spike train B). The next step is to integrate the absolute difference between the two CFs (gray shading area). The final result of the calculation procedure is 7/3.

spike timing between two spike trains. The EMD is calculated in a non-parametric way so that it produces a unique value. Due to its non-parametric approach, the EMD can avoid the dependency of distance outcomes on parameters.

Moreover, there is an efficient way to calculate the restricted version of the EMD as follows (Cohen, 1999). Let F and G be the cumulative functions of f and g, respectively. Then, the EMD is given by

$$\mathrm{EMD}\left(f,g\right) = \int_{-\infty}^{\infty} |F\left(t\right) - G(t)|\,\mathrm{d}t \qquad (3)$$

An example of the calculation procedure above is illustrated in **Figure 2**.

Relationship With Other Measures

The Kullback–Leibler divergence is a distance between two probability distribution functions. Therefore, the Kullback–Leibler divergence computes the difference between two functions at every point on the domain of a random variable, similar to the EMD computation as shown in Eq. 3. However, a difference between the Kullback–Leibler divergence and the EMD is that the EMD depends not only on the difference between the functional values but also on the distance between the points on the domain while the Kullback–Leibler divergence does not. In Eq. 3, the EMD is calculated by (the difference between functions) × (the length to which the difference is maintained), so that the EMD is based on spike timing difference unlike the Kullback–Leibler divergence.

In the Victor-Purpura distance, spikes are shifted if the distance between spikes is small or added/deleted if the distance is large, depending on the time-scale parameter q. On the other hand, in the EMD, no matter how large the distance is, the spikes are always shifted. This indicates that the EMD can be viewed as similar to the Victor-Purpura distance with an extremely high cost of adding/deleting spikes. But, since the parameter q of the Victor-Purpura distance controls the time shift cost only, not adding/deleting explicitly, imposing a high cost on adding/deleting spikes can be implemented by selecting a very small value for q. Consequently, the Victor-Purpura distance with a small time-scale parameter (q) becomes similar to the EMD, with an emphasis on temporal coding.

If two spike trains have the same number of spikes, N, and the Victor-Purpura distance does not take the option of the cost for adding/deleting spikes, the Victor-Purpura distance and the EMD are exactly the same with the time-scale parameter q = 1/N. Hence under those conditions, the EMD can be considered as the average displacement of the spikes.

When two spike trains have different numbers of spikes, the EMD still calculates the average displacement of the spikes to some extent: the displacement of the part of a spike instead of an entire spike. The displacement of the part of a spike only reflects the temporal difference between spike trains. In contrast, the Victor-Purpura distance works in a different way due to the option of the cost for adding/deleting spikes. Since the cost for adding/deleting spikes directly correlates with a difference in the number of spikes between trains, the Victor-Purpura distance can reflect the rate difference. Hence, it has been pointed out that the Victor-Purpura distance is suitable for rate coding but not for temporal coding if the number of spikes is quite different between spike trains (Satuvuori and Kreuz, 2018).

When two spike trains, f and g, have the same number of spikes, we can describe the Victor-Purpura distance without the option of adding/deleting spikes with respect to the EMD as follows: $d_{\mathrm{EMD}}(f, g) = \lim_{q\to 0} (1/q)\, d_{\mathrm{VP}[q]}(f, g)$, where d_{EMD} indicates the EMD between f and g and $d_{\mathrm{VP}[q]}$ indicates the Victor-Purpura distance with the time-scale parameterq. Even if f and g have a different number of spikes, the description above holds if the Victor-Purpura distance is applied to the normalized spike train as in the EMD.

Evaluation

Our new spike train distance was compared to four existing spike train distances: (1) the Victor-Purpura distance (Victor and Purpura, 1996) with parameter values, q = 0.1, 0.2, . . ., 12.8 s^{-1}; (2) the van Rossum distance (van Rossum, 2001) with parameter values, τ = 1, 2, . . ., 16 s. Note that an alternative calculation method (Houghton and Kreuz, 2012) was used here instead of the original one (van Rossum, 2001); (3) the SPIKE-distance (Kreuz et al., 2013); and 4) the RI-SPIKE-distance (Satuvuori et al., 2017).

The tested time-scale parameters of the Victor-Purpura distance and the van Rossum distance were determined as follows. For the Victor-Purpura distance parameter q, the time range of a spike train in which we performed the analysis was set to 0 - 10 s. Then we opted for values of q varying between two opposite cases: q = 0.1 s^{-1} and q = 12.8 s^{-1}. The smallest q = 0.1 s^{-1} in the Victor-Purpura distance made the metric focus on a "spike timing shift" by assigning a cost of 1 to add/delete each spike, whereas it costed at most $(q = 0.1\ s^{-1}) \times (10\ s) = 1$ for time-shifting a spike. Then, the value of q was increased by a factor of two up to the largest q = 12.8 s^{-1}, which turned the algorithm to focus on "spike adding/deleting" by increasing the cost for time-shifting such as $(q = 12.8\ s^{-1}) \times (1\ s) = 12.8$ even for shifting a spike by 1 s.

Similarly, for the van Rossum distance, the smallest value of τ = 1 s makes the convolved range narrow by setting the width of the exponential function to 1 s. Then, the value of τ was increased by a factor of two up to the largest value of τ = 16 s, which makes the convolved range cover the overall spike train by setting the width of the exponential function to 16 s.

Taking spike counts into dissimilarity is a key difference between the EMD and the Victor-Purpura distance or the van Rossum distance. In fact, while the EMD is focused on temporal coding, both the Victor-Purpura distance and the van Rossum distance cover from a mixture of temporal coding and rate coding to pure rate coding by varying the time scale parameter q or τ, as they are so designed originally. We demonstrated such differences between the EMD and the Victor-Purpura distance or the van Rossum distance in the simulations (see section "Results").

A comparison of the five spike train distances was conducted to assess how well each distance represented three aspects of similarity between spike trains: spike timing difference, temporal similarity, and spike time synchrony. Furthermore, each distance's robustness to changes in firing rates was examined for temporal similarity and spike time synchrony.

To avoid potential errors while replicating the existing distance calculation procedures, we directly utilized the available source code for each distance. The code to calculate the Victor-Purpura distance was obtained from http://www-users.med.cornell.edu/~jdvicto/spkdm.html.
The code for the van Rossum distance was from http://wwwold.fi.isc.cnr.it/users/thomas.kreuz/images/vanRossum.m.
The codes for both the SPIKE-distance and the RI-SPIKE-distance were from http://wwwold.fi.isc.cnr.it/users/thomas.kreuz/Source-Code/cSPIKE.html.

For the calculation of the SPIKE-distance and the RI-SPIKE-distance, we always set the time range of the underlying dissimilarity profiles exactly equal to the spike generation interval.

Spike Timing Difference

A pair of spike trains with three spikes each was synthesized to test spike timing difference. The locations of the first and third spikes were fixed and matched between the trains. The second spike of the first train was fixed close to the first spike. Then, the location of the second spike of the second train was moved toward the third spike. This test paradigm was performed in the previous study by Kreuz et al. (2011) to compare several distances. We adopted it here with the inclusion of the van Rossum distance, the RI-SPIKE-distance, and the EMD. In the test, we located the first spike at 0 s and the third at 10 s in the two trains. The second spike of the first train was fixed at 1 s. Then, the second spike of the second train was moved from 1 s to 9 s in steps of 1 s (see **Figure 3A**). We measured the distance for each shift of the second spike of the second train.

Temporal Similarity and Robustness to Firing Rate Change

A simulation experiment was performed to test the robustness of each distance against firing rate changes when measuring temporal similarity between spike trains. Spike trains were generated according to a simple probabilistic model. The probabilistic model was built following a certain firing rate profile. Temporal similarity would increase if a pair of spike trains were generated from a probabilistic model sharing a similar profile and decrease if the profiles become more dissimilar. Note that temporal similarity describe here depends only upon firing rate profiles, not firing rates themselves. The probabilistic model used here consisted of two intervals where each interval

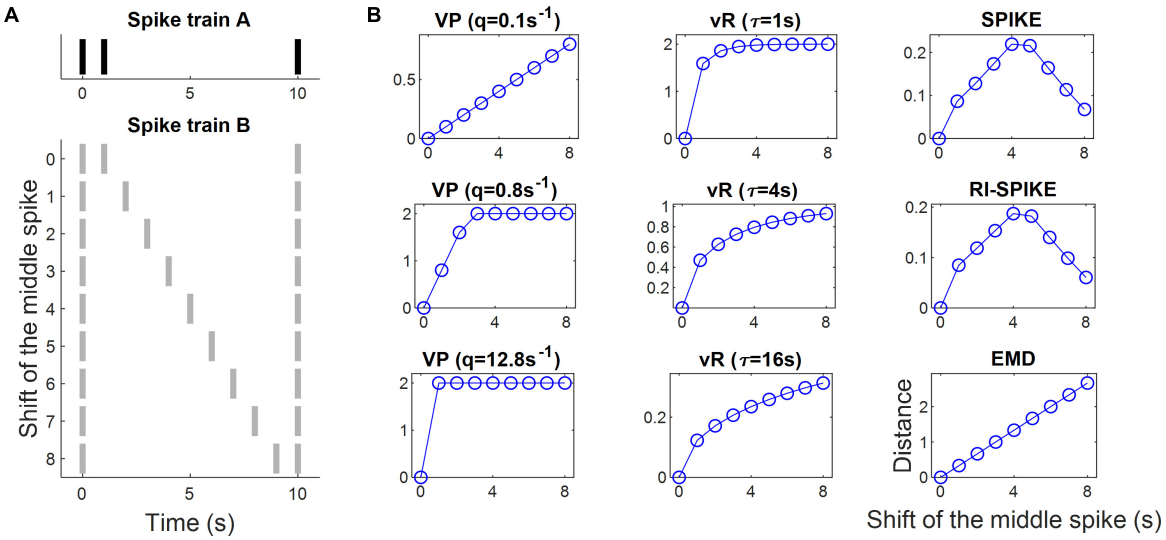

FIGURE 3 | Spike train distance results for the measurement of spike timing differences. **(A)** Spike train A is fixed whereas spike train B is changed as the location of a middle spike is shifted from left to right so that the spike timing difference between A and B increases linearly. **(B)** The spike train distance results of each of the five distances: the Victor-Purpura distance, the van Rossum distance, the SPIKE-distance, the RI-SPIKE-distance, and the EMD. The horizontal axis represents the amount of the shift of the middle spike in train B. The Victor-Purpura distance showed a linear increase in distance only for a certain parameter (e.g., q = 0.1 s). The van Rossum distance showed an increase in distance non-linearly but monotonically. The SPIKE-distance and the RI-SPIKE-distance did not show monotone increases. The EMD showed a linear increase as the spike timing difference increased. VP, Victor-Purpura distance; vR, van Rossum distance.

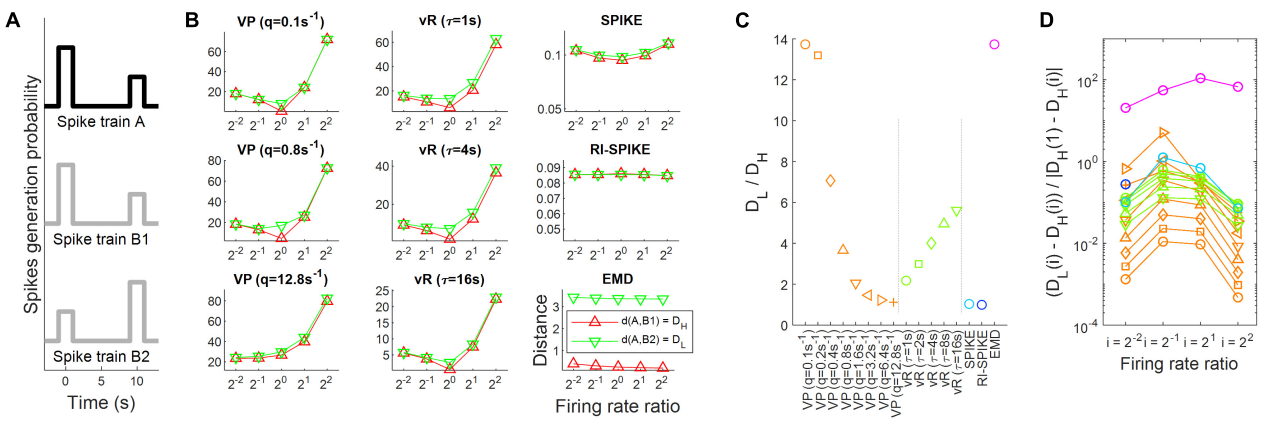

FIGURE 4 | Spike train distance results for the measurement of temporal similarity. **(A)** The probabilistic models of spike generation for spike trains A, B1, and B2 are described. In the simulation, spike trains A and B1 share the same probabilistic model whereas spike trains A and B2 have different probabilistic models. Accordingly, the temporal similarity is high between A and B1, but low between A and B2. **(B)** The spike train distance results of each of the five distances as the ratio of firing rates between the spike trains varies from 2^{-2} to 2^{2}. The red lines represent distances between the spike trains A and B1 and green lines represent those between A and B2. It is clearly shown that the variability of distances by changes in the ratio is larger than that by changes in the temporal similarity for the four distances, including the Victor-Purpura, the van Rossum, the SPIKE- and the RI-SPIKE- distances. In contrast, the distances calculated by the EMD remain almost unchanged as the ratio changes, being robust to the firing rate change. **(C)** Results of spike train distance for measuring temporal similarity. D_L is a distance with a low temporal similarity, and D_H is a distance with a high temporal similarity. **(D)** Quantification of robustness as the firing rate changes. $D_L(i)$ is the distance with a low temporal similarity when the firing rate ratio is i, and $D_H(i)$ is the distance with a high temporal similarity when the firing rate ratio is i. The results of the RI-SPIKE-distance partly disappear because of negative values. VP, Victor-Purpura distance; vR, van Rossum distance.

had a non-zero probability of containing a spike. Spikes in the intervals were randomly generated from a uniform distribution centered at 0 s and 10 s with a halfwidth of 1 s. Then, we built three spike trains denoted as spike trains A, B1, and B2. In the probabilistic model of spike train A, the probability of generating a spike in the first interval was twice as high as that in the second interval. Spike train B1 had the same probabilistic model as spike train A. On the other hand, it was reversed in spike train B2 such that the probability of generating a spike in the second interval was twice that in the first interval (see **Figure 4A**). Hence, the distance between A and B1 should be smaller than that between A and B2, because

temporal patterns would be more similar between A and B1 than between A and B2.

To test the robustness of the distances against firing rate changes, we varied the number of spikes in the trains. We first set the number of spikes in A to $2^3 \times 3$, where 2^3 spikes were generated three times (twice in the first interval and once in the second interval). Then, five levels of the number of spikes were used to vary the firing rates in B1 or B2. The number of spikes in B1 or B2 was varied as $2^1 \times 3$, $2^2 \times 3$, $2^3 \times 3$, $2^4 \times 3$, and $2^5 \times 3$, making the spike count ratios of A to B1 or B2 2^{-2}, 2^{-1}, 2^0, 2^1, and 2^2. If a spike train distance is robust to firing rate changes, distance variability over all the ratios should be negligible compared to the difference in distance between A to B1 and between A and B2. We calculated the difference in distances between these two pairs (A and B1, A and B2) using each of the five distances by varying the firing rates in B1 or B2.

Spike Time Synchrony and Robustness to Firing Rate Change

Another simulation experiment was performed to test the robustness of each distance against firing rate changes when measuring spike time synchrony between spike trains. To this end, a pair of spike trains, denoted as A and B were synthesized. Spike train A was generated to contain eleven equally spaced spikes discharged at 0 s, 1 s, . . ., 10 s. Spike train B was generated according to a probabilistic model, consisting of eleven uniform distributions centered at 0 s, 1 s, . . ., 10 s. Then, we varied the halfwidth of these uniform distributions across ten levels to manipulate the degrees of spike timing jitter; the halfwidth was set as 0.05 s, 0.1 s, . . ., or 0.5 s (see **Figure 5A**). As the halfwidth was increased, spike timing jitter increased, which was likely to desynchronize spike timing more between A and B. It would then result in an increase in the distance between A and B.

Similar to section "Temporal Similarity and Robustness to Firing Rate Change," we varied the number of spikes in B to test the robustness of the distance to firing rate change. The number of spikes in B varied across nine levels to reflect firing rate changes. It varied as 1×11, 2×11, . . ., and 9×11 (the first number in the product indicates the number of spikes randomly generated in each interval of B) so that the ratios of A to B became 1, 2, . . ., and 9, respectively. We expected that if the spike train distance was robust to firing rate changes, variability in the distance across the ratios should be negligible compared to variability in distance according to different degrees of spike timing jitter. We calculated the distances between A and B for each degree of spike timing jitter for each firing rate level in B.

Comparison With Victor and Purpura's Distance

The spike train distance in the present study is closely related to the Victor-Purpura distance. It is important to compare the properties between the Victor-Purpura distance and the EMD. Satuvuori and Kreuz (2018) already discussed the suitability of the Victor-Purpura distance to rate and temporal coding. They suggested that the Victor-Purpura distance is suitable to rate coding in general, but suitable to temporal coding only for similar firing rates, even with a wide range of time-scale parameter q. To verify whether the EMD suffered from a similar issue to the Victor-Purpura distance, we applied the analysis of Satuvuori and Kreuz (2018) to the EMD. Three spike trains were generated in the analysis. Spike train A was generated to contain one spike discharged at 5 s. Spike train B was generated according to a probabilistic model of a uniform distribution centered at 5 s with the halfwidth of 1 s. Spike train C was also generated according to a probabilistic model of a uniform distribution centered at 5 s with the halfwidth of 5 s. Spike train B had five levels of the number of spikes; 2^0, 2^1, 2^2, 2^3, and 2^4. By comparison, spike train C had only one spike as in spike train A (see **Figure 6A**). From the point of view of temporal coding, it was expected that the distance between A and B was smaller than the distance between A and C and the distance between B and C, because spike trains A and B had more similar temporal information compared to C. The Victor-Purpura distance was examined for time-scale parameters in the range from 0.01 to 1000.

RESULTS

Spike Timing Difference

The spike train distance measurements exhibited differences among the five spike train distances tested in this study. The Victor-Purpura distance linearly increased as the spike timing difference increased with one parameter value (q = 0.1 s), but was saturated with the other parameter values (q = 0.8, 12.8 s). Similarly, the van Rossum distance monotonically increased as the spike timing difference increased with one parameter (τ = 16 s), but was saturated with another parameter (τ = 1, 4 s). Both the SPIKE-distance and the RI-SPIKE-distance increased first but later decreased as the spike timing difference increased. The EMD linearly increased as the spike timing difference increased (**Figure 3B**).

Also, we observed that the SPIKE-distance and the RI-SPIKE-distance consider the spike trains to be more similar if a middle spike is close to the edge spikes than if the middle spike is located at an equal distance from both edge spikes. The reason is that these methods focus on the local dissimilarity between spike trains. Two spike trains are locally similar when the middle spike is close to the edge spikes since then it becomes easier to see it as part of a doublet that together is quite synchronous with the single spike in the other spike train. In contrast, if the middle spike is located at an equal distance from both edge spikes, then the distance to the nearest spike in the other train is maximized, increasing local dissimilarity. The van Rossum distance seems to evaluate a similarity of two spike trains based on synchronization of spike timings within a certain temporal range, where the temporal range was determined by the time-scale parameter τ. Then, if two spikes from each spike train occurred within the temporal range, these spikes were deemed to be synchronized. The Victor-Purpura distance with the parameter q = 0.1 s (i.e., emphasizing temporal differences) and the EMD linearly increase as the difference of middle spikes is linearly increased, because these methods focus on equalizing two spike trains. Hence, for instance, if a difference in the latency of neural responses between spike trains is of interest, the Victor-Purpura distance with a small q and the EMD can provide an appropriate measure.

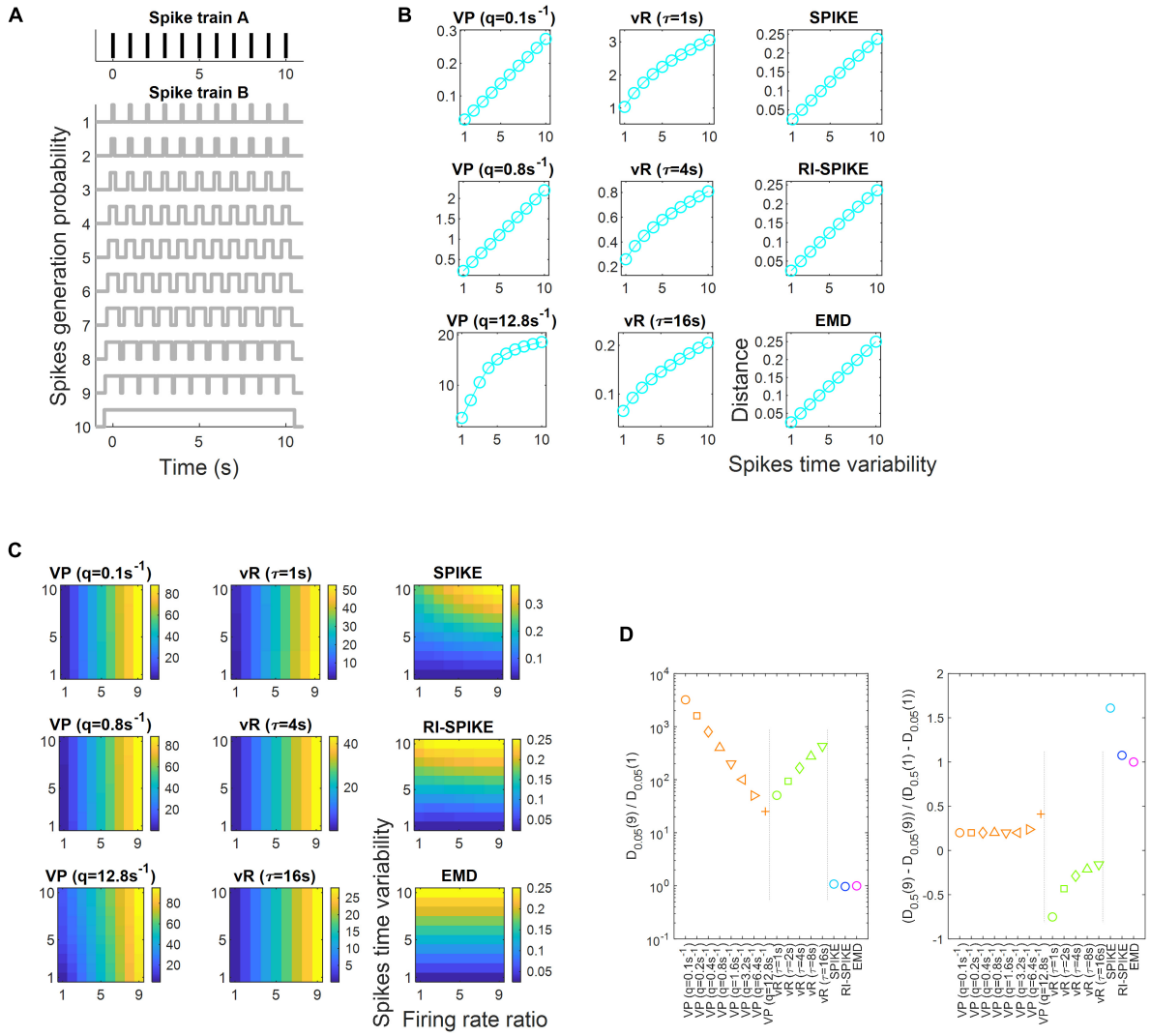

FIGURE 5 | Spike train distance results for the measurement of spike time synchrony. **(A)** Spike train A is fixed to have equally spaced eleven spikes. Spike train B is generated by a probabilistic model of spike generation with various spike timing jitter. The spike timing jitter is manipulated by increasing the halfwidth of eleven uniform distributions each centered at the spike timing of train A. Spike time synchrony between A and B decreases as spike timing jitter increases. **(B)** The spike train distance results of the five distances as the ratio of firing rates of B over A are equal to 1. All the distances exhibit approximately linear increases with increases in spike timing jitter. **(C)** The spike train distance results of each of the five distances as ratios of firing rates of B over A increase from 1 to 9. The index of the vertical axis corresponds to the index of the spike trains in panel **(A)**, where increasing index number indicates increasing spike timing jitter. Distances proposed by Victor-Purpura and van Rossum are significantly affected by the variation in the firing rate ratio, whereas those proposed by the SPIKE-distance, the RI-SPIKE-distance and the EMD are not. **(D)** Results of spike train distance for measuring spike time synchrony. $D_k(n)$ is the distance when the firing rate ratio of one spike train to another was n, and k denotes the halfwidth of the uniform distribution in the spike train. VP, Victor-Purpura distance; vR, van Rossum distance.

The characteristics of distances for small spike timing differences (for example, the middle spike is shifted by 0, 1, or 2 in **Figure 3B**) can provide information about temporal precision of the spike timing. The Victor-Purpura distance (q = 0.1 s) and the EMD are linearly decreasing when the spike timing difference converges to zero. This linear property allows them to have the information about temporal precision, but with no conclusive answer to whether a timing difference between spike trains is precise or not. On the other hand, the van Rossum distance, the SPIKE-distance and the RI-SPIKE-distance are rapidly decreasing when the spike timing difference is nearing zero so that they can provide precise information whether timing difference falls within some range or not.

Temporal Similarity and Robustness to Firing Rate Change

We evaluated distance measurements between a pair of spike trains with a high or low temporal similarity when the ratio of the firing rates between the trains varied. Let D_L be a distance with a low temporal similarity (i.e., between A and B2) and D_H be a distance with a high temporal similarity (i.e., between A and B1)

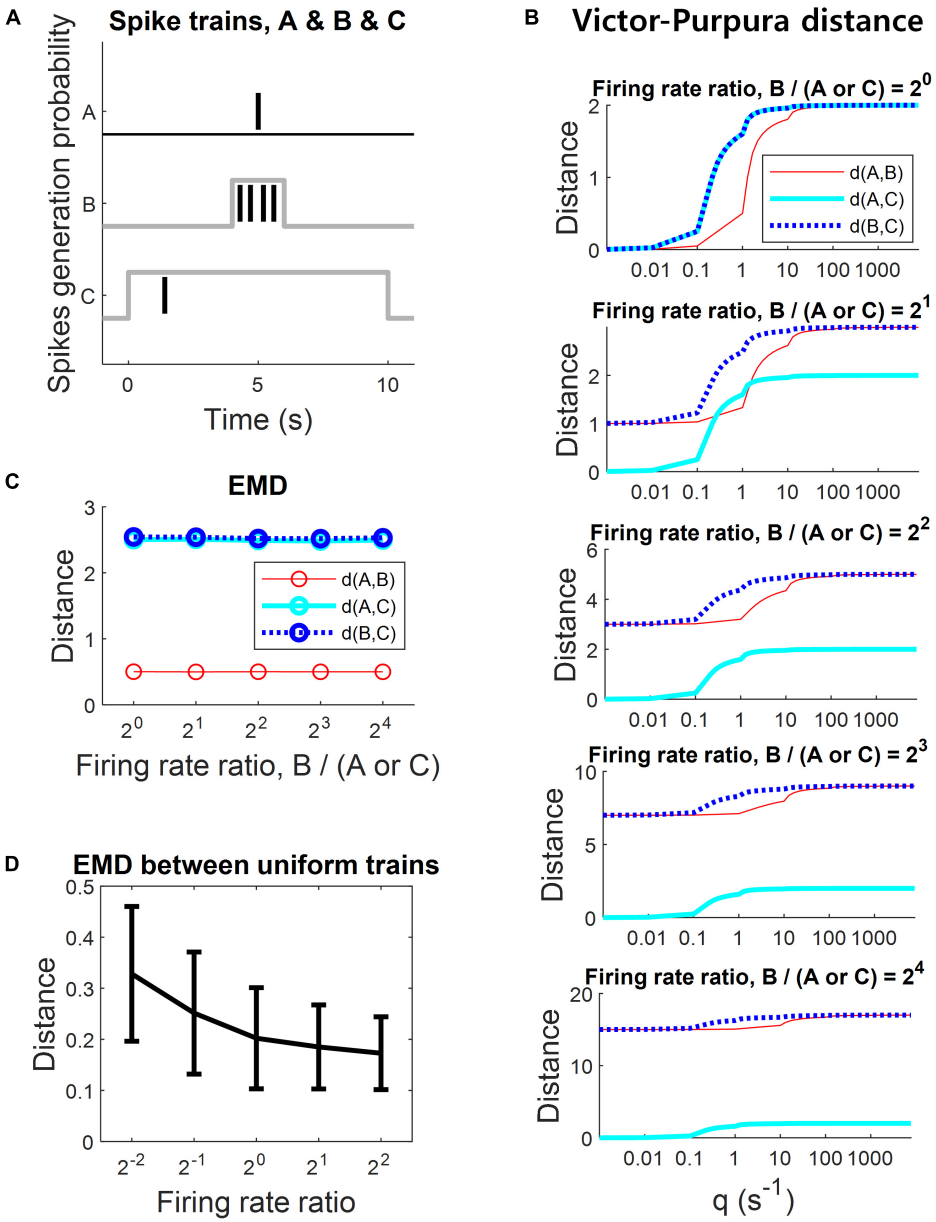

FIGURE 6 | Comparison with the Victor-Purpura distance in terms of suitability for temporal coding with different firing rates. **(A)** Spike train A has only one spike with fixed timing. Spike train B has five levels of spikes with narrow range spike timing jitters. Spike train C has only on spike with a broad range spike timing jitter. The desirable expected results are that the distance between spike trains A and B is smaller than the distance between A and C and the distance between B and C. **(B)** The Victor-Purpura distance with various values of the time-scale parameter q. The Victor-Purpura distance did not show the desirable result with increases in firing rate ratio. **(C)** The EMD showed desirable results overall with increases in firing rate ratio, having a nearly constant scale. **(D)** The EMD between a uniform spike trains with different firing rate ratios. It shows that the EMD is not completely insensitive to firing rate differences.

(see section "Temporal Similarity and Robustness to Firing Rate Change"). First, we calculated the ratios of D_L to D_H from each distance for the case when the firing rates of two spike trains were equal, and the result is summarized in **Figure 4C**. The Victor-Purpura distance (q = 0.1, 0.8 s), the van Rossum distance (τ = 4, 16 s), and the EMD clearly resulted in a smaller distance with a high temporal similarity than with a low temporal similarity (**Figure 4B**). These low and high temporal similarities reflect the global difference between two spike trains in **Figure 4A**, not the local difference. The spike trains B1 and B2 in **Figure 4A** are globally different, but locally similar (near 0 s and 10 s). Since both the SPIKE-distance and the RI-SPIKE-distance focus on the local difference, these distances show less sensitivities for the discrimination between low and high temporal similarity in a global sense. In contrast, the EMD is a global measurement, showing an ability to discriminate global temporal similarity.

Although the RI-SPIKE-distance is robust to firing rate changes just as the EMD is, this point indicates a key difference between the RI-SPIKE-distance and the EMD (see section "Discussion").

Next, to assess the robustness to firing rate changes when unequal firing rates exist between the spike trains, we calculated the ratio $[D_L(i) - D_H(i)]/|D_H(1) - D_H(i)|$, where i denotes the firing rate ratio of spike train B1 (or B2) to that of spike train A for $i = 1/4, 1/2, 2, 4$ [e.g., $D_H(1/2)$ refers to distance measurements when the firing rate ratio is 1/2]. The distance results for each value of i are given in **Figure 4D**. As for the robustness to firing rate changes, the Victor-Purpura distance and the van Rossum distance increased as the ratio of the firing rates deviated from 1, which indicates that variability in the distance across the firing rate ratios was larger than the difference in distances between high and low temporal similarities, revealing that the distances were not robust to firing rate changes. This was not the case for the SPIKE-distance and the RI-SPIKE-distance, where the distances remained at similar levels across the ratios of firing rates although variability in the distance across the ratios was larger than the difference in distances between high and low temporal similarities, showing that they were also not robust to firing rate changes. On the other hand, the EMD showed that variability in the distance across the ratios was much smaller than the difference in distances between high and low temporal similarities, demonstrating its robustness to firing rate changes (**Figure 4B**). The SPIKE-distance, the RI-SPIKE-distance, and the EMD showed the robustness to firing rate changes relative to the Victor-Purpura distance and the van Rossum distance. It implies that those three distances are more suitable for temporal coding.

Spike Time Synchrony and Robustness to Firing Rate Change

Spike train distances with various synchrony levels were measured using each of the five distances and their robustness to firing rate changes was tested. Every distance clearly showed a similar pattern when the ratio of firing rates was 1 such that the spike train distance increased as the degree of spike timing jitter increased (**Figure 5B**). To assess the robustness to firing rate changes, we quantified the effect of the firing rate ratio on the spike train. Here, let $D_k(n)$ be the distance when the firing rate ratio of spike train B to spike train A was n, where k denotes the halfwidth of the uniform distribution in B (see section "Spike Time Synchrony and Robustness to Firing Rate Change"). We first calculated the ratio $D_{0.05}(9)/D_{0.05}(1)$ using each distance and obtained the results as summarized in the left figure of **Figure 5D**. Next, we calculated the ratio $[D_{0.5}(9) - D_{0.05}(9)]/[D_{0.5}(1) - D_{0.05}(1)]$ using each distance and obtained results, which are listed in the right figure of **Figure 5D**. A comparison of these two ratios showed that when the firing rate ratio increased, the Victor-Purpura distance and the van Rossum distance increased rapidly, whereas other distances were almost unchanged. In other words, by using the Victor-Purpura distance and the van Rossum distance, variability in distance across the firing rate ratios was larger than variability in distance due to different degrees of spike timing jitter, showing that the distances were not robust to firing rate changes. The SPIKE-distance, the RI-SPIKE-distance, and the EMD revealed that variability in distance across the ratios was smaller than that among different levels of synchrony, demonstrating that they were robust to firing rate changes. Moreover, the RI-SPIKE-distance and the EMD appeared to be most robust (**Figure 5C**). These results indicate that the Victor-Purpura distance and the van Rossum distance are suitable to measure the dissimilarity due to both rate difference and temporal synchrony. The SPIKE-distance is also suitable to measure the dissimilarity in both rate difference and temporal synchrony although it seems to be less sensitive to rate difference than the Victor-Purpura distance and the van Rossum distance. On the other hand, the RI-SPIKE-distance and the EMD are suitable to measure temporal synchrony, insensitive to rate differences.

Comparison With Victor and Purpura's Distance

The simulation result for the Victor-Purpura distance in the present study was similar to that in the study by Satuvuori and Kreuz (2018). The expected result was that the distance between the spike trains A and B was smaller than those between A and C and between B and C, because the temporal coding between A and B is more similar than that between other pairs (**Figure 6A**). When the firing rate ratio of B to A or C was 2^0 (i.e., the same firing rates), the Victor-Purpura distance showed the expected result for a wide range of time-scale parameters q (**Figure 6B**, top). It indicates that the Victor-Purpura distance is suitable for temporal coding if the firing rate ratio is 1. However, as the firing rate ratio of B to A or C increased, the Victor-Purpura distance started to show unexpected results. The distances between A and B and between B and C were increasing for every time-scale parameter q, reflecting the increased rate difference (**Figure 6B**). It indicates that the Victor-Purpura distance is not suitable for temporal coding if the firing rate ratio deviates from 1. The smaller value of the time-scale parameter q emphasizes the temporal coding. However, the result showed that the Victor-Purpura distance is still rate-sensitive even for a very small value of q. Therefore, the value of q apparently changes sensitivity from pure rate coding to combined rate and temporal coding, not to pure temporal coding (Satuvuori and Kreuz, 2018).

On the other hand, the EMD showed the expected results for all tested firing rate ratios. Furthermore, the distances between every pair of spike train remained nearly constant even as the firing rate ratio changed (**Figure 6C**). It indicates that the EMD is suitable for temporal coding even though the firing rates differ between the spike trains, showing that it does not reflect rate coding. That is, the EMD is sensitive to pure temporal coding in contrast to the Victor-Purpura distance.

Although the EMD is relatively insensitive to firing rate difference than the Victor-Purpura distance, it is uncertain whether the EMD is completely insensitive. In order to test the effect of different rate ratios on the EMD, we calculated the EMD between two Poisson spike trains that were generated uniformly over [0, 1] s with different rates. The spike trains were generated with firing rates of 1, 2, 4, 8, and 16 Hz. Then, the spike trains with 4 Hz were compared to those with other firing rates (including

the identical 4 Hz) so that the firing rate ratios varied over 2^{-2}, 2^{-1}, 2^{0}, 2^{1}, and 2^{2}. The resulting EMD values are provided in **Figure 6D**. The EMD between trains with the same temporal pattern varied across different firing rate ratios although the EMD variation was much smaller than the firing rate ratios variation.

Application to Neural Data

We demonstrated the measurement of a temporal similarity between real neuronal spike trains using the EMD. The neural data is publicly available from Flint et al. (2012), and can be downloaded from https://crcns.org/data-sets/movements/dream. The example of neural spike trains was obtained from the primary motor cortex of a behaving non-human primate (Flint et al., 2012). An example of various levels of temporal similarity measured by the EMD is shown in **Figure 7**, in which the spike trains observed under the different experimental conditions (i.e., different movement directions of the subject's arm) showed mutually different temporal similarity with the base condition at the arm movement direction of 45° (at which the example neuron fired the most).

A neuron in the primary motor cortex (M1) modulates its firing rates with arm movement directions (Georgopoulos et al., 1982). Arm movements induce a certain temporal pattern such that a spike train of a M1 neuron contains more spikes around movement onset and less spikes before and after movement offset. Also, the firing rate of the neuron is maximal at the preferred direction (PD) of arm movement and decreases gradually when the movement direction deviates farther from the PD (Georgopoulos et al., 1982; Schwartz et al., 1988; Kalaska et al., 1989; Caminiti et al., 1990). Hence, the temporal patterns of spike trains between the PD and other directions are expected to be more dissimilar when the movement direction becomes more different from the PD. We found that the EMD could describe various levels of temporal similarity to the base condition for various directions and specifically showed that distance increased as the angle became orthogonal to the PD. In addition, the EMD on the true data (red lines in the inlet graph of **Figure 7A**) revealed a clearer difference between the PD and orthogonal angles than that on the surrogate data with randomized spike timings (gray lines of the inlet graph of **Figure 7A**). Specifically, corresponding to each true spike train, we generated a random spike train by generating spike timings from the uniform distribution while maintaining the number of spikes unchanged. So, if the difference between directions is mainly represented in the number of spikes, the difference between directions should also be maintained in the surrogate data. However, the result demonstrated that the EMD difference between spike trains of different directions was not merely due to the firing rate difference.

A spike train distance shall yield small values between spike trains obtained under similar experimental conditions and large values between spike trains obtained under different conditions. We demonstrated that the EMD satisfied such a criterion using the real neuronal spike data of a non-human primate in **Figure 7**. In **Figure 7**, the EMD showed

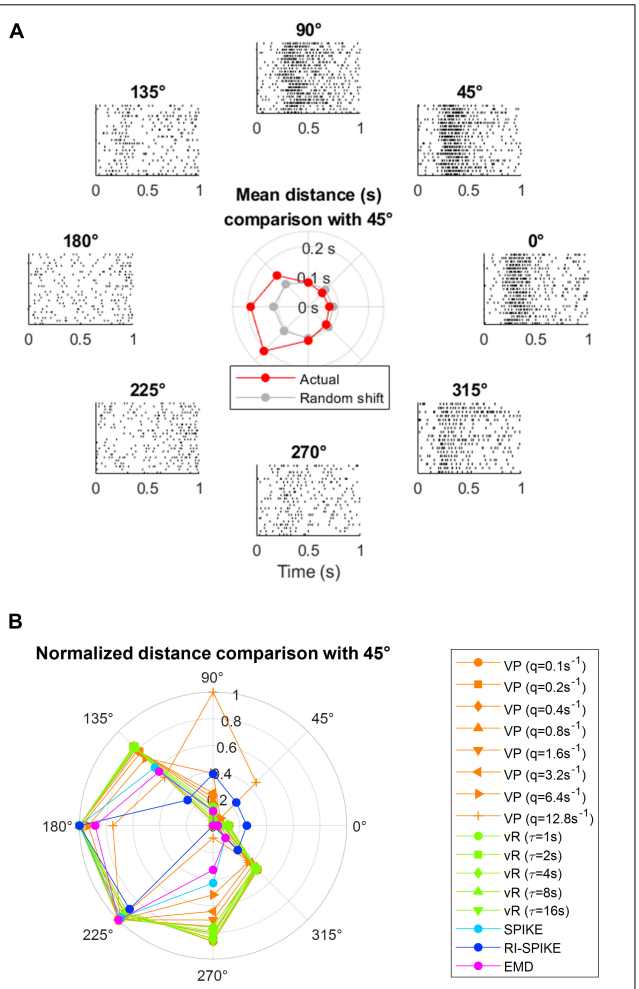

FIGURE 7 | Application of the spike train distance to real neuronal data in the primary motor cortex in a non-human primate (Flint et al., 2012). During the data recordings, the subject moved its arm from the central position toward one of the eight target positions and repeated this movement multiple times for each direction. Multiple spike trains of a single neuron for each of the eight target positions are described at each peripheral location, indicated by a directional angle as 0°, 45°, . . ., and 315°. Each spike train is obtained for 1 s after the onset of a movement cue. The spike trains exhibit different temporal patterns for different directions. **(A)** The direction at 45° is set as the seed direction, where the firing rate is maximum. Then, the spike train distance is calculated between the seed direction and each of other seven directions. The mean spike train distance between each pair is described using red circles in the center. The spike train distance within the seed direction is also calculated for comparison (no calculation between the same spike trains). The EMD from actual data (red line of the inner graph) has a clearer difference between the base and orthogonal angles than the EMD from data of randomly shifted spike timing (gray line of the inner graph), which does not exhibit a temporal pattern, showing the EMD difference is not merely due to the firing rate differences. **(B)** Comparison between the EMD and other spike train distances for the data in panel **(A)**. Each spike distance was normalized such that D_new = (D − D_min)/(D_max − D_min) so that the distance values are filled between 0 and 1, because each spike distance has different magnitude scale. Throughout our spike distance analyses, we have set up a time-scale of spike trains for the Victor-Purpura distance and the van Rossum distance as [0, 10] s, which makes these distances applicable to both rate coding and temporal coding. To be consistent with such parameter settings of all the analyses done in the study, we also maintained the same time-scale range for the analysis of real

(Continued)

FIGURE 7 | Continued
neuronal spike data in panel **(A)**. Since the spike trains of the real neurons we analyzed lasted for 1 s after a task onset, we extended spike trains by multiplying 10 to spike timings, changing the spike train range from [0, 1] s to [0, 10] s and used the same parameter settings as other simulation-based analyses for the Victor-Purpura and the van Rossum distances. This extension of the spike train range does not alter the SPIKE-distance, RI-SPIKE-distance and the EMD because they produce time-scale independent distance outcomes. VP, Victor-Purpura distance; vR, van Rossum distance.

small values when the subject moved the arm in a direction similar to the preferred direction (i.e., similar experimental condition) and large values when the subject moved the arm in a direction dissimilar to the preferred direction (i.e., dissimilar experimental condition) (see red lines in the inlet graph of **Figure 7A**). In particular, the EMD calculated this result based on the temporal pattern rather than on the firing rate difference.

We compared EMD and other spike train distances in terms of an ability to distinguish primary motor cortical spike trains with spiking timing information according to the arm movement directions of a non-human primate. There were eight equally divided arm movement directions in this 2D center-out arm reaching task. As each spike train distance covered a different magnitude scale, each spike train distance was normalized by D_new = (D − D_min)/(D_max − D_min) so that the distance values ranged between 0 and 1. We selected one of the eight directions as an anchor (e.g., 45°) and measured average pairwise distance using each spike distance measure between a set of spike trains corresponding to the anchor direction and each set of spike trains corresponding to other directions. We found that the EMD well represented differences between spike trains according to movement directions such that the distance is 0 at the PD, 1 at the opposite of the PD, and the intermediate values at other directions (**Figure 7B**).

We evaluated how the EMD could be used to discriminate the neural spiking patterns of different upper limb movement directions represented in the primary motor cortex (M1) of a non-human primate (Flint et al., 2012). The non-human primate moved the upper limb in eight different directions while spiking timings of the population of M1 neurons were recorded. There were multiple trials of this task in each direction. As the duration of movements varied across trials, we selected a 1-s epoch after the onset of a go cue. Before spike train distance computation using various methods including the EMD, we normalized the overall spike count of every spike train in order to assess each method's ability to extract movement-related information only from spiking timing patterns. This normalization was performed based on resampling – i.e., randomly selecting a certain number of spikes from the original spike train. In this manner, every resampled spike train could have the same number of spikes for every direction while retaining the temporal pattern of the original spike trains.

For resampling, we first selected 113 out of 196 M1 neurons, which fired spikes enough to produce spike trains suitable for our distance analysis (a neuron was selected if it fired $\geq$50 spikes within the 1-s epoch on average for each direction). For each selected neuron, we randomly chose R spikes from the original spike train, repeating this resampling for every spike train of every direction for that neuron. The number of spikes in a resampled spike train, R, was stochastically determined by generating a random number from the Poisson distribution with the mean rate of 10. The mean rate of 10 was chosen such that the largest number generated from the Poisson distribution with this mean rate was unlikely to exceed the half of 50 (i.e., 25), in order to make resampled spike trains vary over trials. This ensured that the expected number of spikes in every resampled train in every direction was identical, while allowing trial-to-trial variability. Once the resampled spike train was generated, we multiplied 10 to its spike timings to change the spike train range from [0, 1] s to [0, 10] s, in order to adjust the range adequate for pre-defined time-scale parameters of the Victor-Purpura distance and the van Rossum distance. Also, as the SPIKE-distance and the RI-SPIKE-distance calculate the distance in a range from the first spike to the last spike, we added two auxiliary spikes at 0 and 10 s (**Figure 8A**).

For the assessment of each spike train distance method, we randomly selected a single resampled spike train in the k-th direction and calculated distance between it and every other resampled spike trains using a given spike train distance. Those calculated distances were averaged for each direction, yielding the average distances d_i (i = 1,. . .,8) for each of the eight directions. The averaged distances were then normalized over direction such that $\overline{d_i} = \left(d_i - d_{min}\right)/(d_{max} - d_{min})$ as above. The shorter normalized distance toward the correct direction, $\overline{d_k}$, represented better discrimination of the spiking patterns for the correct direction from other directions. The EMD, as well as the Victor-Purpura distance and the van Rossum distance with specific parameter settings, resulted in shorter distances than others (**Figure 8B**). Therefore, it demonstrated that the EMD could decode the directional information of upper limb movements in M1 neurons based on spiking timing patterns.

We applied a clustering analysis [Houghton and Victor (2010) and Victor and Purpura (1996)] to the data shown in **Figure 8A** in order to compare the effect of each distance metrics on decoding the information of movement directions from spike trains. For decoding such directional information based on the shortest distance to the training samples of spike trains, we used the k-nearest neighbor algorithm (Fix and Hodges, 1951). The decoding performance were measured by the normalized transmitted information proposed in the study by Houghton and Victor (2010), which ranges from 0 to 1 where a higher value indicates more accurate decoding. The result demonstrated that the EMD produced the best decoding output (with the number of neighbors, k = 3) (**Figure 8C**).

We also applied the same clustering analysis to the data shown in **Figure 7A** without removing firing rate differences, in order to examine the effect of directionally tuned firing rates on the spike train distances. We observed that the Victor-Purpura distance and the van Rossum distance produced larger normalized transmitted information than the SPIKE-distance,

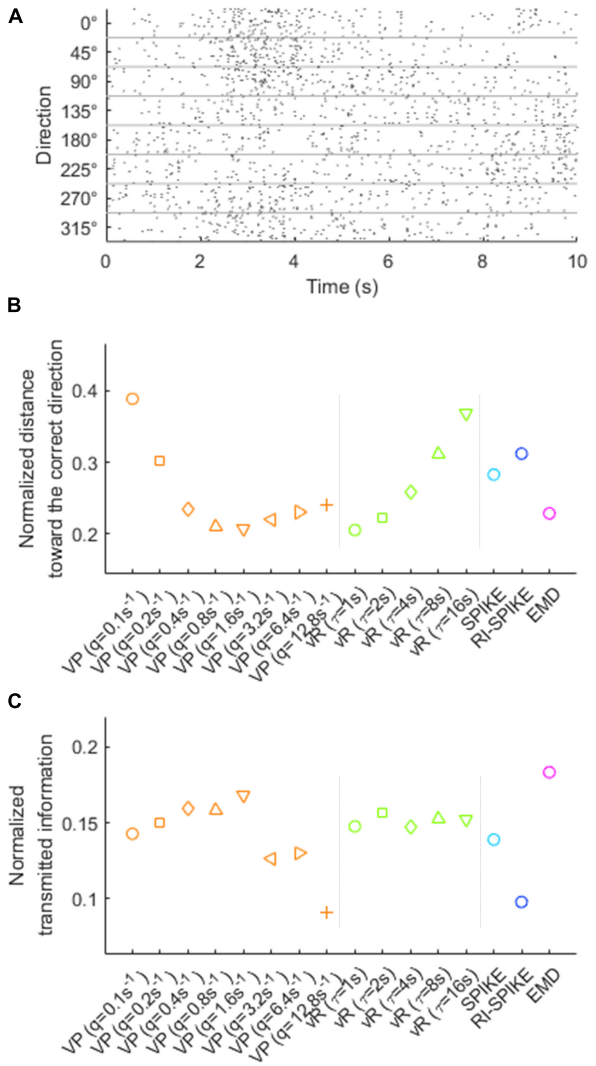

FIGURE 8 | (A) Resampled spike trains of the neuron in **Figure 7**. The spikes in the resampled spike train are randomly chosen from the pool of spike timings in each direction. The spiking patterns in the original spike train is preserved while the number of spikes is controlled. The range of spike trains extended from [0, 1] s to [0, 10] s. **(B)** The directional decoding results using spiking patterns in resampled spike trains. The distance indicates the average distance for each direction. The distance for the direction is normalized by other directions. The small value of the normalized distance toward the correct direction indicates a high magnitude of discrimination of the spiking patterns for the direction from the other directions. **(C)** Decoding directional information from the spike trains in panel **(A)** is performed using the k-nearest neighbor algorithm (k = 3 in our analysis) and evaluated by the normalized transmitted information (see the text). Higher normalized transmitted information indicates better decoding performance. VP, Victor-Purpura distance; vR, van Rossum distance.

the RI-SPIKE-distance, and the EMD regardless of the setting of time-scale parameters. It demonstrates that the first two distances are more suitable for rate coding than last three distances, as also shown in sections "Temporal Similarity and Robustness to Firing Rate Change" and "Spike Time Synchrony and Robustness to Firing Rate Change."

DISCUSSION

In the present study, we applied the EMD to neuroscience as a spike train distance to measure the shortest delivery path length between spike trains. In this distance, a spike train was considered as a function that assigned a non-negative value at spiking time such that the sum of all non-negative values was equal to one. For any two functions in this metric space, one function could be transformed into another function through the iteration of delivering a quantity at a point in the domain of a function to another point. Each delivery created a path whose length could be quantified by the product of the amount of the moving quantity (i.e., a fraction of a spike) and the delivery time. The sum of all delivery paths was then defined as the delivery path length. Among all possible delivery paths, the shortest path was sought, and its length was used as the spike train distance. We demonstrated that our distance sufficiently expressed temporal similarity based on the temporal profile of spiking probabilities and spiking time synchrony between a pair of spike trains, and that it was more robust to differences in absolute firing rates with a common temporal profile of spike probabilities than previous distances.

The metric EMD is induced by the metric based on the temporal events. It means that a distance between two spike trains is entirely measured from distances between spikes within those spike trains. Owing to this property, the EMD can vary linearly in response to linear changes in spike timing (**Figure 3**). This linear property may strengthen the reliability of the EMD for capturing spike timing differences between spike trains and allow one to easily determine how the distance would vary with spike timing variation. On the other hand, the EMD can provide information about temporal precision, but not conclusive information whether spike timing difference is in some range or not, due to this linearity.

The EMD measures a difference between two normalized spike trains, in contrast to other distances that use spike trains *per se* without normalization. This normalization allows the EMD to compare the actual temporal patterns of a pair of spike trains with negligible influences from firing rates. This property makes the EMD more robust to firing rate changes than other distances (**Figures 4**, **5**). It is expected that temporal coding research may take advantage of this property.

Of course, not every spike train distance should be robust against firing rate change. If a distance between spike trains with similar firing rates is smaller than between spike trains with different firing rates, it is suitable for representing rate coding. Yet, if certain cases require information merely from temporal coding, the robustness against firing rate change would be necessary for distance measures. Since the EMD does not reflect rate coding as discussed above, the spike counts would be supplementary to the EMD.

The EMD as a spike train distance is based only on spike timing differences, which allows the EMD to be adequate for temporal coding. The EMD measures the minimum length of spike timing shifts to make two spike trains identical.

To calculate the length, the amount of spikes in two spike trains should be the same and the normalization step making the amount of spikes in each train equal to one is necessary. The normalization step plays a crucial role in the robustness to firing rate changes. Thus, existing methods other than the EMD can be applied to normalized spike trains, which would preserve the robustness to firing rate changes. However, other methods may not be as adequate as the EMD for temporal coding. For example, the Jensen–Shannon divergence, which measures dissimilarity between two probability distributions, can be applied to the normalized spike trains because the normalized spike trains can be considered as a probability distribution. But, since it is not concerned with spike timing differences, it may not provide measurements useful for temporal coding. Also, the Victor-Purpura distance can be applied to the normalized spike trains. However, since the Victor-Purpura distance has the adding/deleting spikes option, it cannot guarantee that the dissimilarity is calculated based only on spike timing differences. Hence, we suggest that the EMD can be advantageous over other methods to provide spike train distance based solely on timing differences between a given pair of spikes and therefore useful for temporal coding schemes.

Precise spike timing is a key element in temporal coding (Butts et al., 2007; Gollisch and Meister, 2008; Johansson and Flanagan, 2009). There are largely two different approaches to measure how much the spike timings of a pair of spike trains match with each other. One way is to measure a global difference between the trains of spike timings, and the other way is to measure local matches between the trains of spike timings. For example, in the global measurement, a spike train (2,3,4,5) can be matched with a spike train (1,2,3,4) by shifting all spike timings by $+1$, and a spike train (1,2,3,5) can be matched with the spike train (1,2,3,4) by shifting its last spike timing by -1. Therefore, the distance between (1,2,3,4) and (2,3,4,5) is larger than the distance between (1,2,3,4) and (1,2,3,5). The Victor-Purpura distance with a short time-scale parameter and the EMD measure spike distances in this way: 0.4 vs. 0.1 in the Victor-Purpura distance with q = 0.1, and 1 vs. 1/4 in the EMD. In the local measurement, the spike train (1,2,3,4) and the spike train (2,3,4,5) are locally matched at three different timings ({2, 3, 4}), and the spike train (1,2,3,4) and the spike train (1,2,3,5) are also locally matched at three different timings ({1, 2, 3}). As such, the distance between (1,2,3,4) and (2,3,4,5) is the same as the distance between (1,2,3,4) and (1,2,3,5). The van Rossum distance, the SPIKE-distance, and the RI-SPIKE-distance measure spike distances in this way. Hence, we can select the global spike train distance measurement if we intend to measure how similar the distributions of spike timings in two spike trains are, or the local spike train distance measurement if we intend to focus on local spike timing matching. In this study, we propose a new method for the global spike train distance measurement.

Although the EMD is less affected by firing rate changes compared to other distances, it is not completely invariant to firing rate changes. For instance, in our simulations, when two spike trains were generated from uniform distributions in an interval from 0 to 1 s with a firing rate of 1 Hz, the EMD between two spike trains was 0.33 ± 0.24 s. However, when two spike trains were generated in the same interval with a firing rate of 10 Hz, the distance decreased to 0.14 ± 0.06 s. Therefore, there was a tendency for the EMD to decrease as the firing rate increased. Also, these results were confirmed with Poisson spike trains in **Figure 6D**. The results showed that the EMD values varied with different firing rate ratios although the variation was very smaller than firing rate ratios. Therefore, meticulous care is required when using the EMD for temporal coding research without considering rate coding completely. However, in cases where two spike trains exhibit certain temporal patterns and those two temporal patterns are different, the EMD would quantify dissimilarity well between two temporal patterns even if they have fairly different firing rates.

The present study mainly addressed the sensitivities of spike train distances to rate and temporal coding. However, in addition to sensitivities, each distance offers a unique feature. The SPIKE-distance and the RI-SPIKE-distance have fine time resolutions and thus can measure differences in local spike patterns.

The EMD also has a definite advantage such that it can be extended to stochastic spike trains as follows. Many noise sources perturb the generation of spikes, inducing a variability of spiking events (Faisal et al., 2008). Due to this variability, we can consider a spike train as a stochastic process. For examples, the peristimulus time histogram based on the average of trials or the probabilistic reconstruction of a spike train (Kass and Ventura, 2001) takes stochastic spike trains into account. In this sense, a spike train distance that can deal with continuous data is needed to compare two stochastic spike trains. Moreover, such a spike train distance should be based on distance metrics defined with deterministic spike trains, as a stochastic spike train can be viewed as a natural variant of a deterministic spike train (Haslinger et al., 2009). The proposed distance, EMD, bases itself in a metric space for deterministic spike trains and can also be applied to stochastic spike trains in the form of normalized continuous data.

DATA AVAILABILITY STATEMENT

The datasets generated for this study can be found in the [CRCNS (Flint et al., 2012)] (https://crcns.org/data-sets/movements/dream).

AUTHOR CONTRIBUTIONS

Both authors designed the study and wrote the manuscript. DS performed simulation and analyzed the data.

FUNDING

This research was supported by the Brain Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (2016M3C7A1904988).

REFERENCES

Butts, D. A., Weng, C., Jin, J., Yeh, C., Lesica, N. A., Alonso, J., et al. (2007). Temporal precision in the neural code and the timescales of natural vision. *Nature* 449, 92–95. doi: 10.1038/nature06105

Caminiti, R., Johnson, P. B., and Urbano, A. (1990). Making arm movements within different parts of space: dynamic aspects in the primate motor cortex. *J. Neurosci.* 10, 2039–2058. doi: 10.1523/jneurosci.10-07-02039.1990

Carrillo-Reid, L., Miller, J. K., Hamm, J. P., Jackson, J., and Yuste, R. (2015). Endogenous sequential cortical activity evoked by visual stimuli. *J. Neurosci.* 35, 8813–8828. doi: 10.1523/JNEUROSCI.5214-14.2015

Chicharro, D., Kreuz, T., and Andrzejak, R. G. (2011). What can spike train distances tell us about the neural code? *J. Neurosci. Methods* 199, 146–165. doi: 10.1016/j.jneumeth.2011.05.002

Cohen, S. (1999). *Finding Color and Shape Patterns in Images*. Doctoral dissertation, Stanford University, Stanford, CA.

Di Lorenzo, P. M., and Victor, J. D. (2003). Taste response variability and temporal codingin the nucleus of the solitary tract of the rat. *J. Neurophysiol.* 90, 1418–1431. doi: 10.1152/jn.00177.2003

Faisal, A. A., Selen, L. P. J., and Wolpert, D. M. (2008). Noise in the nervous system. *Nat. Rev. Neurosci.* 9, 292–303.

Fix, E., and Hodges, J. L. (1951). "Discriminatory analysis, nonparametric discrimination: consistency properties. Technical Report 4. Dayton, OH: USAF School of Aviation Medicine

Flint, R., Lindberg, E. L., Jordan, L. R., Miller, L. E., and Slutzky, M. W. (2012). Accurate decoding of reaching movements from field potentials in the absence of spikes. *J. Neural Eng.* 9:046006. doi: 10.1088/1741-2560/9/4/046006

Fukushima, M., Rauske, P. L., and Margoliash, D. (2015). Temporal and rate code analysis of responses to low-frequency components in the bird's own song by song system neurons. *J. Comp. Physiol. A* 201, 1103–1114. doi: 10.1007/s00359-015-1037-0

Georgopoulos, A. P., Kalaska, J. F., Caminiti, R., and Massey, J. T. (1982). On the relations between the direction of two-dimensional arm movements and cell discharge in primate motor cortex. *J. Neurosci.* 2, 1527–1537. doi: 10.1523/jneurosci.02-11-01527.1982

Gollisch, T., and Meister, M. (2008). Rapid neural coding in the retina with relative spike latencies. *Science* 319, 1108–1111. doi: 10.1126/science.1149639

Harvey, M. A., Saal, H. P., Dammann, J. F. III, and Bensmaia, S. J. (2013). Multiplexing stimulus information through rate and temporal codes in primate somatosensory cortex. *PLoS Biol.* 11:e1001558. doi: 10.1371/journal.pbio.1001558

Haslinger, R., Klinkner, K. L., and Shalizi, C. R. (2009). The computational structure of spike trains. *Neural Comput.* 22, 121–157. doi: 10.1162/neco.2009.12-07-678

Houghton, C., and Kreuz, T. (2012). On the efficient calculation of van Rossum distances. *Netw. Comput. Neural Syst.* 23, 48–58. doi: 10.3109/0954898X.2012.673048

Houghton, C., and Victor, J. D. (2010). "Measuring representational distances–the spike-train metrics approach," in *Visual Population Codes: Toward a Common Multivariate Framework for Cell Recording and Functional Imaging*, eds N. Kriegeskorte, and G. Kreiman, (Cambridge, MA: The MIT Press), 213–250.

Jamali, M., Chacron, M. J., and Cullen, K. E. (2016). Self-motion evokes precise spike timing in the primate vestibular system. *Nat. Commun.* 7:13229. doi: 10.1038/ncomms13229

Johansson, R. S., and Flanagan, J. R. (2009). Coding and use of tactile signals from the fingertips in object manipulation tasks. *Nat. Rev. Neurosci.* 10, 345–359. doi: 10.1038/nrn2621

Kalaska, J. F., Cohen, D. A., Hyde, M. L., and Prud'homme, M. (1989). A comparison of movement direction-related versus load direction- related activity in primate motor cortex, using a two-dimensional reaching task. *J. Neurosci.* 9, 2080–2202.

Kantorovich, L. V. (1940). On one effective method of solving certain classes of extremal problems. *Dokl. Akad. Nauk USSR* 28, 212–215.

Kass, R. E., and Ventura, V. (2001). A spike-train probability model. *Neural Comput.* 13, 1713–1720. doi: 10.1162/08997660152469314

Krause, B. M., Murphy, C. A., Uhlrich, D. J., and Banks, M. I. (2017). PV+ cells enhance temporal population codes but not stimulus-related timing in auditory cortex. *Cereb. Cortex.* 29, 627–647. doi: 10.1093/cercor/bhx345

Kreuz, T., Chicharro, D., Greschner, M., and Andrzejak, R. G. (2011). Time-resolved and time-scale adaptive measures of spike train synchrony. *J. Neurosci. Methods* 195, 92–106. doi: 10.1016/j.jneumeth.2010.11.020

Kreuz, T., Chicharro, D., Houghton, C., Andrzejak, R. G., and Mormann, F. (2013). Monitoring spike train synchrony. *J. Neurophysiol.* 109, 1457–1472. doi: 10.1152/jn.00873.2012

Kreuz, T., Haas, J. S., Morelli, A., Abarbanel, H. D. I., and Politi, A. (2007). Measuring spike train synchrony. *J. Neurosci. Methods* 165, 151–161. doi: 10.1016/j.jneumeth.2007.05.031

Logiaco, L., Quilodran, R., Procyk, E., and Arleo, A. (2015). Spatiotemporal spike coding of behavioral adaptation in the dorsal anterior cingulate cortex. *PLoS Biol.* 13:e1002222. doi: 10.1371/journal.pbio.1002222

Machens, C. K., Prinz, P., Stemmler, M. B., Ronacher, B., and Herz, A. V. M. (2001). Discrimination of behaviorally relevant signals by auditory receptor neurons. *Neurocomputing* 3, 263–268. doi: 10.1016/s0925-2312(01)00382-4

MacLeod, K., Bäcker, A., and Laurent, G. (1998). Who reads temporal information contained across synchronized and oscillatory spike trains? *Nature* 395, 693–608.

Mechler, F., Victor, J. D., Purpura, K. P., and Shapley, R. (1998). Robust temporal coding of contrast by V1 neurons for transient but not for steady-state stimuli. *J. Neurosci.* 18, 6583–6598. doi: 10.1523/jneurosci.18-16-06583.1998

Narayan, R., Graña, G., and Sen, K. (2006). Distinct time-scales in cortical discrimination of natural sounds in songbirds. *J. Neurophysiol.* 96, 252–258. doi: 10.1152/jn.01257.2005

Reich, D. S., Mechler, F., and Victor, J. D. (2001). Temporal coding of contrast in primary visual cortex: when, what, and why. *J. Neurophysiol.* 85, 1039–1050. doi: 10.1152/jn.2001.85.3.1039

Rubner, Y., Tomasi, C., and Guibas, L. J. (2000). The earth mover's distance as a metric for image retrieval. *Int. J. Comput. Vis.* 40, 99–121.

Satuvuori, E., and Kreuz, T. (2018). Which spike train distance is most suitable for distinguishing rate and temporal coding? *J. Neurosci. Methods* 299, 22–33. doi: 10.1016/j.jneumeth.2018.02.009

Satuvuori, E., Mulansky, M., Bozanic, N., Malvestio, I., Zeldenrust, F., Lenk, K., et al. (2017). Measures of spike train synchrony for data with multiple time scales. *J. Neurosci. Methods* 287, 25–38. doi: 10.1016/j.jneumeth.2017.05.028

Schwartz, A. B., Kettner, R. E., and Georgopoulos, A. P. (1988). Primate motor cortex and free arm movements to visual targets in three- dimensional space. I. Relations between single cell discharge and direction of movement. *J. Neurosci.* 8, 2913–2927. doi: 10.1523/jneurosci.08-08-02913.1988

Tabuchi, M., Monaco, J. D., Duan, G., Bell, B., Liu, S., Liu, Q., et al. (2018). Clock-generated temporal codes determine synaptic plasticity to control sleep. *Cell* 175, 1213–1227. doi: 10.1016/j.cell.2018.09.016

van Rossum, M. C. W. (2001). A novel spike distance. *Neural Comput.* 13, 751–763. doi: 10.1162/089976601300014321

Vargas-Irwin, C. E., Franquemont, L., Black, M. J., and Donoghue, J. P. (2015). Linking objects to actions: encoding of target object and grasping strategy in primate ventral premotor cortex. *J. Neurosci.* 35, 10888–10897. doi: 10.1523/JNEUROSCI.1574-15.2015

Vaserstein, L. N. (1969). Markov processes over denumerable products of spaces, describing large systems of automata. *Probl. Peredachi Inf.* 5, 64–72.

Victor, J. D., and Purpura, K. P. (1996). Nature and precision of temporal coding in visual cortex: a metric-space analysis. *J. Neurophysiol.* 76, 1310–1326. doi: 10.1152/jn.1996.76.2.1310

Victor, J. D., and Purpura, K. P. (1998). Spatial phase and the temporal structure of the response to gratings in V1. *J. Neurophysiol.* 80, 554–571. doi: 10.1152/jn.1998.80.2.554

Wang, L., Narayan, R., Graña, G., Shamir, M., and Sen, K. (2007). Cortical discrimination of complex natural stimuli: can single neurons match behavior? *J. Neurosci.* 27, 582–589. doi: 10.1523/jneurosci.3699-06.2007

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.

Copyright © 2019 Sihn and Kim. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 05 February 2020
doi: 10.3389/fnins.2020.00079

Neural Synchrony Gives Rise to Amplitude- and Duration-Invariant Encoding Consistent With Perception of Natural Communication Stimuli

*Michael G. Metzen, Volker Hofmann and Maurice J. Chacron**

Computational Systems Neuroscience Laboratory, Department of Physiology, McGill University, Montreal, QC, Canada

When confronted with a highly variable environment, it remains poorly understood how neural populations encode and classify natural stimuli to give rise to appropriate and consistent behavioral responses. Here we investigated population coding of natural communication signals with different attributes (i.e., amplitude and duration) in the electrosensory system of the weakly electric fish *Apteronotus leptorhynchus*. Our results show that, while single peripheral neurons encode the detailed timecourse of different stimulus waveforms, measures of population synchrony are effectively unchanged because of coordinated increases and decreases in activity. A phenomenological mathematical model reproduced this invariance and shows that this can be explained by considering homogeneous populations whose responses are solely determined by single neuron firing properties. Moreover, recordings from downstream central neurons reveal that synchronous afferent activity is actually decoded and thus most likely transmitted to higher brain areas. Finally, we demonstrate that the associated behavioral responses at the organism level are invariant. Our results provide a mechanism by which amplitude- and duration-invariant coding of behaviorally relevant sensory input emerges across successive brain areas thereby presumably giving rise to invariant behavioral responses. Such mechanisms are likely to be found in other systems that share anatomical and functional features with the electrosensory system (e.g., auditory, visual, vestibular).

Keywords: invariance, weakly electric fish, identity-preserving transformations, neural coding, synchrony

OPEN ACCESS

Edited by:
Daya Shankar Gupta, Camden County College, United States

Reviewed by:
Yong Gu, State Key Laboratory of Neuroscience, Shanghai Institutes for Biological Sciences (CAS), China
Stephen Elisha Clarke, Stanford University, United States

****Correspondence:***
Maurice J. Chacron
maurice.chacron@mcgill.ca

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *29 October 2019*
Accepted: *20 January 2020*
Published: *05 February 2020*

Citation:
Metzen MG, Hofmann V and Chacron MJ (2020) Neural Synchrony Gives Rise to Amplitude- and Duration-Invariant Encoding Consistent With Perception of Natural Communication Stimuli. Front. Neurosci. 14:79. doi: 10.3389/fnins.2020.00079

INTRODUCTION

It is now widely accepted that behavioral responses of vertebrates to natural stimuli are determined by integrating the activities of large neuronal populations (Cohen and Kohn, 2011). However, how such integration is achieved remains poorly understood in general. This is in part because natural stimuli display complex spatiotemporal characteristics (Attias and Schreiner, 1997; Mante et al., 2005; Theunissen and Elie, 2014), as well as the fact that neuronal activities are not independent of one another (Averbeck et al., 2006). Of particular interest is the fact that neurons often fire action potentials synchronously, which is thought to enable neuronal ensembles to better encode specific stimulus features (Gray and Singer, 1989; Dan et al., 1998; Nunez and Malmierca, 2007; Uhlhaas et al., 2009; Brette, 2012; Harris and Gordon, 2015). Increased response selectivity in higher brain

areas (i.e., "sparse coding") has been observed ubiquitously (Vinje and Gallant, 2000; Laurent, 2002; Olshausen and Field, 2004; Vonderschen and Chacron, 2011; Theunissen and Elie, 2014; Sproule et al., 2015) but must be balanced by the fact that neuronal representations also become more invariant to a given sensory input encountered under different contexts (e.g., the same object under different levels of illumination) (Dicarlo and Johnson, 1999; Quiroga et al., 2005; Billimoria et al., 2008; Rust and Dicarlo, 2010; Barbour, 2011; Rust and Dicarlo, 2012; Schneider and Woolley, 2013; Sharpee et al., 2013). The mechanisms that mediate the emergence of invariant representations and the tradeoff with sparse coding remain poorly understood to this day.

Weakly electric fish generate an electric field through their electric organ discharge (EOD) and can sense perturbations through an array of electroreceptor afferents embedded in their skin (Turner et al., 1999). These afferents synapse onto pyramidal cells within the electrosensory lateral line lobe (ELL) which then project to higher brain centers that mediate behavioral responses (Rose, 2004). Natural electrosensory stimuli comprise those caused by objects such as prey (Nelson and Maciver, 1999) as well as those caused by conspecifics (Zakon et al., 2002; Metzen, 2019). In the latter case, natural electrocommunication stimuli (i.e., "chirps") consist of transient increases in EOD frequency that occur on top of the underlying sinusoidal background beat (Zupanc and Maler, 1993; Engler et al., 2000; Bastian et al., 2001; see Zupanc, 2002 for review; Kolodziejski et al., 2005). The responses of electroreceptors and pyramidal cells to natural electrocommunication stimuli have been extensively characterized (Benda et al., 2005, 2006; Marsat et al., 2009; Marsat and Maler, 2010; Vonderschen and Chacron, 2011; Walz et al., 2014; Metzen et al., 2016; Metzen and Chacron, 2017; Allen and Marsat, 2018, 2019). In particular, for single electroreceptor afferents (EAs), it has been shown that their time-dependent firing rates will vary differentially in time when chirps with different attributes (e.g., characterized by different EOD frequency increases and/or durations) are presented on top of beats with different frequencies (Benda et al., 2005, 2006; Walz et al., 2014). At the population level, it has been shown previously that the presentation of natural electrocommunication stimuli gives rise to synchrony in the responses of EAs which is primarily seen for low beat frequencies (Benda et al., 2006; Walz et al., 2014). It is important to note that Walz et al. (2014) did not systematically vary chirp duration or the EOD frequency increase in their study. For ELL pyramidal cells (PCells), it has been previously shown that they will respond differentially to chirps with different attributes through burst firing caused in part by feedback, thus enabling better signal detection (Marsat et al., 2009; Marsat and Maler, 2010, 2012, Vonderschen and Chacron, 2011). Our previous studies have considered the coding of chirps occurring on different phases of the beat and revealed the emergence of invariant neural representations based on synchronous activity at the level of EAs (Aumentado-Armstrong et al., 2015; Metzen et al., 2016; Metzen and Chacron, 2017). However, how EAs encode chirps with different durations and amplitudes (i.e., different EOD frequency increases) has not been systematically investigated to date at either the single neuron or at the population level. Further, how this information is decoded by downstream pyramidal cells to give rise to perception and behavior has not been studied to date.

Here we used a combination of electrophysiological recordings, mathematical modeling, and behavioral assays to investigate how chirps with different amplitudes and durations are represented by peripheral electroreceptor afferent neural populations. Furthermore, we analyzed how this representation is decoded by downstream central ELL pyramidal neurons that represent a bottleneck in the electrosensory pathway and whose responses are further processed by downstream brain areas to generate electrosensory perception and behavior. Our results demonstrate that synchronous activity at the afferent population level gives rise to a representation of natural electrocommunication stimuli that is invariant to variations in stimulus attributes such as duration and amplitude. This representation is decoded by ELL PCells and these responses are further processed by downstream brain areas to generate invariant behavioral responses. Because of anatomical and functional similarities between the electrosensory and other systems (Clarke et al., 2015), the uncovered mechanism for generating invariant neuronal responses is likely to be generally applicable.

MATERIALS AND METHODS

Ethics Statement

The animal study was reviewed and approved by McGill University's animal care committee under protocol number 5285.

Animals

We used a total of $N = 20$ *Apteronotus leptorhynchus* specimens of either sex in this study. Animals were acquired from tropical fish suppliers and acclimated to laboratory conditions according to published guidelines (Hitschfeld et al., 2009).

Surgery and Recordings

Surgical procedures have been described in detail previously (Toporikova and Chacron, 2009; Vonderschen and Chacron, 2011; McGillivray et al., 2012; Deemyad et al., 2013; Metzen et al., 2016). Briefly, animals ($N = 12$) were injected with tubocurarine chloride hydrate (0.1 – 0.5 mg) for immobilization before being transferred to an experimental tank and respirated with a constant flow of water over their gills (∼10 ml/min). To expose the hindbrain for recording, a portion of the animal's head was kept out of water and anesthetized locally with lidocaine ointment (5%). A small craniotomy (∼5 mm^2) was made above the hindbrain for afferent and ELL PCell recordings. We used 3M KCl-filled glass micropipettes (30 MΩ resistance) to record from electroreceptor afferent axons ($N = 60$) as they enter the ELL (Savard et al., 2011; Metzen and Chacron, 2015; Metzen et al., 2015). We recorded from single EAs in response to stimulation and then recombined the activities. This is because previous studies have shown that, as EAs do not display noise correlations, similar results were obtained when considering either simultaneous or non-simultaneous recordings

(Chacron et al., 2005a; Metzen et al., 2015, 2016). Extracellular recordings from ELL PCells within the lateral segment ($N = 40$) were performed with metal-filled micropipettes (Frank and Becker, 1964; Chacron et al., 2009; Chacron and Fortune, 2010; Metzen et al., 2016). The sample sizes are similar to those used in previous studies. Baseline (i.e., in the absence of stimulation) firing rates for EAs and PCells were 368 ± 113 Hz, and 12 ± 8 Hz, respectively, and were similar to previously reported values (Chacron et al., 2005b; Gussin et al., 2007; Metzen et al., 2015). We only recorded from neurons that responded to at least one chirp stimulus waveform. Recordings were digitized at 10 kHz (CED Power 1401 & Spike 2 software, Cambridge Electronic Design) and stored on a computer for subsequent analysis.

Stimulation

The neurogenic electric organ of *A. leptorhynchus* is not affected by injection of curare-like drugs. Stimuli consisted of amplitude modulations of the animal's own EOD and were produced by first generating a sinusoidal waveform train with frequency slightly greater (20 – 30 Hz) than the EOD frequency that was triggered by the EOD zero crossings. This train is synchronized to the animal's EOD and will either increase or decrease the EOD amplitude based on polarity and intensity. This train is then multiplied (MT3 multiplier, Tucker Davis Technologies) with an amplitude modulated waveform (i.e., the stimulus). The resultant signal is then isolated from ground (A395 linear stimulus isolator, World Precision Instruments) and delivered to the experimental tank via two chloridized silver wire electrodes located $\sim$ 15 cm on each side of the animal (Bastian et al., 2002). To elicit neural and neuronal responses, we generated chirps with different attributes by systematically varying both chirp duration (8, 11, 14, 17, and 20 ms) and amplitude (10, 35, 60, 85, and 110 Hz). These ranges were chosen to contain those observed in the current study as well as those observed in previous studies (Zupanc and Maler, 1993; Engler and Zupanc, 2001; Zupanc et al., 2006). It is important to note that the chirp amplitude is not equivalent to the actual spectral frequency content of the resulting AM stimulus which is 50–100 Hz (Zupanc and Maler, 1993). Moreover, we considered chirps occurring at either phase 90° or 270° of the beat cycle, on top of a sinusoidal beat with frequency $f_{beat} = 4$ Hz as done previously (Vonderschen and Chacron, 2011; Metzen et al., 2016). We chose a 4 Hz beat because this was the frequency used in a previous study (Metzen et al., 2016) and is characteristic of the low frequency beat stimuli encountered during interactions of two same-sex conspecifics, during which electrocommunication stimuli like those considered here occur. We chose two beat phases because our previous study has shown that EA synchrony but not single EA firing rate is invariant to different chirp waveforms with given attributes (i.e., duration and amplitude) occurring at eight different beat phases, which presumably led to invariant behavioral responses (Metzen et al., 2016). Further, we showed that ELL PCells were "locally" invariant in that they responded similarly to chirps occurring near the beat through (i.e., "+chirps") and similarly (but in opposite fashion to "+chirps") to chirps occurring near the beat peak (i.e., "−chirps") (Metzen et al., 2016). The two phases chosen here correspond to representative examples of "+chirps" and "−chirps" that will effectively capture variations in neural responses due to chirps occurring at different beat phases. To measure the stimulus intensity, a small dipole was placed close to the animal's skin. Stimulus intensity was adjusted to produce changes in EOD amplitude that were $\sim$20% of the baseline level, as done previously (Metzen et al., 2016; Metzen and Chacron, 2017). Finally, each chirp stimulus (i.e., a chirp with given duration and amplitude) was presented at least 20 times (i.e., 20 trials) in order to average the variability of neural responses.

Modeling

We used the leaky integrate and fire model with dynamic threshold (LIFdt) (Chacron et al., 2000, 2001) that is an extension of the Nelson model using the following set of differential equations to account for various filtering mechanisms (Bastian, 1981; Nelson et al., 1997):

$$\dot{S}_a = -\frac{S_a}{\tau_a} + \frac{G_a}{\tau_a} A(t), \tag{1}$$

$$\dot{S}_b = -\frac{S_b}{\tau_b} + \frac{G_b}{\tau_b} A(t), \tag{2}$$

$$S(t) = -S_a - S_b + (G_a + G_b + G_c) A(t), \tag{3}$$

where $A(t)$ is the stimulus, and $S(t)$ is the filtered stimulus. The G values are gains in units of spikes per second per millivolt ($G_a = 18300\ spikes \times s^{-1} \times mV^{-1}$; $G_b = 850\ spikes \times s^{-1} \times mV^{-1}$; $G_c = 670\ spikes \times s^{-1} \times mV^{-1}$), and the τ values are time constants in units of seconds ($\tau_a = 0.002$ s; $\tau_b = 0.25$ s). The total dimensionless synaptic current arriving at the spike initiation zone is given by:

$$I_{syn} = S(t) + \gamma A_0 + \xi(t), \tag{4}$$

where I_{syn} is the synaptic current, $S(t)$ is the filtered stimulus according to equations (1–3), γ and A_0 are constants. $\xi(t)$ is Gaussian white noise with zero mean and variance of one. In the time window after the absolute refractory period and up to the next action potential, the voltage V and the threshold θ are given by:

$$\dot{V} = -\frac{V}{\tau_v} + \frac{I_{syn}}{\tau_v}, \tag{5}$$

$$\dot{\theta} = \frac{(\theta_0 - \theta)}{\tau_\theta}, \tag{6}$$

where V is the membrane potential, τ_v is the voltage decay constant of the membrane, θ is the spike threshold, and τ_θ is the threshold decay constant. Whenever $V = \theta$, a spike occurs, and V is reset to zero and maintained there for the duration of the refractory period (T_r). The threshold θ is also increased by a fixed amount $\Delta\theta$ and otherwise decays with time constant τ_θ between action potentials. Parameter values used were $\tau_v = 1$ ms; $\tau_\theta = 7.75$ ms; $\theta_0 = 0.08$; $\Delta\theta = 0.001$; $T_r = 1$ ms. Parameter values were chosen based on previous studies (Chacron et al., 2000, 2001; Savard et al., 2011) and were adjusted such that the mean firing rate of our model (392.71 ± 0.02 Hz) was within the experimentally observed range. As such, our model neurons

were homogeneous and the spiking activities of the model neuron stimulated in the same way as our experimental data (i.e., same number of trials and trial length) were used to compute all measures at the single neuron level. The spiking activities of two model neurons with independent realizations of the noise $\xi(t)$ were used to compute all measures at the population level.

Analysis

All analyses were performed using custom-built routines in Matlab (The MathWorks Inc., Natick, MA, United States), these routines are freely available online at http://dx.doi.org/10.6084/m9.figshare.8041136.

Electrophysiology

We used a total of $N = 12$ animals of either sex for electrophysiological recordings (EAs: $N = 5$; PCells: $N = 7$). ELL PCells were recorded within the lateral segment (LS) of the ELL where cells are most sensitive to high frequency communication signals (Marsat et al., 2009). This segment contains about 900 PCells, each receiving convergent input from about 1000 EAs on average (Maler, 2009). Action potential times were defined as the times at which the signal crossed a suitably chosen threshold value. From the spike time sequence, we created a binary sequence $R(t)$ with binwidth $\Delta t = 0.1$ ms and set the content of each bin to equal the number of spikes which fell within that bin. The time-dependent firing rates were obtained by averaging the neural or neuronal responses across repeated presentations of a given stimulus with binwidth 0.1 ms and were smoothed with a 6 ms long boxcar filter. We note that similar results were obtained when systematically varying the size of the boxcar filter between 6.25 ms and 250 ms (**Figures 3**, **4F**, **7**, **8G**).

Synchrony Between the Spiking Activities of Electrosensory Afferents

To quantify neural synchrony, we computed the cross-correlation coefficient between the spiking responses $R_i(t)$ and $R_j(t)$ of neurons i and j as was done previously (Shea-Brown et al., 2008; Metzen et al., 2015, 2016; Metzen and Chacron, 2017). As mentioned before, we randomly combined electrosensory afferents to compute synchrony, as these do not display noise correlations (Chacron et al., 2005a; Metzen et al., 2015). The time varying spiking synchrony was computed as the correlation coefficient between spike count sequences S_i obtained from the binary sequences for non-overlapping 5 ms bins during a time window of 31.25 ms that was translated in steps of 0.25 ms using:

$$\rho = \frac{Cov(S_1, S_2)}{\sqrt{Var\,(S_1)\;Var(S_2)}} \tag{7}$$

Here, $Cov(\ldots)$ is the covariance while $Var(\ldots)$ denotes the variance, and S_1, S_2 are the spike count sequences from neurons 1 and 2, respectively. The time-dependent synchrony measures were then averaged across trials. We note that similar results were obtained when systematically varying the time window length between 6.25 ms and ∼60 ms but that synchrony values decreased for longer lengths up to 250 ms (**Figures 5G**, **6G**).

Quantifying Neural Response Invariance

The invariance score for either parameter (i.e., duration or amplitude) was defined as (Metzen et al., 2016; Metzen and Chacron, 2017):

$$Invariance = 1 - \frac{\sum_{i \neq j}\left[\frac{D\left(FR_i(t), FR_j(t)\right)}{D\left(S_i(t), S_j(t)\right)}\right]}{N_{chirps}\left(N_{chirps} - 1\right)}, \tag{8}$$

where $N_{chirps} = 10$ and the sum runs over indices i and j representing different values of the parameter (i.e., duration or amplitude) for all possible combinations of i ≠ j. $D(x,y)$ is a distance metric between x and y that was computed as (Aumentado-Armstrong et al., 2015; Metzen et al., 2016; Metzen and Chacron, 2017):

$$D(x, y) = \frac{\sqrt{\langle (x - \langle x \rangle - y + \langle y \rangle)^2 \rangle}}{\max\left[\frac{\max(x) - \min(x)}{\sqrt{2}}, \frac{\max(y) - \min(y)}{\sqrt{2}}\right]}, \tag{9}$$

where $< \ldots >$ denotes an average over an evaluation window of 30 ms after chirp onset that is shown as a gray band in the figures, $FR_i(t)$ is the peri-stimulus time histogram (PSTH) response of a given cell to chirp stimulus waveform $S_i(t)$, and max(. . .), min(. . .) denote the maximum and minimum values, respectively. All responses were normalized prior to computing the distance metric. We note that, according to equation (8), the distance between responses to two different stimulus waveforms is normalized by the distance between the stimulus waveforms themselves. A value of one indicates perfect invariance, whereas a value of zero indicates that a neuron whose response faithfully encodes the detailed timecourse of the different stimulus waveforms will not be considered invariant according to our definition. It is important to note that, unlike the detectability measure described below, our invariance measure is based on the timecourse of the actual neural responses and not solely on their minimum and maximum values. Thus, in order to obtain a high invariance score, it is not sufficient for different neural responses to merely have the same minimum and maximum values, they actually have to have a similar timecourse. It is furthermore important to note that the invariance score was computed from the PSTH responses which are averaged over trials to reduce variability. It is thus unlikely that the invariance scores reported in the current study are due to large response variability. Invariance scores were computed for each individual cell and subsequently averaged across the respective populations. We computed duration and amplitude invariance for synchronous activity as described above except that we used the timecourse of the varying correlation coefficient instead of spike counts as an input.

Detectability

To determine the detectability of a stimulus waveform resulting from a chirp with a specific amplitude or duration within the ongoing beat, we computed the distance $D(x,y)$ [equation

(9)] between the chirp waveform and the corresponding beat waveform (i.e., the beat waveform when no chirp occurred) as done previously (Aumentado-Armstrong et al., 2015; Metzen and Chacron, 2017). A value of one indicates perfect detectability, whereas a value of zero indicates that the chirp waveform is identical to the beat waveform. The neuronal detectability of a chirp (using either single unit firing rate or synchronous activity) was computed using:

$$Detectability_{neuronal} = abs\left(\frac{R_{chirp} - R_{beat}}{R_{chirp} + R_{beat}}\right), \qquad (10)$$

where $R_{chirp} = R_{max} - R_{min}$ (i.e., the difference between the maximum and minimum values of the response). R_{chirp} was computed over a time window of 15 ms for EA firing rate and of 60 ms for EA synchrony and PCell firing to account for differences in the timecourse of responses as done previously (Metzen and Chacron, 2017). $R_{beat} = R_{max} - R_{min}$ is the difference between the maximum and minimum values of the response (i.e., either of EA firing rate, EA synchrony, or PCell firing rate) to the undisturbed beat during one beat cycle, respectively. We note that this measure is similar to the chirp selectivity index used in previous studies (Vonderschen and Chacron, 2011; Aumentado-Armstrong et al., 2015).

Behavior

Apteronotus leptorhynchus has been shown to robustly give chirp echo responses when stimulated with chirps (Hupé and Lewis, 2008; Gama Salgado and Zupanc, 2011). Measuring this chirp echo response has been effectively used before to infer the perceptual abilities of these animals under different stimulus conditions (Metzen et al., 2016; Metzen and Chacron, 2017). Moreover, chirping behavior was shown to be identical in freely moving and restrained fish (Hitschfeld et al., 2009). We therefore measured the chirp echo behavioral response by restraining fish ($N = 8$) in a "chirp chamber" as described previously (Metzen and Chacron, 2014; Metzen et al., 2016). Stimuli were delivered by two electrodes spaced 10 cm from each other located on the right side of the animal (**Figure 9A**, S1 and S2). The EOD was measured between electrodes placed near the head and tail, amplified (Axoclamp 2B, Molecular Devices), digitized at 10 kHz sampling rate using CED 1401plus hardware and Spike2 software (Cambridge Electronic Design), and stored on a computer hard disk for offline analysis. Previous studies have shown that stimulation with low frequency (<10 Hz) beats will induce chirping behavior but that this habituates over time (Bastian et al., 2001). As such, we initially habituated the animal to a 4 Hz beat stimulus lasting 60 s in order to minimize the probability of chirp responses being elicited due the beat alone. Computing the baseline chirp rate during the first (control) and last (habituated) 30 s of the habituation period showed a significant drop in chirp rate down to 0 (control: 0.5919 ± 1.137 chirps $\times$ s^{-1}; habituated: 0 ± 0 chirps $\times$ s^{-1}; paired t-test; $p = 1.28 \times 10^{-5}$), indicating that the animals were habituated to the beat signal. It is therefore highly unlikely that any echo response observed after stimulus chirp onset was due to the underlying beat as the chirp rate immediately before stimulus chirp onset was null. We then randomly interspersed chirp stimuli at variable intervals ($15 \text{ s} \pm 3 \text{ s}$) and the recording was started 200 ms before chirp onset. To analyze the chirp echo response, we first extracted the time varying EOD frequency of each fish tested. Echo response chirps after stimulus chirp onset were identified as increases in the animal's own EOD frequency that exceeded 30 Hz (Bastian et al., 2001). The time of occurrence of echo response chirps was defined as the time at which the EOD frequency excursion was maximal. The echo response chirp rate was computed as the number of echo response chirps during a time window of 1 s following the stimulus chirp onset since previous studies have shown that the majority of responses occur during this time window (Zupanc et al., 2006). Invariance scores for behavior were computed as described above for neural and neuronal responses except that we used the behavioral PSTHs computed from the echo responses using a 1 s boxcar filter as responses (Metzen et al., 2016). We note that we randomly varied the beat phase at which the chirp occurred between 0° and 315° in increments of 45° for either of the two chirp parameters (i.e., duration and amplitude) to avoid habituation.

Chirp Statistics

In order to quantify the distribution of chirp attributes duration and amplitude in naturally occurring electrocommunication signals, we analyzed the chirps elicited by our fish population ($N = 8$) during the habituation period to a 4 Hz beat stimulus of 60 s duration. To do so, we extracted the time-varying EOD frequency by computing the inverse of the timing difference between successive zero crossings as done previously (Metzen et al., 2016). Chirp amplitude was computed as the difference between the baseline EOD amplitude and the maximal EOD frequency during a chirp event. The chirp duration was defined as the full width at half-maximum of the EOD frequency excursion. The time of occurrence of the chirp was defined as the time at which the EOD frequency is maximal (Aumentado-Armstrong et al., 2015). As mentioned above, previous studies have shown that, in response to stimulation with low frequency (<10 Hz) beats, animals will emit chirps. However, these studies have also shown that the characteristics of the emitted chirps (e.g., duration and amplitude) will depend on the stimulation protocol such as the beat frequency as well as stimulus intensity (Zupanc and Maler, 1993; Bastian et al., 2001; Gama Salgado and Zupanc, 2011). In contrast, our measurements of chirp statistics were made under the same conditions (i.e., same beat contrast and frequency) than those used to investigate neural and echo responses to these, thereby making them more directly comparable.

Statistics

Statistical significance was assessed through a paired t-test or a one-way analysis of variance (ANOVA) with the Bonferroni method of correcting for multiple comparisons at the $p = 0.05$ level. Values are reported as boxplots unless otherwise stated. Errorbars indicate mean $\pm$ SD. On each box, the central mark indicates the median, and the bottom and top edges of the box indicate the 25th and 75th percentiles, respectively. Outliers are

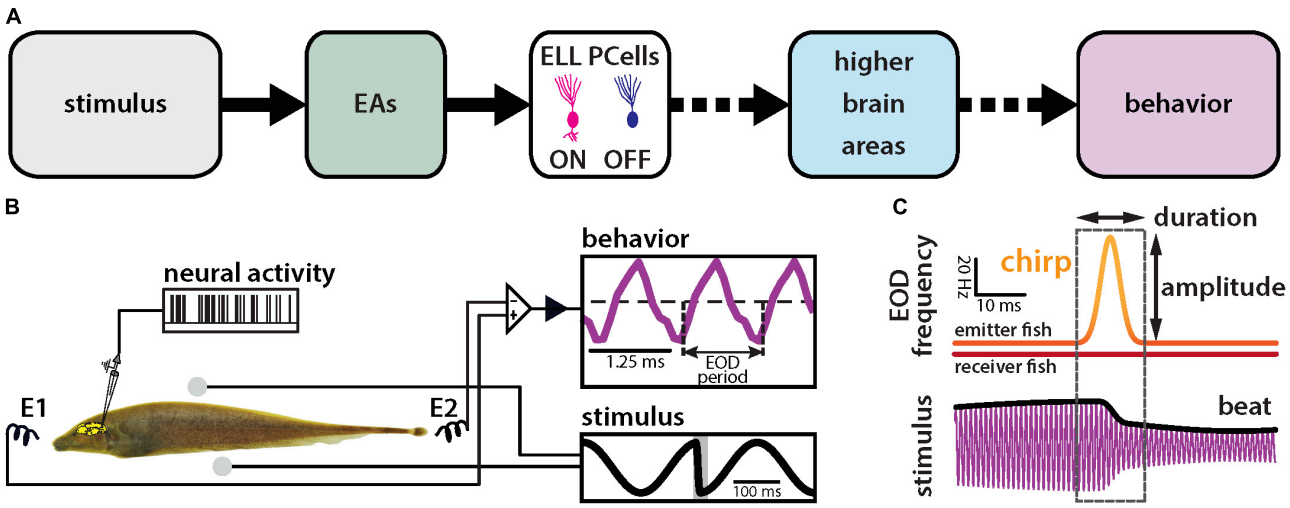

FIGURE 1 | Schematic of the experimental setup. **(A)** Schematic showing successive brain areas involved in the processing of electrosensory stimuli. **(B)** Experimental setup: the animal's electric field (i.e., the behavior, purple) is monitored by a pair of electrodes located in front and behind the animal (E1 and E2) while neural and neuronal activity is recorded. The stimulus (black) is delivered using a separate set of electrodes positioned on each side (spheres). The shaded gray rectangle in the lower right inset shows the timewindow used for analysis. **(C)** During a chirp, the emitter fish's EOD frequency (top orange trace) is transiently increased by a maximum of amplitude for a brief duration (dashed box) while the receiver fish's EOD frequency (red trace) remains constant. This can be characterized by the duration and amplitude of the frequency excursion. The chirp results in a phase reset of the beat (bottom black trace).

plotted individually using the " symbol. 95% confidence intervals were estimated using the t-distribution.

RESULTS

We investigated how natural electrocommunication signals (i.e., "chirps") with different attributes are encoded by peripheral electroreceptor afferents (EAs) and their downstream target pyramidal cells (PCells) within the ELL to give rise to behavior (**Figure 1A**). To do so, we used an immobilized preparation in which neural, neuronal and behavioral responses can be recorded simultaneously (**Figure 1B**). Under natural conditions, chirps occur during social interactions in which the emitter fish sends the signal to the receiver fish. This signal consists of a transient increase in the emitter's fish's EOD frequency with given time duration and amplitude (i.e., the amount by which the EOD frequency increases; **Figure 1C**, top panel). Interactions between the two fish's EOD frequencies gives rise to a sinusoidal background beat (**Figure 1C**, bottom panel, black). The chirp signal perturbs the underlying beat amplitude when considering the stimulus sensed by the receiver fish (**Figure 1C**, bottom panel, black).

We first investigated how chirp amplitude and duration were distributed in emitted chirps of our fish population used for behavior (N = 8). To do so, fish were stimulated with a background beat of 4 Hz, which is characteristic of the low frequency stimuli encountered during natural interactions between same-sex conspecifics, and the resulting chirps were detected and analyzed (see section "Materials and Methods"). While previous studies have shown that fish will emit chirps when stimulated by beats alone (Zupanc and Maler, 1993; Bastian et al., 2001; Gama Salgado and Zupanc, 2011), these have shown that the attributes of the emitted chirps (e.g., their duration and amplitude) can vary based on beat attributes such as frequency (Bastian et al., 2001). Thus, in order to ensure that our results can be directly comparable, we measured the characteristics of emitted chirps (i.e., amplitude, duration) using the same beat stimulus (i.e., same frequency and amplitude) that was used to stimulate neurons and behavioral echo responses as described below.

We found that, for both duration (**Figure 2A**) and amplitude (**Figure 2B**), the distributions were relatively narrow (duration: mean: 11.45 ms, SD: 0.75 ms; amplitude: mean: 39.1 Hz, SD: 8.0 Hz). These results are consistent with previous ones (Bastian et al., 2001). We further found that the different stimulus waveforms resulting from chirps with different duration and amplitude became progressively more different from the background beat itself (**Figures 2C,D**), consistent with previous findings (Benda et al., 2005; Walz et al., 2014). We quantified these differences by computing stimulus detectability and found larger values for higher values of either duration (**Figure 2E**) or amplitude (**Figure 2F**). To better understand experimental results, we used a phenomenological mathematical model of EA activity that closely reproduces experimental results (see section Materials and Methods). The model consists of a leaky integrate and fire formalism with dynamic threshold (LIFdt) for spike generation to which input in the form of the filtered stimulus based on single neuron properties found experimentally (Xu et al., 1996), noise, and a constant bias are given (**Figure 2G**). The model afferents were simulated using independent sources of noise (see section Materials and Methods), which assumes that there are no noise correlations and is consistent with available experimental data (Chacron et al., 2005a; Metzen et al., 2016).

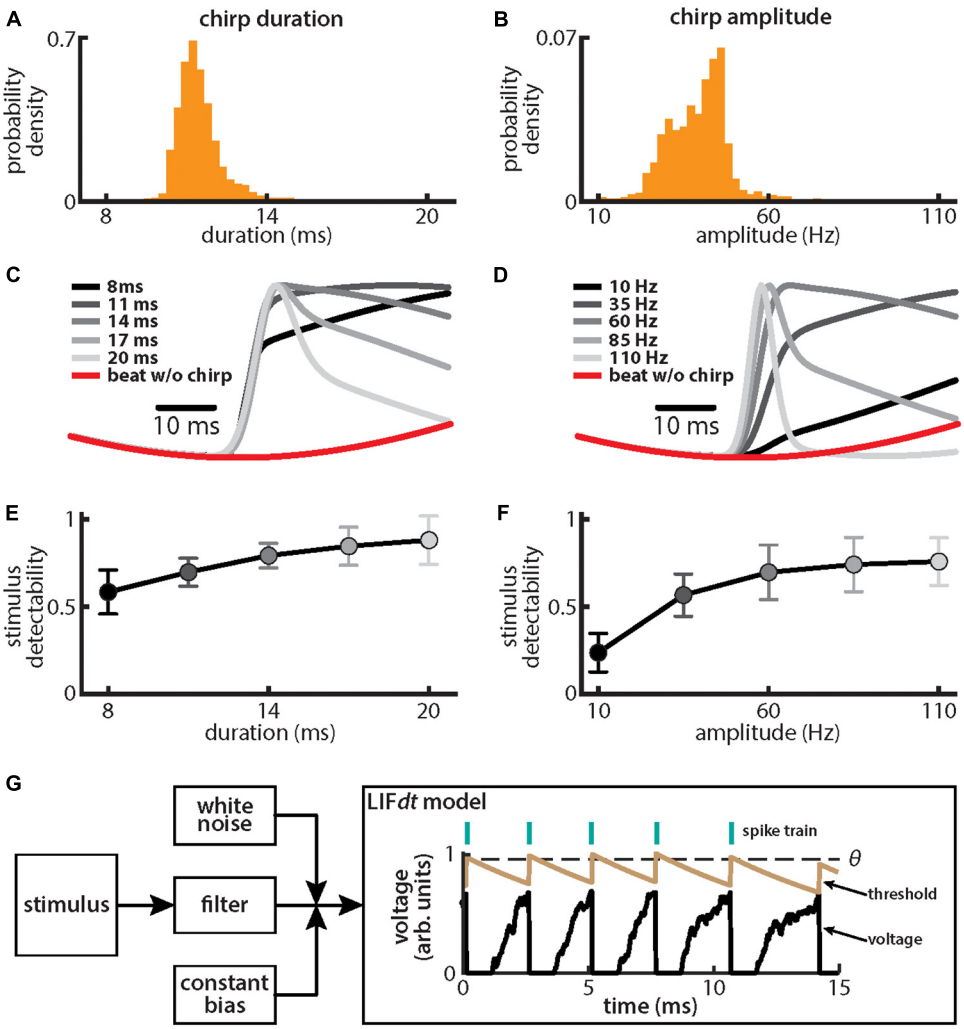

FIGURE 2 | Chirps with different durations and amplitudes give rise to heterogeneous waveforms. **(A)** Probability density for chirp duration measured from chirps that were emitted by animals when stimulated with a 4 Hz beat stimulus. **(B)** Same as **(A)**, but for chirp amplitude. **(C)** Chirp waveforms for different durations, but a fixed amplitude (shaded gray). Also shown is the stimulus waveform of the beat without a chirp (red). Values are plotted as mean ± SE. **(D)** Same as **(C)**, but for chirp amplitude. **(E)** Chirp stimulus detectability as a function of duration. Values are plotted as mean ± SE. **(F)** Same as **(E)**, but for amplitude. **(G)** Schematic of the leaky integrate and fire model with dynamic threshold (LIFdt) where a stimulus is passed to a filter with and a white noise term as well as a current bias is added. The voltage (black curve) and threshold (brown curve) trace obtained with the LIFdt model showing the firing rule. When voltage becomes greater than the threshold θ, a spike is said to have occurred, and the voltage is reset to zero, whereas threshold is incremented by a constant $\Delta\theta$. The threshold is kept constant to simulate the absolute refractory period T_r (equal to one EOD cycle) and then decays exponentially with time constant τ_θ to its equilibrium value θ_0. Parameter values used are given in the section "Materials and Methods."

Single Peripheral Afferents Respond Differentially to Natural Electrocommunication Stimuli With Different Durations and Amplitudes

We first investigated how chirps with different durations were encoded by single EAs (**Figure 3**). We found that responses to these consisted of patterns of increases and decreases in firing activity that faithfully encoded the stimulus waveform (**Figures 3A,B**; green dots showing raster plots and green curves showing the trial-averaged firing rate), consistent with previous results (Benda et al., 2005, 2006; Walz et al., 2014). EA firing activity increased when the chirp waveform (**Figure 3A**, black) occurred near the beat trough (**Figure 3A**, green) but instead decreased when the chirp waveform occurred near the beat peak (**Figure 3B**, green). Superimposing the different responses emphasized differences (**Figure 3C**, top green). Simulations of our LIFdt model's response to the different waveforms were in good qualitative agreement with experimental data (**Figures 3A–D**, cyan). Overall, stimulus detectability computed from single EA responses increased with increasing duration (**Figure 3D**, green, ANOVA with Bonferroni correction, $p = 3.311 \times 10^{-173}$, df = 295, $n = 60$) but were much lower than those computed from the stimulus (**Figure 3D**, black), which

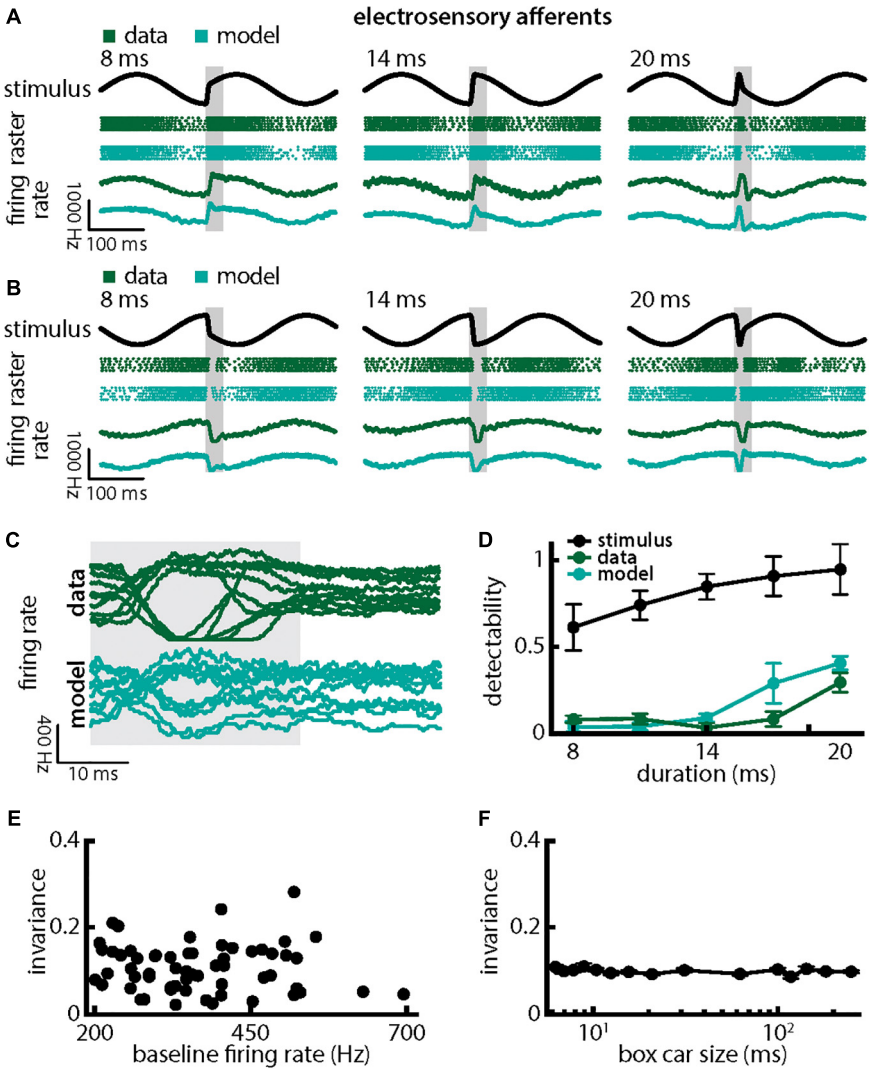

FIGURE 3 | Single peripheral electroreceptor afferents respond differentially to chirps with different durations. **(A)** Example stimulus waveforms (top, black) for chirps with different durations (left: 8 ms; middle: 14 ms; right: 20 ms) occurring at the same phase of a 4 Hz beat, raster plots of an example afferent (middle top, green) and model neuron (middle bottom, cyan) showing responses to 5 out of 20 randomly chosen presentations (i.e., trials) of each stimulus and the corresponding firing rates of both neurons averaged over all 20 trials (bottom). The horizontal bars (shaded gray) represent the chirp window used for evaluation. The gray band shows the evaluation time window used to compute invariance (see below). **(B)** Same as **(A)**, but when the chirp occurred at a different phase of the beat. **(C)** Superimposed trial-averaged firing rate responses of an example afferent (green) and our model (cyan) to chirps of different durations. **(D)** Population-averaged detectability values computed from firing rate responses of the afferents (green) and the model (cyan) as a function of duration. Also shown is the stimulus detectability (black) as a function of duration. **(E)** Invariance as a function of the baseline (i.e., in the absence of stimulation but in the presence of the animal's unmodulated EOD) firing rate for our afferent dataset. No significant correlation was observed (Pearson's correlation coefficient, $r = -0.0483$; $p = 0.7139$). **(F)** Population-averaged invariance as a function of the boxcar low-pass filter size used to obtain the time dependent firing rate from spiking activity.

is due to the fact that EA responses to chirps with different durations differed from one another rather than due to variability between individual responses to a given chirp. This is because the invariance measure is computed using the trial-averaged time dependent firing rates (i.e., firing rates averaged over repeated presentations of the stimulus waveform associated with a chirp with given duration and amplitude) rather than single-trial responses which are more variable (see Section Materials and Methods). Detectability computed from our model's responses closely matched values from experimental data (**Figure 3D**, compare green and cyan). Afferent heterogeneities as quantified by the baseline firing rate (i.e., in the absence of stimulation) did not affect invariance as no significant correlation was observed (**Figure 3E**, Pearson's correlation coefficient, $r = -0.0811$; $p = 0.3238$). Finally, our invariance results were robust to changes in filter settings used to obtain the trial-averaged time-dependent firing rate from spiking activity (**Figure 3F**).

Qualitatively similar results were obtained when we varied chirp stimulus amplitude (**Figure 4**). Responses consisted of patterns of increases and decreases in firing activity that

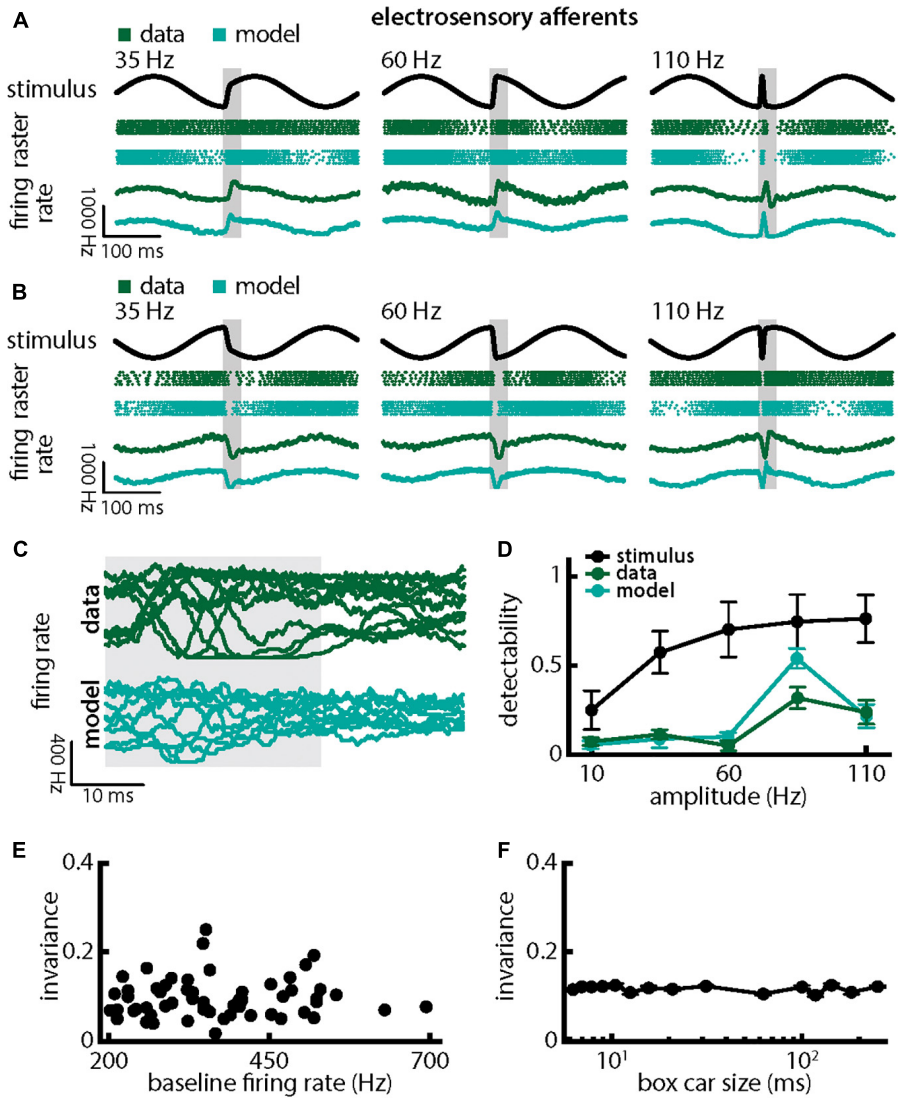

FIGURE 4 | Single peripheral electroreceptor afferents respond differentially to chirps with different amplitudes. **(A)** Example stimulus waveforms (top, black) for chirps with different amplitudes (left: 35 Hz; middle: 60 Hz; right: 110 Hz) occurring at the same phase of a 4 Hz beat, raster plots of an example afferent (middle top, green) and model neuron (middle bottom, cyan) showing responses to 5 out of 20 randomly chosen presentations (i.e., trials) of each stimulus and the corresponding firing rates of both neurons averaged over all 20 trials (bottom). The horizontal bars (shaded gray) represent the chirp window used for evaluation. The gray band shows the evaluation time window used to compute invariance (see below). **(B)** Same as **(A)**, but when the chirp occurred at a different phase of the beat. **(C)** Superimposed trial-averaged firing rate responses of an example afferent (green) and our model (cyan) to chirps of different amplitudes. **(D)** Population-averaged detectability values computed from firing rate responses of the afferents (green) and the model (cyan) as a function of amplitude. Also shown is the stimulus detectability (black) as a function of amplitude. **(E)** Invariance as a function of the baseline (i.e., in the absence of stimulation but in the presence of the animal's unmodulated EOD) firing rate for our afferent dataset. No significant correlation was observed (Pearson's correlation coefficient, $r = 0.0386$; $p = 0.7694$). **(F)** Population-averaged invariance as a function of the boxcar size used to obtain the time dependent firing rate from spiking activity.

faithfully encoded the stimulus waveform (**Figures 4A,B**; green). Superimposing the different responses again emphasized differences (**Figure 4C**, top, green). Detectability also increased with increasing amplitude (**Figure 4D**, green curve; ANOVA with Bonferroni correction, $p = 9.61 \times 10^{-133}$, df = 295, $n = 60$). Results obtained from numerical simulations of our model were in good qualitative agreement with experimental data overall (**Figures 4A–D**, cyan). Afferent heterogeneities as quantified by the baseline (i.e., in the absence of stimulation) firing rate also did not affect invariance as no significant correlation was observed (**Figure 4E**, Pearson's correlation coefficient, $r = 0.0386$; $p = 0.7694$). Finally, our invariance results were robust to changes in filter settings used to obtain the trial-averaged time-dependent firing rate from spiking activity (**Figure 4F**). Thus, we conclude that single peripheral afferents respond differentially to the different stimulus waveforms associated with changes in both duration and amplitude for natural electrocommunication stimuli.

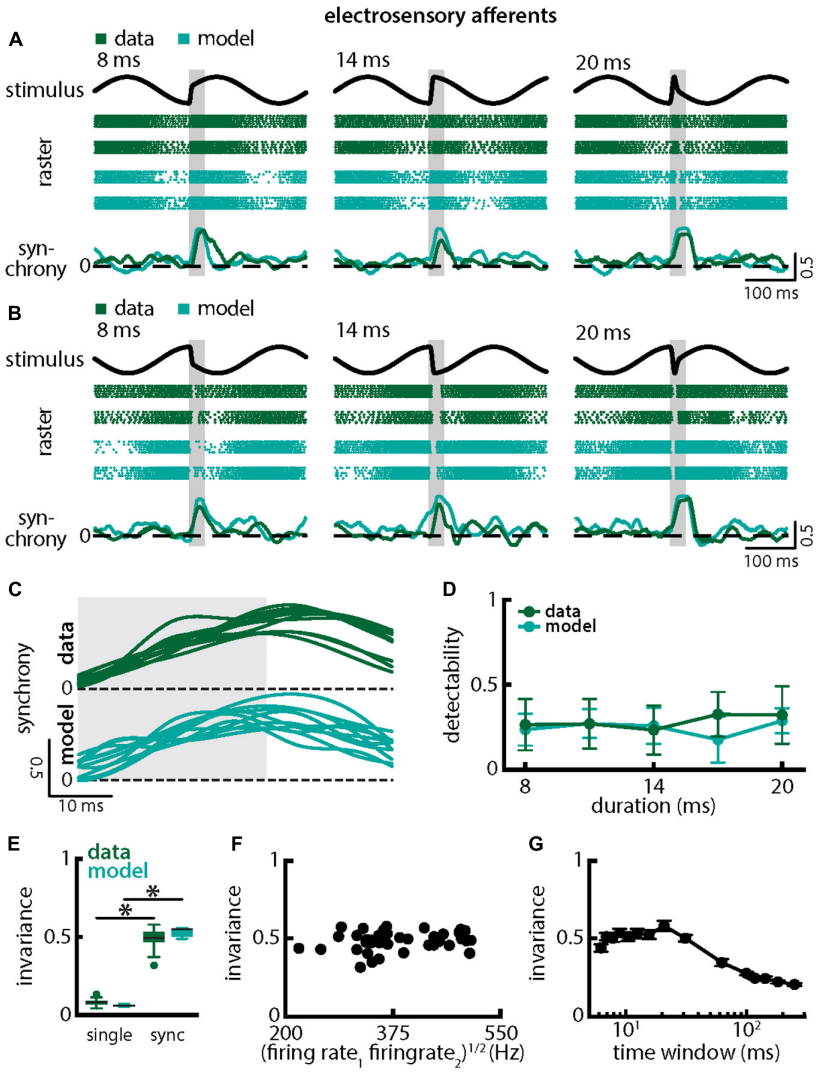

FIGURE 5 | Synchrony provides an invariant representation of chirps with different durations. **(A)** Example stimulus waveforms (top) for chirps with different durations (left: 8 ms; middle: 14 ms; right: 20 ms) occurring at the same phase of the 4 Hz beat, raster plots of two example afferents (middle top, green) and model neurons (middle bottom, cyan) showing responses to 5 out of 20 randomly chosen presentations (i.e., trials) of each stimulus and the time varying synchrony averaged over all 20 trials (bottom) from the shown example afferent (green) and model (cyan) pairs. The horizontal bars (shaded gray) represent the chirp window used for evaluation. The gray band shows the evaluation time window used to compute invariance (see below). **(B)** Same as **(A)**, but when the chirp occurred at a different phase of the beat. **(C)** Superimposed trial-averaged synchrony responses from the same example pair of afferents for experimental data (green) and from a pair of model afferents (cyan) for chirps of different durations. **(D)** Population-averaged neuronal detectability values computed from the spiking synchrony from the afferents (green) and the model (cyan) as a function of duration. **(E)** Population-averaged invariance values computed for the single afferents (green) and for the model (cyan) from single afferent activity (left) and from synchrony (right) for chirp duration. "*" indicates statistical significance at the $p = 0.05$ level using a paired t-test. **(F)** Invariance as a function of the geometric mean of the afferent baseline firing rates for our dataset. No significant correlation was observed (Pearson's correlation coefficient; $r = 0.1924$; $p = 0.2343$). **(G)** Invariance as a function of time window length. Invariance was more or less independent of time window length for values up to ~60 ms.

Afferent Populations Respond With Similar Increases in Synchrony to Stimuli With Different Durations and Amplitudes and Thus Provide an Invariant Representation of Both Stimulus Attributes

We next investigated how afferent populations encode natural electrocommunication stimuli with varying duration (**Figure 5**). Our results show that the spiking activities of afferent pairs were more synchronized in response to all stimulus waveforms (**Figures 5A,B**, green). We thus quantified the time-varying synchrony from pair-wise correlations between afferent activities which ranges between -1 (perfect anti-synchrony) and 1 (perfect synchrony) with 0 indicating lack of synchrony (see section Materials and Methods). It is important to note that the synchrony measure was averaged over trials (i.e., repeated presentations of the stimulus waveform associated with a chirp

with given duration and amplitude, see section Materials and Methods) in order to ensure that changes are not due to trial-to-trial variability in the neural responses. We found that synchronous activity was much higher when a chirp had occurred than during the background beat. Synchrony transiently increases in a similar fashion in response to all chirps of different durations and irrespective of whether the stimulus occurred at the beat peak or trough (**Figures 5A,B**, bottom panels, 5C, top green). Overall, synchrony at the population level was a much better detector of the stimulus than the single afferent activity, as quantified by higher detectability values especially for lower durations (compare **Figure 5D** and **Figure 3D**, green; ANOVA with Bonferroni correction, $p = 0.002$, df = 295, $n = 60$). We quantified invariance (see section Materials and Methods) using both the single afferent activity as well as the synchrony measure and found significantly higher values for the latter (single neuron: mean: 0.08 ± 0.02 SD; max: 0.13; min: 0.04; synchrony: mean: 0.49 ± 0.06 SD; max: 0.58; min: 0.32; $p = 7.791 \times 10^{-42}$, *t*-test; **Figure 5E**). Similar results were observed when systematically varying the time scale at which synchrony was computed up to ∼60 ms (**Figure 5G**). Results obtained from simulations of our model at the population level were in good qualitative agreement with experimental data (**Figures 5A–E**, compare green and cyan throughout). Our model shows that the experimentally observed invariance of synchrony at the population level can be explained by the temporal filtering properties observed in electroreceptor afferents and further suggests that EA heterogeneities are not necessary to observe synchrony in EA pairs. Confirming this prediction, afferent heterogeneities as quantified by the geometric mean of the baseline firing rates of each pair did not affect invariance as no significant correlation was observed (**Figure 5F**, Pearson' correlation coefficient, $r = 0.1924$; $p = 0.2343$). We conclude that synchronous activity at short timescales in receptor afferents displays invariance to variations in chirp duration.

Qualitatively similar results were obtained when investigating changes in chirp stimulus amplitude (**Figure 6**). The spiking activities of afferents were always more synchronized following the stimulus presentation (**Figures 6A,B**, green) thereby giving rise to similar increases in the synchrony measure (**Figures 6A,B**, bottom panels, green, 6C, top green). Stimulus detectability was also higher when considering synchrony than single neuron activity (**Figure 6D**, ANOVA with Bonferroni correction, $p = 7.435 \times 10^{-9}$, df = 295, $n = 60$). As chirps with different amplitude all gave rise to increases in synchrony with similar a timecourse (**Figure 6C**, top green), invariance was larger than when considering single neuron activity (**Figure 6E**, synchrony: mean: 0.49 ± 0.06 SD; max: 0.58; min: 0.32; amplitude: mean: 0.37 ± 0.07 SD; max: 0.54; min: 0.21; single neuron: mean: 0.06 ± 0.02 SD; max: 0.09; min: 0.04; $p = 6.672 \times 10^{-30}$, *t*-test). Similar results were observed when systematically varying the time scale at which synchrony was computed up to ∼60 ms (**Figure 6G**). Finally, results from modeling were in good qualitative agreement with experimental data (**Figures 6A–E** compare green and cyan throughout). Our model further confirms that the experimentally observed invariance of synchrony at the population level can be explained by the temporal filtering properties observed in electroreceptor afferents and further suggests that EA heterogeneities are not necessary to observe synchrony in EA pairs. Indeed, afferent heterogeneities as quantified by the geometric mean of the baseline firing rates of each pair did not affect invariance as no significant correlation was observed (**Figure 6F**, Pearson's correlation coefficient, $r = -0.2148$; $p = 0.1832$). We conclude that synchronous activity at short timescales in receptor afferents displays invariance to variations in chirp amplitude.

Single ELL Pyramidal Neuron Responses Decode Synchronous Afferent Activity as Their Responses Are More Invariant to Duration and Amplitude Than Those of Single Afferents

So far, our results have shown that, while single afferents respond differentially to natural electrocommunication stimuli with different durations and amplitudes, this is not the case when looking at the population level. This is because their activities are more synchronized irrespective of duration or amplitude, which leads to an invariant representation of electrocommunication stimuli. Information transmitted by neural activity is of course only useful if it is actually decoded by downstream neurons. As such, we next investigated the responses of ELL PCells that receive input from afferents to natural electrocommunication stimuli with different durations and amplitudes. PCells can be classified as either ON or OFF-type based on whether they respond to increased stimulation with increases or decreases in firing rate, respectively (Saunders and Bastian, 1984).

When varying chirp duration (**Figure 7**), we found that single ELL PCells responded similarly to stimuli occurring at a given background beat phase (**Figures 7A,B**). Specifically, ON-type cells responded with increases in firing rate that were largely independent of chirp duration when the stimulus occurred at the beat trough (**Figure 7A**, magenta). In contrast, OFF-type cells responded with decreases in firing rate that were largely independent of duration for these stimuli (**Figure 7A**, blue). When the chirp stimulus instead occurred at the beat peak, the situation was reversed as ON-type cells responded with decreases in firing rate (**Figure 7B**, magenta) while OFF-type cells responded instead with increases in firing rate (**Figure 7B**, blue) that were in both cases largely independent of chirp duration. This is best seen by superimposing the different responses (**Figure 7C**). Stimulus detectability computed from ELL PCell activity was qualitatively similar for ON- and OFF-type cells (**Figure 7D**). This detectability was furthermore similar to that computed from afferent synchrony and thus significantly higher than that computed from single afferent activity (ON-type: $p = 5.456 \times 10^{-17}$; OFF-type: $p = 2.046 \times 10^{-13}$; ANOVA with Bonferroni correction). As such, invariance values among ON- and OFF-type ELL PCells were similar (**Figure 7E**, left; $p = 0.525$, *t*-test) but larger than those for single afferents (compare with **Figure 5E**; ON-type: $p = 5.456 \times 10^{-17}$; OFF-type: $p = 2.046 \times 10^{-13}$; ANOVA with Bonferroni correction). It should be noted that invariance values computed from single ELL PCells were lower owing to the fact that each cell type responded differentially when the stimulus occurred on different beat

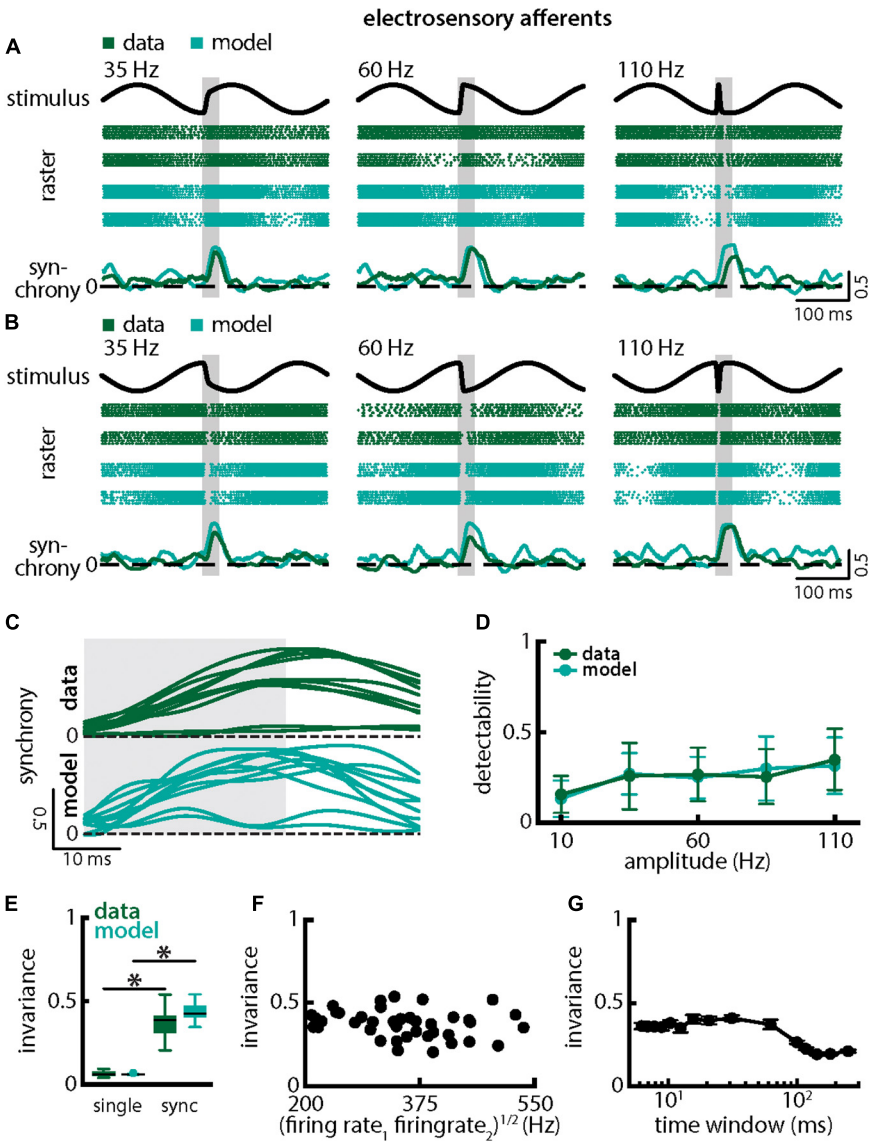

FIGURE 6 | Synchrony provides an invariant representation of chirps with different amplitudes. **(A)** Example stimulus waveforms (top, black) for chirps with different amplitudes (left: 35 Hz; middle: 60 Hz; right: 110 Hz) occurring at the same phase of the 4 Hz beat, raster plots of two example afferents (middle top, green) and model neurons (middle bottom, cyan) showing responses to 5 out of 20 randomly chosen presentations (i.e., trials) of each stimulus and the time varying synchrony averaged over all 20 trials (bottom) from the shown example afferent (green) and model (cyan) pairs. The horizontal bars (shaded gray) represent the chirp window used for evaluation. The gray band shows the evaluation time window used to compute invariance (see below). **(B)** Same as **(A)**, but when the chirp occurred at a different phase of the beat. **(C)** Trial-averaged synchrony responses from the same example pair of afferents for experimental data (green) and from a pair of model afferents (cyan) for chirps of different amplitudes. **(D)** Population-averaged neuronal detectability values computed from the spiking synchrony from the afferents (green) and the model (cyan) as a function of amplitude. **(E)** Population-averaged invariance values computed for the afferents (green) and for the model (cyan) from single afferent activity (left) and from synchrony (right) for chirp amplitude. "*" indicates statistical significance at the $p = 0.05$ level using a paired t-test. **(F)** Invariance as a function of the geometric mean of the afferent baseline firing rates for our dataset. No significant correlation was observed (Pearson's correlation coefficient, $r = -0.2148$; $p = 0.1832$). **(G)** Invariance as a function of time window length. Invariance was more or less independent of time window length for values up to ~60 ms.

phases (compare **Figures 7A–C**). Invariance scores computed for a given phase were significantly higher (**Figure 7E**, right; $p = 8.380 \times 10^{-8}$; paired t-tests), which further confirms that single ELL PCell responses are more invariant than those of single afferents. It is important to note that the higher invariance scores seen for ELL PCells to chirps with different durations as compared to afferents is thus primarily due to the fact that trial-averaged firing rate responses were more similar to one another rather than variability. This is because the invariance measure is computed using the trial-averaged time dependent firing rates (i.e., firing rates averaged over repeated presentations of the stimulus waveform associated with a chirp with given

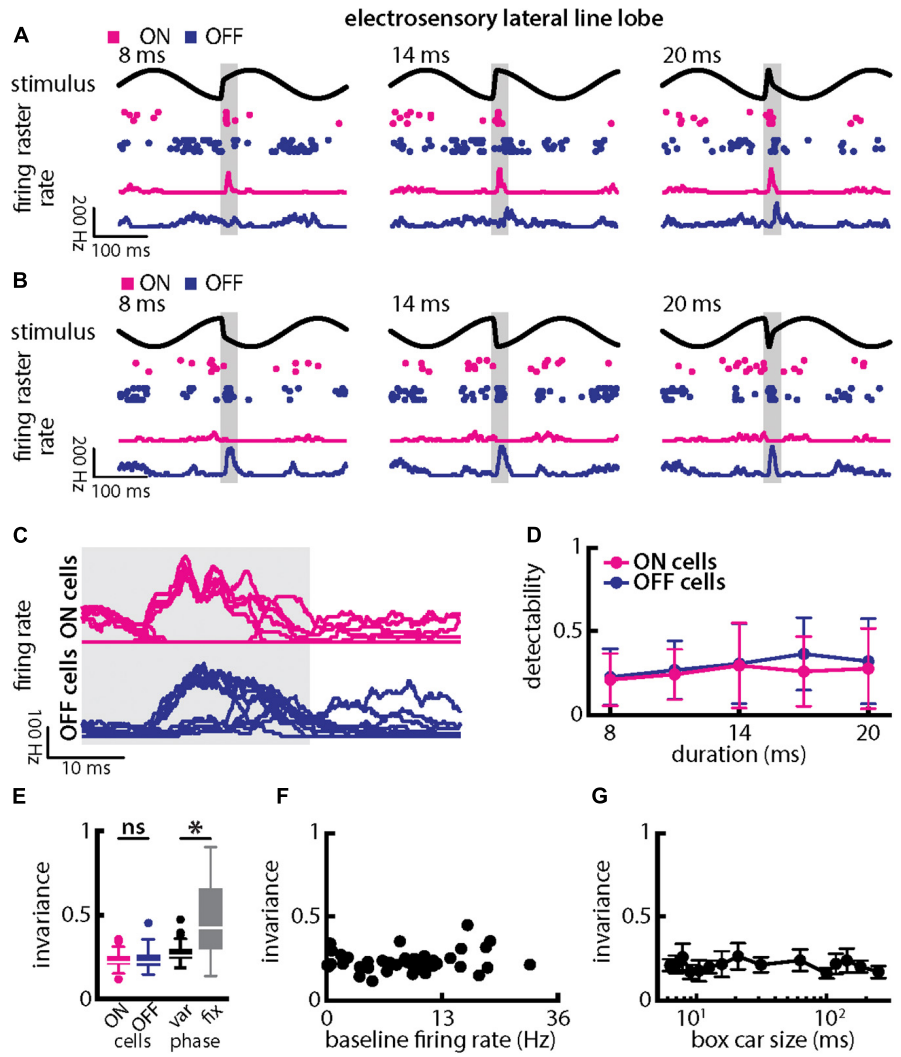

FIGURE 7 | Central electrosensory neurons display more invariant representation of chirps with varying duration than peripheral electroreceptor afferents. **(A)** Example stimulus waveforms (top, black) for chirps with different durations (left: 8 ms; middle: 14 ms; right: 20 ms) occurring at the same phase of the 4 Hz beat, raster plots of an example ON-PCell (middle top, magenta) and OFF-PCell (middle bottom, blue) showing responses to 5 out of 20 randomly chosen presentations (i.e., trials) of each stimulus and corresponding firing rates of the same PCells averaged over all 20 trials (bottom). The gray band shows the evaluation time window used to compute invariance (see below). **(B)** Same as **(A)**, but when the chirp occurred at a different phase of the beat. **(C)** Superimposed trial-averaged firing rate responses of the same example ON (top panel) and OFF-type (bottom panel) PCells to chirps with different durations. **(D)** Averaged detectability values computed from firing rate responses for our ON-type PCell population (magenta) and our OFF-type PCell population (blue) as a function of duration. **(E)** Left: Population-averaged invariance values computed from ON (magenta) and OFF-type (blue) PCells for duration. Right: Invariance computed for all chirp phases used (left) and when only the phase that elicited excitatory responses in our PCell population was used (right) for varying chirp duration values. "*" indicates statistical significance at the $p = 0.05$ level using a paired t-test. **(F)** Invariance as a function of baseline firing rate. No significant correlation was observed (Pearson's correlation coefficient, $r = 0.1505$; $p = 0.3539$). **(G)** Population-averaged invariance as a function of the boxcar size used to obtain the time dependent firing rate from spiking activity.

duration and amplitude) rather than single-trial responses which are more variable (see section Materials and Methods). Pyramidal cell heterogeneities as quantified by the baseline firing rate did not affect invariance as no significant correlation was observed (**Figure 7F**; Pearson's correlation coefficient, $r = 0.1505$; $p = 0.3539$). Invariance scores were furthermore robust to changes in the filter settings used to obtain the time-dependent firing rate from spiking activity (**Figure 7G**). Further, we note that previous studies have shown that some midbrain neurons receive balanced input from ON- and OFF-type ELL PCells (McGillivray et al., 2012; Aumentado-Armstrong et al., 2015) whose responses would then be expected to be more invariant as seen previously for other chirp attributes (Metzen et al., 2016).

Qualitatively similar results were obtained when varying chirp amplitude. Overall, responses of ON- and OFF-type were largely independent of chirp amplitude when the stimulus occurred at a given background beat phase (**Figures 8A–C**, magenta and blue). Stimulus detectability was higher than that of single

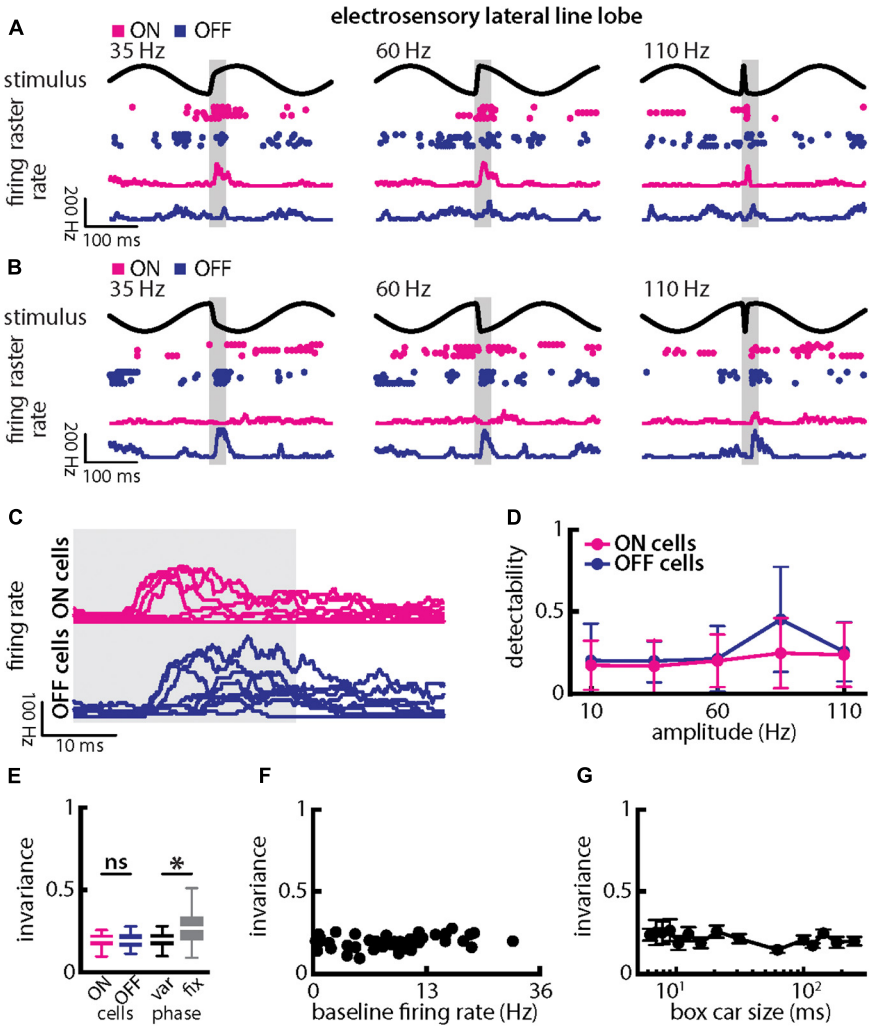

FIGURE 8 | Central electrosensory neurons display more invariant representation of chirps with varying amplitude than peripheral electroreceptor afferents. **(A)** Example stimulus waveforms (top) for chirps with different amplitude (left: 35 Hz; middle: 60 Hz; right: 110 Hz) occurring at the same phase of the 4 Hz beat, raster plots of an example ON-PCell (middle top, magenta) and OFF-PCell (middle bottom, blue) showing responses to 5 out of 20 randomly chosen presentations (i.e., trials) of each stimulus and corresponding firing rates the same PCells averaged over all 20 trials (bottom). The gray band shows the evaluation time window used to compute invariance (see below). **(B)** Same as **(A)**, but when the chirp occurred at a different phase of the beat. **(C)** Superimposed trial-averaged firing rate responses of the same example ON (top panel) and OFF-type (bottom panel) PCells to chirps with different amplitudes. **(D)** Population-averaged detectability values computed from firing rate responses for our ON-type PCell population (magenta) and our OFF-type PCell population (blue) as a function of amplitude. **(E)** Left: Population-averaged invariance values computed from ON (magenta) and OFF-type (blue) PCells for amplitude. Right: Invariance computed for all chirp phases used (left) and when only the phase that elicited excitatory responses in our PCell population was used (right) for varying chirp amplitude values. "*" indicates statistical significance at the $p = 0.05$ level using a paired t-test. **(F)** Invariance as a function of baseline firing rate. No significant correlation was observed (Pearson's correlation coefficient, $r = 0.2619$; $p = 0.1026$). **(G)** Population-averaged invariance as a function of the boxcar size used to obtain the time dependent firing rate from spiking activity.

afferents (**Figure 8D**, ON-type: $p = 2.686 \times 10^{-20}$; OFF-type: $p = 1.28 \times 10^{-20}$; ANOVA with Bonferroni correction). As such, invariance values, although similar for ON- and OFF-type cells (**Figure 8D**, left; $p = 0.954$, t-test), were significantly higher than those obtained for single afferents (compare with **Figure 6E**; ON-type: $p = 2.686 \times 10^{-20}$; EAs vs. OFF-type: $p = 1.28 \times 10^{-20}$; ANOVA with Bonferroni correction). Invariance scores were lower owing to the fact that each cell type responded differentially when the stimulus occurred on different beat phases. Invariance scores computed for a given phase were significantly higher than those computed across phases (**Figure 8E**, right; $p = 8.380 \times 10^{-8}$; paired t-tests), owing to the fact that ON- and OFF-type cells responded differentially when stimuli occurred at different phases of the background beat (compare **Figures 8A–C**). Pyramidal cell heterogeneities as quantified by the baseline firing rate did not affect invariance as no significant correlation was observed (**Figure 8F**, Pearson's correlation coefficient, $r = 0.2619$; $p = 0.1026$). Invariance scores were robust to changes in the filter settings used to obtain the time-dependent firing rate from spiking activity (**Figure 8G**).

Overall, our results strongly suggest that single ELL PCells decode synchronous afferent activity elicited by natural electrocommunication stimuli with different durations and amplitudes. This is because their response detectability and invariance are more consistent with those obtained from afferent synchrony than those obtained from single afferent activity.

Weakly Electric Fish Display Behavioral Responses That Are Invariant to Natural Electrocommunication Stimuli With Varying Duration and Amplitude

Finally, we investigated behavioral responses to chirps with different amplitudes and duration (**Figure 9**). To do so, we took advantage of the fact that *A. leptorhynchus* display "chirp echo responses" when stimulated with chirps (Hupé and Lewis, 2008) (**Figure 9A**; see Materials and Methods). Our results show that the behavioral responses elicited by chirp stimuli with different durations (**Figure 9B**) or amplitudes (**Figure 9C**) were similar to one another and that echo response rates were similar across different chirp durations (**Figure 9D**) as well as different chirp amplitudes (**Figure 9E**). Consequently, invariance values computed from behavioral responses were significantly higher than those obtained for either single afferents or PCells (**Figure 9F**; duration: EAs: $p = 1.739 \times 10^{-15}$; PCells: $p = 0.033$; **Figure 9G**; amplitude: EAs: $p = 1.028 \times 10^{-15}$; PCells: $p = 5.125 \times 10^{-5}$; ANOVA with Bonferroni correction). It is important to note that the behavioral responses (i.e., echo response rates) were most likely elicited by the chirp stimuli rather than the beat (see section Materials and Methods). It should furthermore be noted that differences in the timecourse of echo response rates that were most likely due to estimation error and/or fluctuations actually limited behavioral invariance values obtained here. These should thus be seen as lower bounds as is further discussed below. Our results thus show that behavioral responses were invariant to both chirp duration and amplitude, consistent with the hypothesis that changes in synchronous afferent activity, rather than changes in the single afferent firing rate, are decoded by ELL PCells. Our results have thus revealed that neural synchrony can be used to generate a neuronal representation that is invariant to stimuli with different attributes and how this representation is further processed downstream to presumably give rise to behavior. Moreover, the duration and amplitude of the echo response chirps elicited by the fish did not significantly change for different chirp parameters (**Figure 9H**; KS tests, $p \geq 0.1161$ in all cases).

DISCUSSION

Summary of Results

Here we investigated how electrosensory neural populations encoded natural electrocommunication stimuli with varying attributes (i.e., duration and amplitude) in order to mediate behavior. Despite the fact that both attributes were narrowly distributed under natural conditions, recordings from peripheral afferents revealed that, while single neurons encoded the different stimulus waveforms associated with different durations or amplitudes, all waveforms gave rise to increased synchrony either through excitation or inhibition at the population level. A phenomenological mathematical model reproduced experimental data showing that afferent responses at both the single neuron and population levels could be accounted for by single neuron filtering and spiking properties. Recordings from downstream central electrosensory neurons (i.e., ELL PCells) revealed that they decode information carried by synchronous activities of afferents as their responses were more invariant than those of single afferents. Specifically, ON-type cells were excited when afferents are excited synchronously while OFF-type cells were instead excited when afferents are inhibited synchronously. It is likely that ELL PCell responses are further processed by downstream brain areas to give rise to the observed invariant behavioral responses to natural electrocommunication stimuli. Our results thus reveal that neural synchrony can be used to generate an invariant representation to natural electrocommunication stimuli with different attributes as well as the mechanisms by which this representation is decoded by downstream neurons to presumably lead to behavioral responses at the organismal level.

Feature Invariant Representations of Natural Electrocommunication Stimuli: Functional Consequences for Coding and Perception

The results of the current study have shown that electrosensory pathway encodes natural electrocommunication stimuli with different attributes. These are unlike those considered previously in which a natural electrocommunication stimulus with given attributes (i.e., the same amplitude and duration) occurred at different phases of the underlying background signal (Metzen et al., 2016; Metzen and Chacron, 2017), as seen under natural conditions (Walz et al., 2013; Aumentado-Armstrong et al., 2015). While single afferents encoded the resulting different stimulus waveforms differentially, synchrony between afferents at the population level provided an invariant representation that is decoded by downstream neurons to give rise to behavior (Metzen et al., 2016). Such invariant responses are desirable from a functional point of view because the probability at which natural electrocommunication signals occur is independent of the phase of the background signal at the time of emission (Walz et al., 2013; Aumentado-Armstrong et al., 2015). As such, these responses enable the organism to correctly perceive that different waveforms are actually generated due to the same electrocommunication signal (i.e., with given duration and amplitude).

As such, our result showing that natural electrocommunication stimuli with different amplitudes and durations are encoded in an invariant fashion by the electrosensory pathway is surprising. This is because, unlike the background beat phase considered above, both chirp amplitude and duration are instead narrowly distributed under natural conditions. Indeed, previous studies have shown that the natural electrocommunication signals differ in terms of duration and amplitude across different

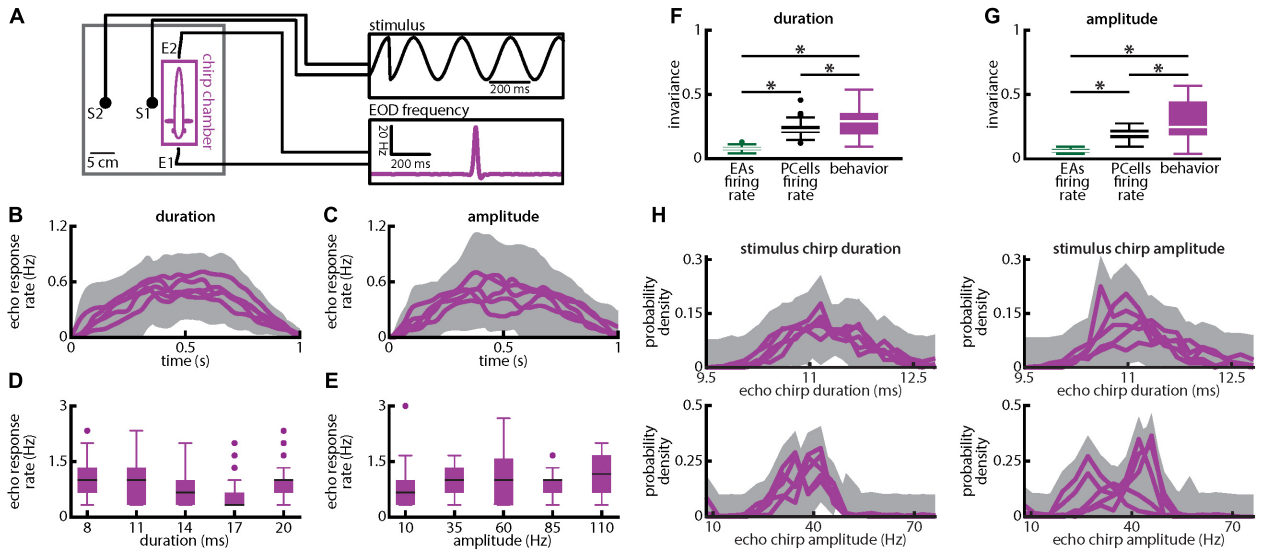

FIGURE 9 | Weakly electric fish display invariant behavioral responses to chirps with varying duration and amplitude. **(A)** Experimental setup. Each fish ($N = 8$) was placed in an enclosure within a tank (chirp chamber). Stimuli were applied via two electrodes (S1 and S2) perpendicular to the fish's rostro-caudal axis. The fish's EOD frequency was recorded by a pair of electrodes positioned at the head and tail of the animal (E1 and E2). Behavioral responses consisted of communication stimuli characterized by transient increases in EOD frequency in response to the presented stimulus. **(B,C)** Population-averaged time dependent echo response rates for chirps of different durations **(B)** and amplitudes **(C)**. The shaded gray bands represent the 95% confidence intervals. **(D,E)** Population-averaged behavioral echo response rate (purple) for different durations **(D)** and amplitudes **(E)**. **(F,G)** Population-averaged invariance scores computed from behavioral responses (purple) in comparison to the neuronal invariance scores using single afferents (green) and PCells (black) obtained for different durations **(F)** and amplitudes **(G)**. "*" indicates statistical significance at the $p = 0.05$ level using a one-way ANOVA with Bonferroni correction. **(H)** *Top:* Probability distributions of echo response duration for stimulus chirps of different durations (left) and amplitudes (right). *Bottom:* Probability distributions of echo response amplitude for stimulus chirps of different durations (left) and amplitudes (right). In all four cases, the probability distributions were not significantly different from one another (KS tests, $p \geq 0.1161$ in all cases).

Apteronotid species and could thus be used in theory to distinguish between con- and hetero-specifics (Petzold et al., 2016). This is even more surprising because we considered chirps with attributes that were well outside of the range observed for chirps emitted by fish (see **Figure 2**). However, our results show that such "un-natural" chirps gave rise to neural (in terms of EA synchrony) and behavioral responses that closely resembled those observed for more "natural" chirps. Our results thus provide evidence against (but do not disprove, see below) the hypothesis that differences in chirp duration and amplitude are encoded by the electrosensory system and can be perceived by the organism. Specifically, they suggest that despite large differences in their attributes, such stimuli are all ultimately perceived similarly. If correct, then this hypothesis greatly complicates the problem of distinguishing between conspecific and heterospecific individuals based on chirp characteristics. Our results support the proposal that the functional role of chirps is to temporarily suppress electrosensory neuronal responses to other stimuli (i.e., temporarily "blind" the opponent) (Hupé and Lewis, 2008). This is because peripheral afferent activities will then be synchronized irrespective of stimulus attributes. Further evidence for this hypothesis comes from previous electrophysiological studies showing that both ELL (Marsat et al., 2009; Vonderschen and Chacron, 2011) and TS (Vonderschen and Chacron, 2011) neurons are best at detecting the presence of natural electrocommunication stimuli rather than at discriminating between differences in stimulus attributes.

That said, it is important to note that our results do not imply that weakly electric fish cannot distinguish between chirps with different attributes. Specifically, our results do not rule out the possibility that the animals can actually perceive differences in chirp amplitude and duration but simply do not report them behaviorally. Indeed, it is possible that ELL pyramidal cells other than the ones considered here (i.e., in other segments) could actually decode information about stimulus attributes carried by single peripheral afferents. This possibility is however unlikely because previous studies have shown that the ELL pyramidal cells within the lateral segment considered here give the strongest responses to natural electrocommunication stimuli (Marsat et al., 2009). It is furthermore important to note that previous studies have shown that the invariant neuronal responses due to synchrony and the invariant behavioral responses with given attributes occurring at different phases of the underlying background both deteriorate when higher beat frequencies are considered (Metzen and Chacron, 2017). This is because EA synchrony during the chirp is much weaker for higher beat frequencies and thus more commensurate with that seen during the beat (Walz et al., 2014). Importantly, our previous results showing that both neural and behavioral invariance deteriorate for higher beat frequencies provides a strong link between changes in invariance due to EA synchrony and changes in behavioral invariance (Metzen and Chacron, 2017). Under natural conditions, the beat frequency can reach much higher values (e.g., 400 Hz) than the one considered

in the current study and we predict that, as seen for phase-invariance, both the duration and amplitude-invariant neural and behavioral responses seen here would deteriorate when higher beat frequencies are used. Future studies should thus investigate how increasing the beat frequency affects invariant coding and perception of electrocommunication stimuli with different durations and amplitudes.

We further hypothesize that the invariant neuronal and behavioral responses to natural electrocommunication stimuli considered here would break down when the stimulus contrast is increased beyond that explored in this study which is experienced when fish are located ∼13 cm from one another (Yu et al., 2019). Indeed, higher contrasts are experienced when two conspecifics move closer (i.e., within 5 cm) to one another (Yu et al., 2012, 2019), or when the beat frequency is increased. This is because we predict that peripheral afferents will then display stronger phase locking (i.e., only fire during specifics phases) to the background signal, which will increase their synchrony. In the case of increasing beat frequency, this is due to their known high-pass tuning characteristics (Xu et al., 1996; Metzen and Chacron, 2017). As such, we propose that weakly electric fish will be able to discriminate between natural electrocommunication stimuli with different attributes whenever these are produced when both animals are when in close proximity to another or during high frequency beats. Further studies are needed to test this hypothesis. If true, our hypothesis would provide an explanation as to recent field results showing that natural electrocommunication signals are sometimes produced when both animals are located close to one another or during high frequency beats (Henninger et al., 2018).

Our results have shown that invariant responses of EA synchrony to chirps with varying amplitude and duration are likely decoded by ELL PCells to presumable lead to behavior. However, it should be noted that our study only considered synchrony between EA pairs whereas the PCells within the lateral segment considered here receive input from ∼1000 EAs (Maler, 2009). Previous studies have shown that PCells display ion channels such as persistent sodium which would favor detection of coincident EA activity (Noonan et al., 2003). Further, modeling studies have suggested that the tuning properties of PCells within the lateral segment emerge because they actually detect coincident EA input (Middleton et al., 2009). However, integration of EA input by PCells within the lateral segment has not been systematically studied experimentally. For example, the so-called "synchrony receptive fields" (Brette, 2012) (e.g., the fraction of EA firing synchronously needed to elicit PCell firing, or the time window during which EA activity can be considered synchronous) remain unknown to date. While previous results (Marsat et al., 2009; Marsat and Maler, 2010; Metzen et al., 2016; Metzen and Chacron, 2017) and the results of the current study are consistent with the hypothesis that PCells within the lateral segment detect coincident EA activity, further studies are needed to fully test this hypothesis.

Further, we note that our behavioral invariance values were actually lower than those obtained for EA synchrony. As mentioned above, this is likely due to the fact that the former were limited by fluctuations and we predict that behavioral invariance values are actually higher. Further studies are however needed to understand how the activities of PCell population are integrated downstream. Previous studies have shown that some midbrain electrosensory neurons display invariant responses to beat phase (i.e., a neural correlated of the observed behavioral invariance to beat phase) by integrating input from ON- and OFF-type cells (Metzen et al., 2016). We hypothesize that this mechanism will give rise to responses in midbrain neurons that are fully invariant to chirps of different amplitudes or durations irrespective of the beat phase at which they occur. While there is anecdotal evidence that such neurons exist (see Figures 2C, 8A of Vonderschen and Chacron, 2011), the responses of midbrain neurons to stimulation protocols similar to the ones used in the current study have not been systematically investigated to date and should be the focus of future studies.

It is important to note here that both EAs and ELL PCells display significant heterogeneities in terms of baseline activity as well as responses to stimuli (Bastian, 1981; Bastian and Nguyenkim, 2001; Bastian et al., 2002, 2004; Gussin et al., 2007; Savard et al., 2011). While it is clear that heterogeneous populations are advantageous for coding (Stocks, 2000; Padmanabhan and Urban, 2010; Brette, 2012; Mejias and Longtin, 2012), our modeling and experimental data suggest that heterogeneities are not necessary to observe the phenomena described in the current study. Specifically, our modeling, which was based on a homogeneous neural population, reproduced our experimental data both at the single neuron and population levels for EAs. Moreover, we found no significant correlation between invariance and the baseline firing rate, which is strongly correlated with morphological differences in ELL PCells (Bastian and Nguyenkim, 2001; Bastian et al., 2004). Further studies are needed in order to investigate the effects of neural heterogeneities on invariance coding at both the EA and ELL PCell level. For the former, these should investigate how EA heterogeneities influence the so-called "synchrony receptive fields" of ELL PCells mentioned above. For the latter, the effects of PCell heterogeneities should also be more systematically investigated. This is particularly important as previous studies have shown that a strong factor contributing to PCell heterogeneities is the amount of descending input (i.e., feedback) that is received from higher brain centers. The effect of such feedback has been mostly studied at the single neuron level (Bastian, 1986; Chacron et al., 2005b; Marsat and Maler, 2012; Huang et al., 2018, 2019; Metzen et al., 2018) and further studies are needed to understand whether and, if so, how such feedback can facilitate detection of EA synchrony by ELL PCells.

Finally, we note that the electrocommunication stimuli considered in the present study primarily occur during agonistic encounters and, as such, correspond to the "type II chirps" described previously. It is important to note that *A. leptorhynchus* emit other types of natural communication stimuli that are not considered here (Zakon et al., 2002). In particular, they tend to emit another type of communication signal termed

"type I chirps" during mating behavior. Electrophysiological studies have shown that neuronal responses to these are fundamentally different (Marsat and Maler, 2010; Vonderschen and Chacron, 2011; Allen and Marsat, 2018). Future studies are needed to investigate how electrosensory neuronal populations encode other natural electrocommunication signals not considered here. In particular, it will be important to consider the fact that ELL PCell trial-to-trial variabilities to repeated stimulus presentations are correlated (Chacron and Bastian, 2008; Simmonds and Chacron, 2015; Hofmann and Chacron, 2018), which has been ignored by previous studies (Marsat et al., 2009; Marsat and Maler, 2010; Allen and Marsat, 2018). Such "noise" correlations can have profound influence on coding by neuronal populations (Averbeck et al., 2006; Cohen and Kohn, 2011; Doiron et al., 2016; Franke et al., 2016; Zylberberg et al., 2016) and are likely to be found in all ELL PCells as they are due to shared input from peripheral afferents (Hofmann and Chacron, 2017, 2018).

Implications for Other Systems

Here we have provided the first experimental evidence that synchrony can enable the emergence of a neuronal representation that is invariant to stimuli with different attributes such as amplitude and duration. Such invariant representations are also seen in other systems (auditory: Bendor and Wang, 2005; visual: Zoccolan et al., 2007; olfactory: Martelli et al., 2013). In all cases, tolerance to variations in identity-preserving transformations such as size, contrast, or viewpoint progressively increases in neurons at higher processing stages (Dicarlo and Cox, 2007). The mechanisms leading to such an increase in invariance are not fully understood to date. Our results showing how neural synchrony, which is observed ubiquitously in the central nervous system (Uhlhaas et al., 2009; Harris and Gordon, 2015), gives rise to a neuronal representation that is invariant to both amplitude and duration is thus likely to be shared by other systems/species. This is because invariance to stimulus amplitude has been observed in the visual (Anderson et al., 2000), auditory (Billimoria et al., 2008; Barbour, 2011), somatosensory (Pei et al., 2010), and olfactory (Storace and Cohen, 2017) systems. The fact that the electrosensory system studied here displays both anatomical and functional similarities with other systems (Clarke et al., 2015) suggests that neural synchrony also plays a role in mediating the emergence and refinement of such representations in other systems. Further studies are however needed to test this prediction.

DATA AVAILABILITY STATEMENT

The datasets (analyzed) and codes for this study can be found in the figshare repository (doi: 10.6084/m9.figshare.8041136).

ETHICS STATEMENT

The animal study was reviewed and approved by McGill University's animal care committee under protocol number 5285.

AUTHOR CONTRIBUTIONS

MM, VH, and MC: conceptualization, funding acquisition, methodology, validation, writing, review, and editing. MM: data curation, formal analysis, visualization, and writing – original draft. MM and VH: investigation. MC: project administration, resources, software, and supervision.

FUNDING

This research was supported by the Canadian Institutes of Health Research (MM and MC) and the Deutsche Forschungsgemeinschaft DFG (VH; HO5912/1-1). The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

REFERENCES

Allen, K. M., and Marsat, G. (2018). Task-specific sensory coding strategies are matched to detection and discrimination performance. *J. Exp. Biol.* 221(Pt 6):jeb170563. doi: 10.1242/jeb.170563

Allen, K. M., and Marsat, G. (2019). Neural processing of communication signals: the extent of sender-receiver matching varies across species of *Apteronotus*. *eNeuro* 6:ENEURO.0392-18.2019. doi: 10.1523/ENEURO.0392-18.2019

Anderson, J. S., Lampl, I., Gillespie, D. C., and Ferster, D. (2000). The contribution of noise to contrast invariance of orientation tuning in cat visual cortex. *Science* 290, 1968–1972. doi: 10.1126/science.290.5498.1968

Attias, H., and Schreiner, C. E. (1997). Temporal low-order statistics of natural sounds. *Adv. Neural Inf. Process. Syst.* 9, 27–33.

Aumentado-Armstrong, T., Metzen, M. G., Sproule, M. K., and Chacron, M. J. (2015). Electrosensory midbrain neurons display feature invariant responses to natural communication stimuli. *PLoS Comput. Biol.* 11:e1004430. doi: 10.1371/journal.pcbi.1004430

Averbeck, B. B., Latham, P. E., and Pouget, A. (2006). Neural correlations, population coding and computation. *Nat. Rev. Neurosci.* 7, 358–366.

Barbour, D. L. (2011). Intensity-invariant coding in the auditory system. *Neurosci. Biobehav. Rev.* 35, 2064–2072. doi: 10.1016/j.neubiorev.2011.04.009

Bastian, J. (1981). Electrolocation. 1. How the electroreceptors of *Apteronotus albifrons* code for moving-objects and other electrical stimuli. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 144, 465–479. doi: 10.1007/bf01326832

Bastian, J. (1986). Gain control in the electrosensory system mediated by descending inputs to the electrosensory lateral line lobe. *J. Neurosci.* 6, 553–562. doi: 10.1523/jneurosci.06-02-00553.1986

Bastian, J., Chacron, M. J., and Maler, L. (2002). Receptive field organization determines pyramidal cell stimulus-encoding capability and spatial stimulus selectivity. *J. Neurosci.* 22, 4577–4590. doi: 10.1523/jneurosci.22-11-04577.2002

Bastian, J., Chacron, M. J., and Maler, L. (2004). Plastic and non-plastic cells perform unique roles in a network capable of adaptive redundancy reduction. *Neuron* 41, 767–779. doi: 10.1016/s0896-6273(04)00071-6

Bastian, J., and Nguyenkim, J. (2001). Dendritic modulation of burst-like firing in sensory neurons. *J. Neurophysiol.* 85, 10–22. doi: 10.1152/jn.2001.85.1.10

Bastian, J., Schniederjan, S., and Nguyenkim, J. (2001). Arginine vasotocin modulates a sexually dimorphic communication behavior in the weakly electric fish *Apteronotus leptorhynchus*. *J. Exp. Biol.* 204, 1909–1923.

Benda, J., Longtin, A., and Maler, L. (2005). Spike-frequency adaptation separates transient communication signals from background oscillations. *J. Neurosci.* 25, 2312–2321. doi: 10.1523/jneurosci.4795-04.2005

Benda, J., Longtin, A., and Maler, L. (2006). A synchronization-desynchronization code for natural communication signals. *Neuron* 52, 347–358. doi: 10.1016/j.neuron.2006.08.008

Bendor, D., and Wang, X. (2005). The neuronal representation of pitch in primate auditory cortex. *Nature* 436, 1161–1165. doi: 10.1038/nature03867

Billimoria, C. P., Kraus, B. J., Narayan, R., Maddox, R. K., and Sen, K. (2008). Invariance and sensitivity to intensity in neural discrimination of natural sounds. *J. Neurosci.* 28, 6304–6308. doi: 10.1523/JNEUROSCI.0961-08.2008

Brette, R. (2012). Computing with neural synchrony. *PLoS Comput. Biol.* 8:e1002561. doi: 10.1371/journal.pcbi.1002561

Chacron, M. J., and Bastian, J. (2008). Population coding by electrosensory neurons. *J. Neurophysiol.* 99, 1825–1835. doi: 10.1152/jn.01266.2007

Chacron, M. J., and Fortune, E. S. (2010). Subthreshold membrane conductances enhance directional selectivity in vertebrate sensory neurons. *J. Neurophysiol.* 104, 449–462. doi: 10.1152/jn.01113.2009

Chacron, M. J., Longtin, A., and Maler, L. (2001). Simple models of bursting and non-bursting P-type electroreceptors. *Neurocomputing* 38, 129–139. doi: 10.1016/s0925-2312(01)00549-5

Chacron, M. J., Longtin, A., St-Hilaire, M., and Maler, L. (2000). Suprathreshold stochastic firing dynamics with memory in P-type electroreceptors. *Phys. Rev. Lett.* 85, 1576–1579. doi: 10.1103/physrevlett.85.1576

Chacron, M. J., Maler, L., and Bastian, J. (2005a). Electroreceptor neuron dynamics shape information transmission. *Nat. Neurosci.* 8, 673–678. doi: 10.1038/nn1433

Chacron, M. J., Maler, L., and Bastian, J. (2005b). Feedback and feedforward control of frequency tuning to naturalistic stimuli. *J. Neurosci.* 25, 5521–5532. doi: 10.1523/jneurosci.0445-05.2005

Chacron, M. J., Toporikova, N., and Fortune, E. S. (2009). Differences in the time course of short-term depression across receptive fields are correlated with directional selectivity in electrosensory neurons. *J. Neurophysiol.* 102, 3270–3279. doi: 10.1152/jn.00645.2009

Clarke, S. E., Longtin, A., and Maler, L. (2015). Contrast coding in the electrosensory system: parallels with visual computation. *Nat. Rev. Neurosci.* 16, 733–744. doi: 10.1038/nrn4037

Cohen, M. R., and Kohn, A. (2011). Measuring and interpreting neuronal correlations. *Nat. Neurosci.* 14, 811–819. doi: 10.1038/nn.2842

Dan, Y., Alonso, J.-M., Usrey, W. M., and Reid, R. C. (1998). Coding of visual information by precisely correlated spikes in the lateral geniculate nucleus. *Nat. Neurosci.* 1, 501–507. doi: 10.1038/2217

Deemyad, T., Metzen, M. G., Pan, Y., and Chacron, M. J. (2013). Serotonin selectively enhances perception and sensory neural responses to stimuli generated by same-sex conspecifics. *Proc. Natl. Acad. Sci. U.S.A.* 110, 19609–19614. doi: 10.1073/pnas.1314008110

Dicarlo, J. J., and Cox, D. D. (2007). Untangling invariant object recognition. *Trends Cogn. Sci.* 11, 333–341. doi: 10.1016/j.tics.2007.06.010

Dicarlo, J. J., and Johnson, K. O. (1999). Velocity invariance of receptive field structure in somatosensory cortical area 3b of the alert monkey. *J. Neurosci.* 19, 401–419. doi: 10.1523/jneurosci.19-01-00401.1999

Doiron, B., Litwin-Kumar, A., Rosenbaum, R., Ocker, G. K., and Josic, K. (2016). The mechanics of state-dependent neural correlations. *Nat. Neurosci.* 19, 383–393. doi: 10.1038/nn.4242

Engler, G., Fogarty, C. M., Banks, J. R., and Zupanc, G. K. (2000). Spontaneous modulations of the electric organ discharge in the weakly electric fish, *Apteronotus leptorhynchus*: a biophysical and behavioral analysis. *J. Comp. Physiol. A* 186, 645–660. doi: 10.1007/s003590000118

Engler, G., and Zupanc, G. K. (2001). Differential production of chirping behavior evoked by electrical stimulation of the weakly electric fish, *Apteronotus leptorhynchus*. *J. Comp. Physiol. A* 187, 747–756. doi: 10.1007/s00359-001-0248-8

Frank, K., and Becker, M. C. (1964). "Microelectrodes for recording and stimulation," in *Physical Techniques in Biological Research*, ed. W. L. Nastuk (New York, NY: Academic Press), 23–84.

Franke, F., Fiscella, M., Sevelev, M., Roska, B., Hierlemann, A., and Da Silveira, R. A. (2016). Structures of neural correlation and how they favor coding. *Neuron* 89, 409–422. doi: 10.1016/j.neuron.2015.12.037

Gama Salgado, J. A., and Zupanc, G. K. H. (2011). Echo response to chirping in the weakly electric brown ghost knifefish (*Apteronotus leptorhynchus*): role of frequency and amplitude modulations. *Can. J. Zool.* 89, 498–508. doi: 10.1139/z11-014

Gray, C., and Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 86, 1698–1702. doi: 10.1073/pnas.86.5.1698

Gussin, D., Benda, J., and Maler, L. (2007). Limits of linear rate coding of dynamic stimuli by electroreceptor afferents. *J. Neurophysiol.* 97, 2917–2929. doi: 10.1152/jn.01243.2006

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

Henninger, J., Krahe, R., Kirschbaum, F., Grewe, J., and Benda, J. (2018). Statistics of natural communication signals observed in the wild identify important yet neglected stimulus regimes in weakly electric fish. *J. Neurosci.* 38, 5456–5465. doi: 10.1523/JNEUROSCI.0350-18.2018

Hitschfeld, É. M., Stamper, S. A., Vonderschen, K., Fortune, E. S., and Chacron, M. J. (2009). Effects of restraint and immobilization on electrosensory behaviors of weakly electric fish. *ILAR J.* 50, 361–372. doi: 10.1093/ilar.50.4.361

Hofmann, V., and Chacron, M. J. (2017). Differential receptive field organizations give rise to nearly identical neural correlations across three parallel sensory maps in weakly electric fish. *PLoS Comput. Biol.* 13:e1005716. doi: 10.1371/journal.pcbi.1005716

Hofmann, V., and Chacron, M. J. (2018). Population coding and correlated variability in electrosensory pathways. *Front. Integr. Neurosci.* 12:56. doi: 10.3389/fnint.2018.00056

Huang, C. G., Metzen, M. G., and Chacron, M. J. (2018). Feedback optimizes neural coding and perception of natural stimuli. *eLife* 7:e38935. doi: 10.7554/eLife.38935

Huang, C. G., Metzen, M. G., and Chacron, M. J. (2019). Descending pathways mediate adaptive optimized coding of natural stimuli in weakly electric fish. *Sci. Adv.* 5:eaax2211. doi: 10.1126/sciadv.aax2211

Hupé, G., and Lewis, J. (2008). Electrocommunication signals in free swimming brown ghost knifefish, *Apteronotus leptorhynchus*. *J. Exp. Biol.* 211, 1657–1667. doi: 10.1242/jeb.013516

Kolodziejski, J. A., Nelson, B. S., and Smith, G. T. (2005). Sex and species differences in neuromodulatory input to a premotor nucleus: a comparative study of substance P and communication behavior in weakly electric fish. *J. Neurobiol.* 62, 299–315. doi: 10.1002/neu.20095

Laurent, G. (2002). Olfactory network dynamics and the coding of multidimensional signals. *Nat. Rev. Neurosci.* 3, 884–895. doi: 10.1038/nrn964

Maler, L. (2009). Receptive field organization across multiple electrosensory maps. I. Columnar organization and estimation of receptive field size. *J. Comp. Neurol.* 516, 376–393. doi: 10.1002/cne.22124

Mante, V., Frazor, R. A., Bonin, V., Geisler, W. S., and Carandini, M. (2005). Independence of luminance and contrast in natural scenes and in the early visual system. *Nat. Neurosci.* 8, 1690–1697. doi: 10.1038/nn1556

Marsat, G., and Maler, L. (2010). Neural heterogeneity and efficient population codes for communication signals. *J. Neurophysiol.* 104, 2543–2555. doi: 10.1152/jn.00256.2010

Marsat, G., and Maler, L. (2012). Preparing for the unpredictable: adaptive feedback enhances the response to unexpected communication signals. *J. Neurophysiol.* 107, 1241–1246. doi: 10.1152/jn.00982.2011

Marsat, G., Proville, R. D., and Maler, L. (2009). Transient signals trigger synchronous bursts in an identified population of neurons. *J. Neurophysiol.* 102, 714–723. doi: 10.1152/jn.91366.2008

Martelli, C., Carlson, J. R., and Emonet, T. (2013). Intensity invariant dynamics and odor-specific latencies in olfactory receptor neuron response. *J. Neurosci.* 33, 6285–6297. doi: 10.1523/JNEUROSCI.0426-12.2013

McGillivray, P., Vonderschen, K., Fortune, E. S., and Chacron, M. J. (2012). Parallel coding of first- and second-order stimulus attributes by midbrain electrosensory neurons. *J. Neurosci.* 32, 5510–5524. doi: 10.1523/JNEUROSCI.0478-12.2012

Mejias, J. F., and Longtin, A. (2012). Optimal heterogeneity for coding in spiking neural networks. *Phys. Rev. Lett.* 108:228102.

Metzen, M. G. (2019). Encoding and perception of electro-communication signals in *Apteronotus leptorhynchus*. *Front. Integr. Neurosci.* 13:39. doi: 10.3389/fnint.2019.00039

Metzen, M. G., and Chacron, M. J. (2014). Weakly electric fish display behavioral responses to envelopes naturally occurring during movement: implications for neural processing. *J. Exp. Biol.* 217, 1381–1391. doi: 10.1242/jeb.098574

Metzen, M. G., and Chacron, M. J. (2015). Neural heterogeneities determine response characteristics to second-, but not first-order stimulus features. *J. Neurosci.* 35, 3124–3138. doi: 10.1523/JNEUROSCI.3946-14.2015

Metzen, M. G., and Chacron, M. J. (2017). Stimulus background influences phase invariant coding by correlated neural activity. *eLife* 6:e24482. doi: 10.7554/eLife.24482

Metzen, M. G., Hofmann, V., and Chacron, M. J. (2016). Neural correlations enable invariant coding and perception of natural stimuli in weakly electric fish. *eLife* 5:e12993. doi: 10.7554/eLife.12993

Metzen, M. G., Huang, C. G., and Chacron, M. J. (2018). Descending pathways generate perception of and neural responses to weak sensory input. *PLoS Biol.* 16:e2005239. doi: 10.1371/journal.pbio.2005239

Metzen, M. G., Jamali, M., Carriot, J., Avila-Akerberg, O., Cullen, K. E., and Chacron, M. J. (2015). Coding of envelopes by correlated but not single-neuron activity requires neural variability. *Proc. Natl. Acad. Sci. U.S.A.* 112, 4791–4796. doi: 10.1073/pnas.1418224112

Middleton, J. W., Longtin, A., Benda, J., and Maler, L. (2009). Postsynaptic receptive field size and spike threshold determine encoding of high-frequency information via sensitivity to synchronous presynaptic activity. *J. Neurophysiol.* 101, 1160–1170. doi: 10.1152/jn.90814.2008

Nelson, M. E., and Maciver, M. A. (1999). Prey capture in the weakly electric fish *Apteronotus albifrons*: sensory acquisition strategies and electrosensory consequences. *J. Exp. Biol.* 202, 1195–1203.

Nelson, M. E., Xu, Z., and Payne, J. R. (1997). Characterization and modeling of P-type electrosensory afferent responses to amplitude modulations in a wave-type electric fish. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 181, 532–544. doi: 10.1007/s003590050137

Noonan, L., Doiron, B., Laing, C., Longtin, A., and Turner, R. W. (2003). A dynamic dendritic refractory period regulates burst discharge in the electrosensory lobe of weakly electric fish. *J. Neurosci.* 23, 1524–1534. doi: 10.1523/jneurosci.23-04-01524.2003

Nunez, A., and Malmierca, E. (2007). Corticofugal modulation of sensory information. *Adv. Anat. Embryol. Cell Biol.* 187, 1 p following table of contents, 1–74.

Olshausen, B. A., and Field, D. J. (2004). Sparse coding of sensory inputs. *Curr. Opin. Neurobiol.* 14, 481–487. doi: 10.1016/j.conb.2004.07.007

Padmanabhan, K., and Urban, N. N. (2010). Intrinsic biophysical diversity decorrelates neuronal firing while increasing information content. *Nat. Neurosci.* 13, 1276–1282. doi: 10.1038/nn.2630

Pei, Y. C., Hsiao, S. S., Craig, J. C., and Bensmaia, S. J. (2010). Shape invariant coding of motion direction in somatosensory cortex. *PLoS Biol.* 8:e1000305. doi: 10.1371/journal.pbio.1000305

Petzold, J. M., Marsat, G., and Troy Smith, G. (2016). Co-adaptation of electric organ discharges and chirps in south American ghost Knifefishes (Apteronotidae). *J. Physiol. Paris* 110, 200–215. doi: 10.1016/j.jphysparis.2016.10.005

Quiroga, R. Q., Reddy, L., Kreiman, G., Koch, C., and Fried, I. (2005). Invariant visual representation by single neurons in the human brain. *Nature* 435, 1102–1107. doi: 10.1038/nature03687

Rose, G. J. (2004). Insights into neural mechanisms and evolution of behaviour from electric fish. *Nat. Rev. Neurosci.* 5, 943–951. doi: 10.1038/nrn1558

Rust, N. C., and Dicarlo, J. J. (2010). Selectivity and tolerance ("invariance") both increase as visual information propagates from cortical area V4 to IT. *J. Neurosci.* 30, 12978–12995. doi: 10.1523/JNEUROSCI.0179-10.2010

Rust, N. C., and Dicarlo, J. J. (2012). Balanced increases in selectivity and tolerance produce constant sparseness along the ventral visual stream. *J. Neurosci.* 32, 10170–10182. doi: 10.1523/JNEUROSCI.6125-11.2012

Saunders, J., and Bastian, J. (1984). The physiology and morphology of two types of electrosensory neurons in the weakly electric fish, *Apteronotus leptorhynchus*. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol.* 154, 199–209. doi: 10.1007/bf00604985

Savard, M., Krahe, R., and Chacron, M. J. (2011). Neural heterogeneities influence envelope and temporal coding at the sensory periphery. *Neuroscience* 172, 270–284. doi: 10.1016/j.neuroscience.2010.10.061

Schneider, D. M., and Woolley, S. M. (2013). Sparse and background-invariant coding of vocalizations in auditory scenes. *Neuron* 79, 141–152. doi: 10.1016/j.neuron.2013.04.038

Sharpee, T. O., Kouh, M., and Reynolds, J. H. (2013). Trade-off between curvature tuning and position invariance in visual area V4. *Proc. Natl. Acad. Sci. U.S.A.* 110, 11618–11623. doi: 10.1073/pnas.1217479110

Shea-Brown, E., Josic, K., De La Rocha, J., and Doiron, B. (2008). Correlation and synchrony transfer in integrate-and-fire neurons: basic properties and consequences for coding. *Phys. Rev. Lett.* 100:108102.

Simmonds, B., and Chacron, M. J. (2015). Activation of parallel fiber feedback by spatially diffuse stimuli reduces signal and noise correlations via independent mechanisms in a cerebellum-like structure. *PLoS Comput. Biol.* 11:e1004034. doi: 10.1371/journal.pcbi.1004034

Sproule, M. K. J., Metzen, M. G., and Chacron, M. J. (2015). Parallel sparse and dense information coding streams in the electrosensory midbrain. *Neurosci. Lett.* 607, 1–6. doi: 10.1016/j.neulet.2015.09.014

Stocks, N. G. (2000). Suprathreshold stochastic resonance in multilevel threshold systems. *Phys. Rev. Lett.* 84, 2310–2313. doi: 10.1103/physrevlett.84.2310

Storace, D. A., and Cohen, L. B. (2017). Measuring the olfactory bulb input-output transformation reveals a contribution to the perception of odorant concentration invariance. *Nat. Commun.* 8:81. doi: 10.1038/s41467-017-00036-2

Theunissen, F. E., and Elie, J. E. (2014). Neural processing of natural sounds. *Nat. Rev. Neurosci.* 15, 355–366. doi: 10.1038/nrn3731

Toporikova, N., and Chacron, M. J. (2009). Dendritic SK channels gate information processing *in vivo* by regulating an intrinsic bursting mechanism seen *in vitro*. *J. Neurophysiol.* 102, 2273–2287. doi: 10.1152/jn.00282.2009

Turner, R. W., Maler, L., and Burrows, M. (1999). Electroreception and electrocommunication. *J. Exp. Biol.* 202, 1455–1458.

Uhlhaas, P. J., Pipa, G., Lima, B., Melloni, L., Neuenschwander, S., Nikolic, D., et al. (2009). Neural synchrony in cortical networks: history, concept and current status. *Front. Integr. Neurosci.* 3:17. doi: 10.3389/neuro.07.017.2009

Vinje, W. E., and Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science* 287, 1273–1276. doi: 10.1126/science.287.5456.1273

Vonderschen, K., and Chacron, M. J. (2011). Sparse and dense coding of natural stimuli by distinct midbrain neuron subpopulations in weakly electric fish. *J. Neurophysiol.* 106, 3102–3118. doi: 10.1152/jn.00588.2011

Walz, H., Grewe, J., and Benda, J. (2014). Static frequency tuning accounts for changes in neural synchrony evoked by transient communication signals. *J. Neurophysiol.* 112, 752–765. doi: 10.1152/jn.00576.2013

Walz, H., Hupe, G. J., Benda, J., and Lewis, J. E. (2013). The neuroethology of electrocommunication: how signal background influences sensory encoding and behaviour in *Apteronotus leptorhynchus*. *J. Physiol. Paris* 107, 13–25. doi: 10.1016/j.jphysparis.2012.07.001

Xu, Z., Payne, J. R., and Nelson, M. E. (1996). Logarithmic time course of sensory adaptation in electrosensory afferent nerve fibers in a weakly electric fish. *J. Neurophysiol.* 76, 2020–2032. doi: 10.1152/jn.1996.76.3.2020

Yu, N., Hupe, G., Garfinkle, C., Lewis, J. E., and Longtin, A. (2012). Coding conspecific identity and motion in the electric sense. *PLoS Comput. Biol.* 8:e1002564. doi: 10.1371/journal.pcbi.1002564

Yu, N., Hupe, G., Longtin, A., and Lewis, J. E. (2019). Electrosensory contrast signals for interacting weakly electric fish. *Front. Integr. Neurosci.* 13:36. doi: 10.3389/fnint.2019.00036

Zakon, H., Oestreich, J., Tallarovic, S., and Triefenbach, F. (2002). EOD modulations of brown ghost electric fish: JARs, chirps, rises, and dips. *J. Physiol. Paris* 96, 451–458. doi: 10.1016/s0928-4257(03)00012-3

Zoccolan, D., Kouh, M., Poggio, T., and Dicarlo, J. J. (2007). Trade-off between object selectivity and tolerance in monkey inferotemporal cortex. *J. Neurosci.* 27, 12292–12307. doi: 10.1523/jneurosci.1897-07.2007

Zupanc, G. K., Sirbulescu, R. F., Nichols, A., and Ilies, I. (2006). Electric interactions through chirping behavior in the weakly electric fish, *Apteronotus leptorhynchus*. *J. Comp. Physiol. A Sens. Neural Behav. Physiol.* 192, 159–173. doi: 10.1007/s00359-005-0058-5

Zupanc, G. K. H. (2002). From oscillators to modulators: behavioral and neural control of modulations of the electric organ discharge in the gymnotiform fish, *Apteronotus leptorhynchus*. *J. Physiol. Paris* 96, 459–472. doi: 10.1016/s0928-4257(03)00002-0

Zupanc, G. K. H., and Maler, L. (1993). Evoked chirping in the weakly electric fish *Apteronotus leptorhynchus*: a quantitative biophysical analysis. *Can. J. Zool.* 71, 2301–2310. doi: 10.1139/z93-323

Zylberberg, J., Cafaro, J., Turner, M. H., Shea-Brown, E., and Rieke, F. (2016). Direction-selective circuits shape noise to ensure a precise population code. *Neuron* 89, 369–383. doi: 10.1016/j.neuron.2015.11.019

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.

Copyright © 2020 Metzen, Hofmann and Chacron. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 18 February 2020
doi: 10.3389/fnins.2020.00124

Acoustic Change Responses to Amplitude Modulation in Cochlear Implant Users: Relationships to Speech Perception

Ji-Hye Han[1,2] *and Andrew Dimitrijevic*[1,3]*

[1] Communication Sciences Research Center, Cincinnati Childs Hospital Medical Center, Cincinnati, OH, United States, [2] Laboratory of Brain & Cognitive Sciences for Convergence Medicine, College of Medicine, Hallym University, Chuncheon, South Korea, [3] Department Otolaryngology-Head and Neck Surgery, Sunnybrook Health Sciences Centre, Faculty of Medicine, University of Toronto, Toronto, ON, Canada

Objectives: The ability to understand speech is highly variable in people with cochlear implants (CIs) and to date, there are no objective measures that identify the root of this discrepancy. However, behavioral measures of temporal processing such as the temporal modulation transfer function (TMTF) has previously found to be related to vowel and consonant identification in CI users. The acoustic change complex (ACC) is a cortical auditory-evoked potential response that can be elicited by a "change" in an ongoing stimulus. In this study, the ACC elicited by amplitude modulation (AM) change was related to measures of speech perception as well as the amplitude detection threshold in CI users.

Methods: Ten CI users (mean age: 50 years old) participated in this study. All subjects participated in behavioral tests that included both speech and amplitude modulation detection to obtain a TMTF. CI users were categorized as "good" (n = 6) or "poor" (n = 4) based on their speech-in noise score (<50%). 64-channel electroencephalographic recordings were conducted while CI users passively listened to AM change sounds that were presented in a free field setting. The AM change stimulus was white noise with four different AM rates (4, 40, 100, and 300 Hz).

Results: Behavioral results show that AM detection thresholds in CI users were higher compared to the normal-hearing (NH) group for all AM rates. The electrophysiological data suggest that N1 responses were significantly decreased in amplitude and their latencies were increased in CI users compared to NH controls. In addition, the N1 latencies for the poor CI performers were delayed compared to the good CI performers. The N1 latency for 40 Hz AM was correlated with various speech perception measures.

Conclusion: Our data suggest that the ACC to AM change provides an objective index of speech perception abilities that can be used to explain some of the variation in speech perception observed among CI users.

Keywords: acoustic change complex, amplitude modulation, temporal modulation transfer function, cochlear implants, N1

OPEN ACCESS

Edited by:
Andreas Bahmer, Universitätsklinikum Würzburg, Germany

Reviewed by:
Kaibao Nie, University of Washington, United States
Barbara Cone, University of Arizona, United States

***Correspondence:**
Andrew Dimitrijevic
andrew.dimitrijevic@sunnybrook.ca; cibrainlab.com

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *02 August 2019*
Accepted: *31 January 2020*
Published: *18 February 2020*

Citation:
Han J-H and Dimitrijevic A (2020) Acoustic Change Responses to Amplitude Modulation in Cochlear Implant Users: Relationships to Speech Perception. Front. Neurosci. 14:124. doi: 10.3389/fnins.2020.00124

INTRODUCTION

Cochlear implants (CIs) provide electrical stimulation to the auditory nerve that can, in turn, be interpreted by the brain as sound including speech. However, the behavioral benefits gained from CIs vary significantly among recipients; after cochlear implantation, some users achieve highly improved speech perception even in challenging listening situations such as in background noise while others gain very little or no improvement. Nonetheless, the source of the variability in CI performance is still unknown. In general, factors explaining this variation in individual speech perception ability include the bottom-up processing of the auditory periphery to acoustic features (including spectral and temporal information) and top-down cognitive processing at the cortex level (Moberly et al., 2016). However, demographic factors such as age at implantation and duration of deafness merely explain 20% of the variability in CI outcomes (Lazard et al., 2012).

At present, there are no reliable clinically available biomarkers for measuring CI outcomes to help us understand the source of outcome variability. Since the age at implantation can be as low as 1 year of age, developing objective markers is important for assessing pediatric CI users and candidates who have unreliable behavioral responses. Currently used objective measures such as the stapedius reflex, electrically evoked compound action potentials, and electrically evoked auditory brainstem responses have shown poor correlation with speech perception (Abbas and Brown, 1991; Hirschfelder et al., 2012; Lundin et al., 2015). Unlike these peripheral measures, cortical activity measured at the sensory and source levels has nevertheless shown some reliable relationships with behavioral performance in adult CI users in research settings (Han et al., 2016; Gransier et al., 2019).

Psychoacoustic studies have shown that speech perception through a CI relies predominantly on temporal cues because spectral information cannot be effectively delivered due to a limited number of spectral channels and channel interactions (Shannon et al., 1995; Nie et al., 2006). A CI processes the incoming sound, including speech, by applying a series of filter banks to extract the temporal envelope. This envelope then modulates the amplitude of a pulse train that stimulates the auditory nerve. Speech inherently has amplitude modulation (AM) at multiple rates with syllables in the 1–4 Hz range, phonemic information in the 15–50 Hz range and fine structure at higher rates (Rosen, 1992). Therefore, encoding AM is an important feature needed for successful speech perception (Fu, 2002; Edwards and Chang, 2013). Temporal processing is assessed behaviorally by estimating the minimum AM depth needed to detect modulation at various AM rates. The resulting behavioral AM threshold as a function of rate is referred to as the temporal modulation transfer function (TMTF). The shape of the TMTF resembles a low-pass filter with a cut-off frequency near 50–100 Hz (Viemeister, 1979). Compared with normal-hearing (NH) individuals, the TMTF of CI users has a higher overall AM threshold that is more pronounced at higher frequencies resulting in a lower frequency TMTF filter cutoff and subsequently this property is associated with reduced speech perception ability (Won et al., 2011). The ability to detect high-frequency AM (50–300 Hz) is correlated to speech perception in CI users including tone (Luo et al., 2008), consonants (Cazals et al., 1994), and word recognition (Won et al., 2011) and phonemes (De Ruiter et al., 2015). Recently, low frequency AM rate discrimination at 4 Hz shortly after CI activation time was shown to be a predictor of speech perception at 6 months post-activation (Erb et al., 2019).

Previously, we showed that in NH listeners, the N1 cortical evoked potential to AM changes resembles a low-pass filter shape, and the "N1 TMTF" is similar in shape to the behavioral TMTF (Han and Dimitrijevic, 2015). In that study, the N1 acoustic change complex (ACC) to AM changes were smaller at high versus low AM rates. In the present study, we wanted to determine if N1 ACC responses to AM could be elicited in CI users. We hypothesized that the N1 ACC to AM would be related to speech perception ability in CI users.

MATERIALS AND METHODS

Subjects

Ten adult CI users (five females, all self-reported right-handed) were recruited through Cincinnati Children's Hospital Medical Center according to an Institutional Review Board (IRB)-approved protocol. Their ages ranged from 21 to 84 years (mean age: 50 years). All CI subjects were native speakers of American English based on self-report and had been using his/her CI for at least 1 year prior to enrolling in the study. All CI subjects were postlingually deafened and had severe to profound bilateral hearing loss prior to implantation. They were all bilateral CI users. **Table 1** shows the demographic information of the CI users. A composite score based on the average percent scores over a number of speech perception tasks in background noise was the basis for classifying "good" and "poor" performers. There were six good performers with composite speech perception scores $\geq$50% and four poor users with scores <50%. For the control group (data from a previous study, Han and Dimitrijevic, 2015) 10 healthy NH individuals (six females, mean age = 25.5 years) were recruited. All of them were right-handed and had an audiometric hearing threshold of $\leq$20 dB HL (hearing level) at octave test frequencies from 250 to 8000 Hz. Participants were compensated for their participation, and informed consent was obtained from all of them prior to participation in the study.

Behavioral Testing

The TigerSpeech software (House Ear Institute)[1] was used for the behavioral testing. Consonant and vowel perceptions were measured using a forced-choice paradigm based on a previous report (Fu, 2002). Each of 16 consonants was presented five times ("a/Consonant/a" format, male voice), giving a total of 80 tokens. Similarly, each of 60 vowels was presented five times ("h/Vowel/d" format, male voice), giving a total of 60 vowels. Participants were instructed to indicate which consonant or vowel was heard by choosing the appropriately labeled button on the computer screen, and the performances on the vowel and

[1] http://www.tigerspeech.com/

TABLE 1 | Clinical features of the 10 cochlear implant participants.

CI user	Age (years)	Gender	Stimulated ear	Duration of deafness (year)	CI use (year)	Device/Processor	Processing strategy	Etiology of hearing loss	Composite score	Study group
01	21	M	Left	10	9	Nucleus/CI24RE	ACE	Unknown	78	Good
02	32	F	Left	20	9	Nucleus/CI24RE	ACE	Congenital	26	Poor
03	34	F	Right	33	12	Nucleus/Esprint 22	SPEAK	Hereditary	24	Poor
04	37	F	Right	37	11	Nucleus/CI24RE	ACE	Congenital	16	Poor
05	45	F	Right	37	4	Nucleus/CI512	ACE	Unknown	48	Poor
06	54	M	Left	15	4	Med El/Opus 2	FSP	Meniere's Disease	59	Good
07	59	M	Right	11	1	Nucleus/CI24RE	ACE	Noise induced	55	Good
08	63	F	Right	35	3	Nucleus/CI512	ACE	Genetic	55	Good
09	69	M	Left	22	2	Med El/Opus 2	FSP	Genetic	51	Good
10	84	M	Left	12	10	Nucleus/CI24RE	ACE	Unknown	56	Good

consonant perception tasks were quantified as percent correct. Sentence and word perceptions were measured using the SPIN (Speech-in-Noise) test (Kalikow et al., 1977). A total of 50 sentences were presented and participants were instructed to repeat each word in the sentence. The number of keywords (the terminal word in a sentence) correctly identified out of 50 was expressed as a percentage. We chose to proceed with electrophysiological testing on the CI side with the higher speech composite score.

The behavioral threshold for AM detection at 4, 40, 100, and 300 Hz was performed in a separate task using a three-interval forced choice with trial-by-trial feedback (Levitt, 1971). The task consisted of presenting three consecutive noise stimuli (1 s duration) one of which was amplitude modulated. The subjected needed to identify which interval had the AM stimulus. The AM depth was varied adaptively. The AM threshold refers to the minimum depth that the subject could detect the AM stimulus (average of the last nine reversals). The process was repeated for all four modulation rates. The depth of AM was defined as the percent ratio between maximum and minimum amplitudes such that 0% had no modulation, 100% was fully modulated (Picton et al., 2003).

Stimuli

Stimuli were constructed in Matlab using continuous white noise with occasional changes consisting of AM of 1-s duration occurring every 2.2 s on average (the random inter-stimulus interval varied from 1.8 to 2.6 s) and lasting for 1.0 s. Each stimulus with a change in AM as well as the baseline segment was generated from completely novel randomized noise in Matlab. The AM was changed at rates of 4, 40, 100, and 300 Hz. To avoid differences in the overall level that can occur when AM is introduced, the AM portion was multiplied by a factor that equated the root-mean-square of the preceding 1 s (no modulation).

Stimuli were presented in free field through a single speaker at 0° azimuth 1.5 m away from the subject. All stimuli were presented at the most comfortable level for each subject. To estimate the loudness of the stimuli for CI users, an intensity corresponding to loudness level of "7" on an 11-point scale (a 0 to 10: inaudible to too-loud linear scale) was applied (Hoppe et al., 2001). The stimuli were presented to the NH listeners at 70 dB SPL, while the intensity level was variable (70 to 85 dB SPL) for the CI users. The stimuli were calibrated using a Brüel and Kjaer (Investigator 2260) sound level meter set on both A and slow-time weighting with a half-inch free-field microphone.

Recordings

The electrophysiological data were collected using a 64-channel actiCHamp Brain Products recording system (Brain Products GmbH, Inc., Munich, Germany). Although our CI users were bilaterally implanted, the electrophysiological testing was carried out using one of the CIs while the other was turned off. The side with the higher speech composite score was used for all testing, yielded a total of five on each side. An electrode cap was placed on the scalp with electrodes placed at equidistant locations, the infracerebral cap covering a larger area than is typical in a 10–20 system (Hine and Debener, 2007; Han and Dimitrijevic, 2015). The reference channel was located at the vertex (Cz) while the ground electrode was located on the midline 50% of the distance to the nasion. Continuous data were digitized at 1000 Hz and stored for offline analysis.

Data Processing

Electrophysiological data were analyzed using Brain Vision Analyzer ver. 2.0 (Brain Products GmbH, Inc., Munich, Germany). Data were high-pass filtered (0.01 Hz) to remove baseline drift and down-sampled to 512 Hz. Visual inspection of the data included the removal of extreme stereotypical artifacts related to subject movement (exceeding 500 mV). Independent component analysis (Delorme and Makeig, 2004) implemented in Brain Vision Analyzer (with an identical algorithm to EEGLAB; Delorme and Makeig, 2004) was applied to reduce ocular and cardiac as well as CI artifacts. This approach decomposed the electroencephalographic (EEG) signal into maximally temporally independent components (ICs). Afterward, when an IC was deemed to be an artifact, its corresponding IC weight was set to zero, thereby minimizing its contribution to the data. In this study, ICs related to the CI were removed when the IC waveform morphology had an

abrupt peak within ∼10 ms of the onset/offset of the sound and resembled the AM envelope. The topography of the ICs showed an activation centroid near the location of the CI. Another indication of CI artifact was component energy at the AM modulation frequency. This was performed by computing the frequency spectrum of the IC. The IC with highest energy at the AM rate was removed. This procedure was helpful for CI artifact identification especially at the higher modulation rates (100 and 300 Hz). On average, five ICs or less were removed per CI subject.

After IC artifact reduction, the channel data for the electrodes near the CI were interpolated, the data referenced to average reference, and segmented into epochs −200 to 1500 ms with the AM change stimulus occurring at 0 ms and averaged. The auditory N1 responses, observed by pooling three electrodes in the frontal-central (FC) regions. Manual peak identification occurred over latencies in the 100–200 ms range. Peaks were verified by examining topography and polarity inversions at the mastoid. If no N1 peak was apparent, then this data was considered missing and was not analyzed further.

Procedures

During the EEG recording, participants were seated in a sound-attenuated booth, asked to watch a silent, closed-captioned movie of their choice, and instructed to ignore the background sounds. A total of 400 trials for each of the four AM change stimulus frequencies were conducted across eight blocks. The total recording time was approximately 1.5 h, and subjects were encouraged to take breaks between blocks.

Statistical Analysis

Repeated-measures analysis of variance (ANOVA) was used to assess statistical significance for both the psychoacoustics and EEG recordings. Details of the repeated-measures ANOVA factors are given with the results. The non-parametric Mann-Whitney *U* test was conducted to compare differences between the good and poor CI groups, along with *post hoc* analysis using Tukey's honest significant difference test. Spearman's rank-order correlation was computed to examine relationships between the speech test scores and the N1 amplitude/latency measures.

RESULTS

Psychoacoustics

The minimum AM depth needed for detection of modulation for 4, 40, 100, and 300 Hz was, on average, 44, 37, 49, and 77%, respectively (**Figure 1**) where greater values indicate poorer performance requiring higher modulation depth for detection. The repeated-measures ANOVA revealed a main effect for AM rate [$F(3,27) = 37.7$, $p = 0.0001$], while the *post hoc* analysis showed that the AM threshold for 300 Hz was significantly higher than those for 4 Hz ($p = 0.0002$), 40 Hz ($p = 0.0002$), and 100 Hz ($p = 0.0002$). No significant difference in AM threshold was found between 4, 40, and 100 Hz ($p > 0.05$).

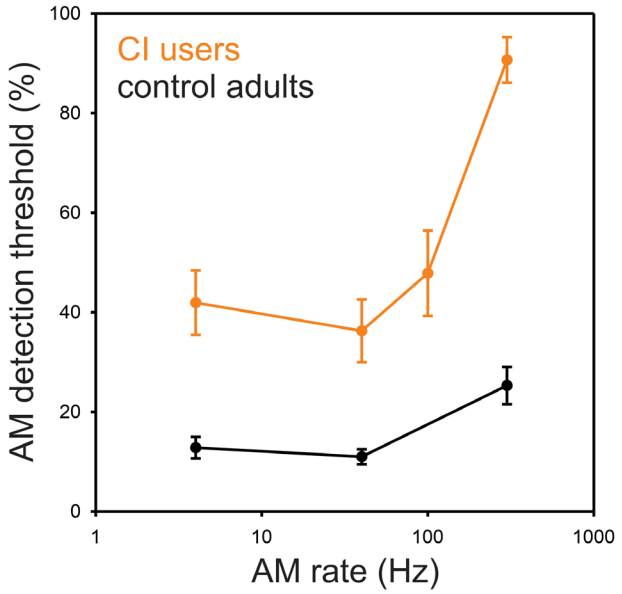

FIGURE 1 | Behavioral AM detection thresholds as a function of AM rate in CI and normal-hearing groups. Shown are the mean detection thresholds across 10 CI and 10 NH participants. Note that the AM detection thresholds were measured at 4, 40, 100, and 300 Hz for CI users, while the thresholds at 100 Hz were not measured for NH participants. The AM detection thresholds in CI users were higher than NH for all AM rates. NH data redrawn from Han and Dimitrijevic (2015).

Cortical Potentials

AM Change: CI vs. NH

Grand mean data are shown in **Figure 2A** illustrating the cortical potentials at FC electrodes for the AM changes at 4, 40, 100, and 300 Hz with a schematic of the stimulus overlaid. The N1 responses to AM change were robust in some cases, although not all CI participants had a measurable response. The N1 responses from CI users for AM changes at the four frequencies were as follows: all of them at 4 Hz, nine at 40 Hz, eight at 100 Hz, and five at 300 Hz. The N1 responses occurred close to 150 ms after the AM change but its peak latency was prolonged with an increase in AM rate. The NH data (redrawn from Han and Dimitrijevic, 2015) shows an "off" response to the change (i.e., 100% AM change back to 0% AM) change at about 1.2 s. This was not observed in the CI data.

In general, N1 responses in the CI group decreased in amplitude and their latency was increased compared to the NH group (**Figure 2B** shows the N1 amplitudes and latencies as a function of AM rate for the NH and CI groups). In NH listeners, the N1 amplitude was the greatest at 40 Hz whereas the amplitudes decreased from 4 Hz to 40 Hz for CI users. In addition, the N1 latencies in the CI users were modulated as a function of AM rate, while no latency differences revealed for NH listeners.

Repeated-measures ANOVA was used to examine the effect of AM rate (4, 40, and 300 Hz) and group (CI vs. NH) for N1

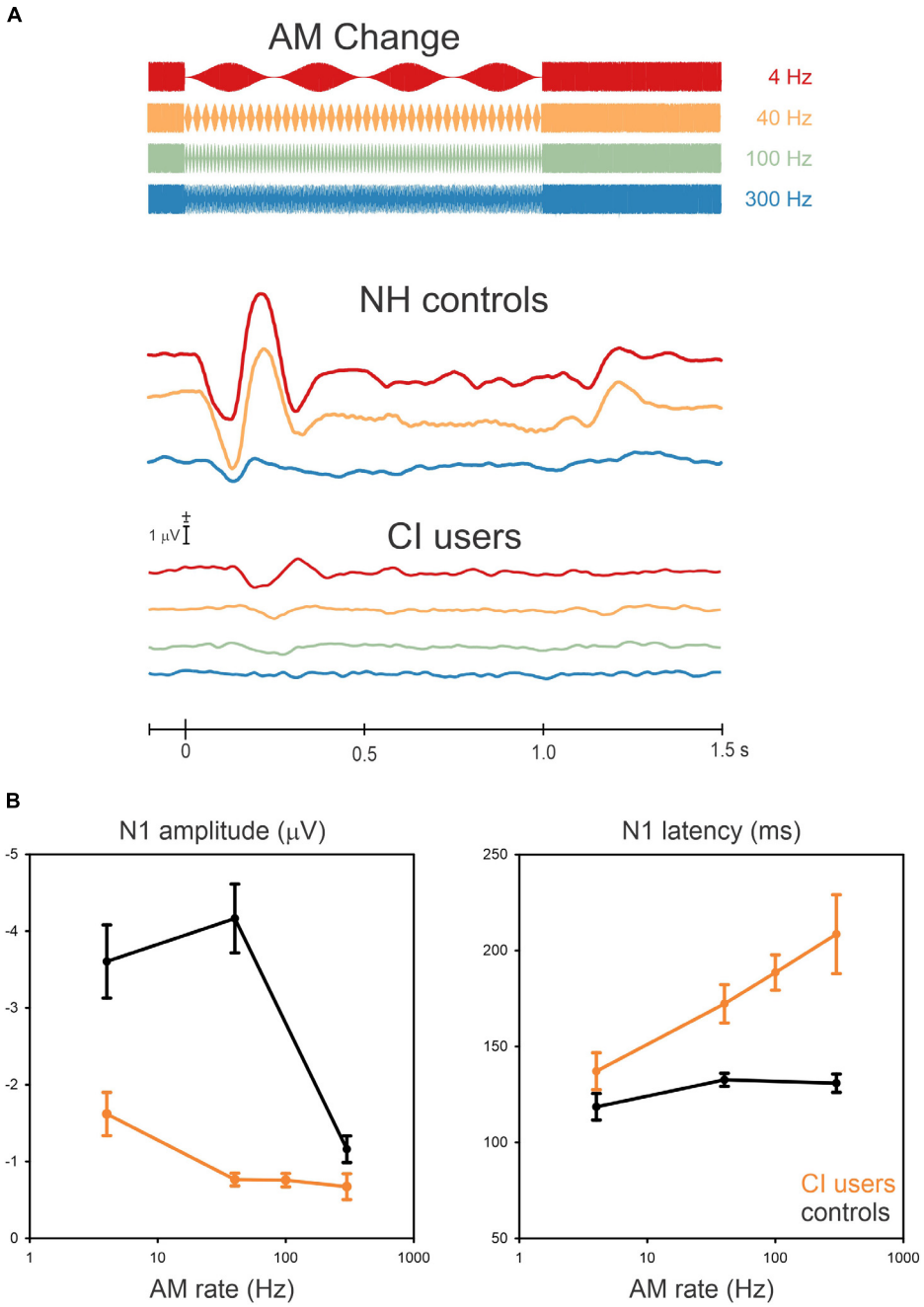

FIGURE 2 | Grand mean waveforms to the AM change stimulus **(A)** and mean N1 amplitudes and latencies **(B)** are shown for NH controls and CI users. **(A)** shows responses recorded at frontal-central electrodes to the 4 (red), 40 (yellow), 100 (Green, only for CI users), and 300 Hz (blue). **(B)** shows the mean averaged N1 amplitude and latency as a function of AM rate across 10, NH and 10 CI subjects. Error bars: standard error of the mean. Overall, AM amplitudes in CI users are smaller and delayed compared to NH for all AM rates.

amplitude and latency. For N1 amplitude, there was a significant main effect for AM rate [$F(2,36) = 46.4$; $p < 0.0001$] as well as group [$F(1,18) = 42.5$; $p < 0.0001$]. Meanwhile, the *post hoc* analysis showed that for the CI group, the N1 amplitude at 4 Hz was significantly larger than at 40 Hz ($p = 0.007$), 100 Hz ($p = 0.007$), and 300 Hz ($p = 0.0003$). Regarding the group effect, the *post hoc* testing revealed that the N1 amplitudes in the NH group were larger than the CI group ($p = 0.0002$), and for N1 latency, a significant effect of AM rate [$F(2,36) = 23.4$; $p < 0.0001$] was found such that the N1 latencies increased as the AM rate increased. The *post hoc* analysis also revealed that the N1 latency at 4 Hz was shorter than at 100 Hz ($p = 0.001$)

and 300 Hz (p = 0.0001), while the N1 latency at 40 Hz was shorter than at 300 Hz (p = 0.003). No significant differences were found between 4 and 40 Hz, 40 and 100 Hz, and 100 and 300 Hz ($p > 0.05$). A significant group effect was also found for N1 latency [$F(2,36) = 31.3$; $p < 0.0001$], with the analysis showing that the N1 latencies for the CI group were delayed compared to the NH group (p = 0.0002).

AM Change: Good vs. Poor CI Performers

Statistical analysis for a comparison between the good and poor CI groups was conducted for the 4 and 40 Hz AM rates only because the N1 responses at 100 and 300 Hz were not measurable in the majority of the CI subjects. For the N1 latency at 40 Hz, a significant group difference was observed such that the latencies for the good CI group were shorter than those for the poor CI group (U = 2.00; p = 0.04). **Figure 3** shows the latencies for the good (n = 6) and poor (n = 4) CI performers for AM at 4 and 40 Hz. No other differences between the good and poor CI groups were found ($p > 0.05$).

N1-Behavior Relationship

Figure 4 shows significant negative Spearman correlations between N1 latency for the 40 Hz AM rate and various speech perception measures including vowel (r = −0.75; $p < 0.05$), consonant (r = −0.82; $p < 0.05$), word (r = −0.74; $p < 0.05$), and sentence (r = −0.71; $p < 0.05$) perception in quiet conditions, as well as vowel (r = −0.84; $p < 0.05$) and word (r = −0.72; $p < 0.05$) perception in noise. The results indicate that shorter N1 latencies for AM at 40 Hz were associated with higher speech perception in the CI users. No significant relationships were observed for the N1 responses at different AM rates and behavioral thresholds in AM change detection (although 40 Hz AM detection threshold versus 40 Hz N1 amplitude approached significance (r = 0.59; p = 0.09).

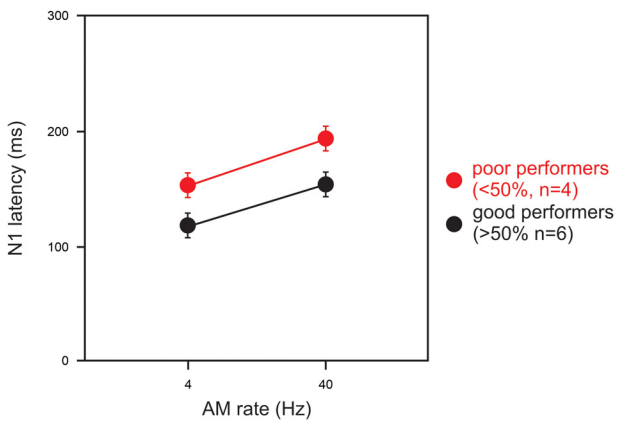

FIGURE 3 | A comparison of N1 latencies between good and poor CI performers. Good performers (n = 6) had composite speech perception scores above 50% and poor users (n = 4) had scores below 50%. Note the N1 latency in poor CI performers were delayed than good CI performers for 4 and 40 Hz AM. Note that the N1 latency for 100 and 300 Hz AM were not shown since not all subject had responses for the AM rates. Errors bars: standard error of the mean.

Correlation analysis was also performed between N1 amplitudes/latencies and demographic variables such as subject age and duration of deafness, no significant relationships were observed.

DISCUSSION

The present study examined the N1 ACC-to-AM change in CI users and revealed four findings. First, although the overall N1 amplitudes were smaller for the CI group, the N1 responses to AM change were robust for low AM frequencies but less so for high ones; this pattern of N1 activity is similar to the psychoacoustic TMTFs in that the AM thresholds are low at slow AM rates and high at fast AM rates. The N1 TMTF pattern in the NH group resembled a low-pass filter shape whereas for CI users this shape was not observed. Second, N1 latency increased with an increase in AM rate. Third, for the AM rates at 4 and 40 Hz, the N1 latencies were longer for the poor CI performers compared to the good performers. Finally, there was significant correlation between the N1 latency for the AM rate at 40 Hz and speech perception.

AM Change as a Paradigm to Assess Cortical Temporal Processing in CI Users

Previously, we developed a novel paradigm to quantify how the central auditory system encodes the detection of AM (Han and Dimitrijevic, 2015). The selected AM rates were based on timescales relevant for speech: syllables occur at slow rates near 4 Hz, formant transitions at 40–100 Hz, and fine structure near 300 Hz (Rosen, 1992). The TMTF quantifies temporal processing by measuring the ability to detect small temporal modulations in a sound as a function of AM rate. In CI users, a larger decay of the AM rate in behavioral AM thresholds has been previously observed compared to the NH control (Cazals et al., 1994; Won et al., 2011). For a direct comparison, we normalized the N1 and behavioral TMTFs in CI users using a similar approach to our previous report (Han and Dimitrijevic, 2015) and plotted the results in **Figure 5**. The CI behavioral TMTF resembles a low-pass filter shape similar to our previous NH data (Figure 10; Han and Dimitrijevic, 2015). However, in contrast to our previous findings in NH, the CI N1 did not have low-pass filter shape rather it continued to decrease in amplitude with increasing AM rate. The reasons for this discrepancy between the behavioral and N1 TMTF are not clear. One possibility is that they are measured differently. Behavioral TMTFs quantify the minimum AM depth needed for detection of modulation whereas the N1 response we recorded was a suprathreshold, 100% AM depth stimulus. Perhaps using AM depths closer to behavioral threshold may reveal N1 TMTF functions resembling those of behavioral TMTFs. The driving factor for the N1 TMTF low-pass filter shape in NH is that the response to 40 Hz is large and similar in magnitude to the 4 Hz response. In CI users, the 40 Hz AM change response was smaller than the 4 Hz response thus yielding a linear function. This pattern is in contrast to electrically evoked ASSRs (EASSRs) in CI users where 40 Hz responses are larger than 4 Hz (Luke et al., 2015) and represents a

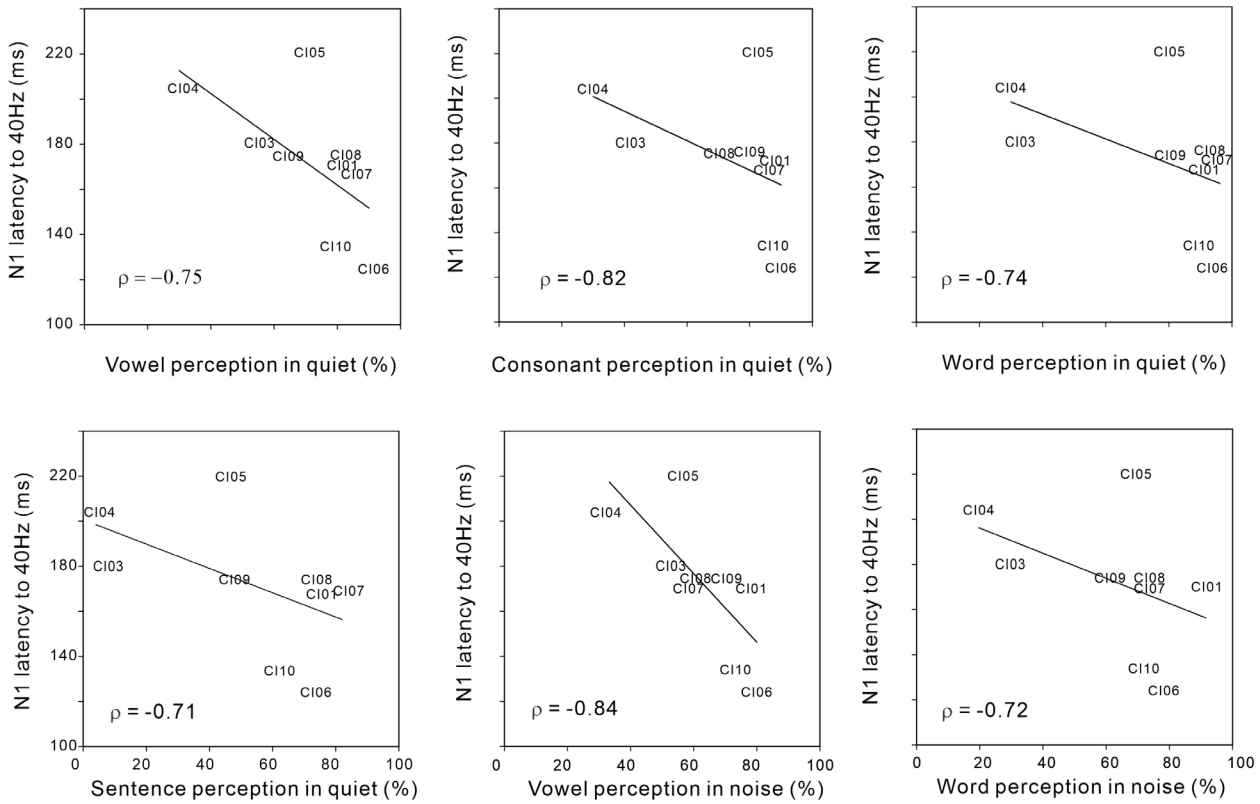

FIGURE 4 | Significant Spearman correlations between N1 latency to 40 Hz AM and various speech perception measures in CI users. Note that the N1 latency to 40 Hz AM decreased as speech perception performances were better.

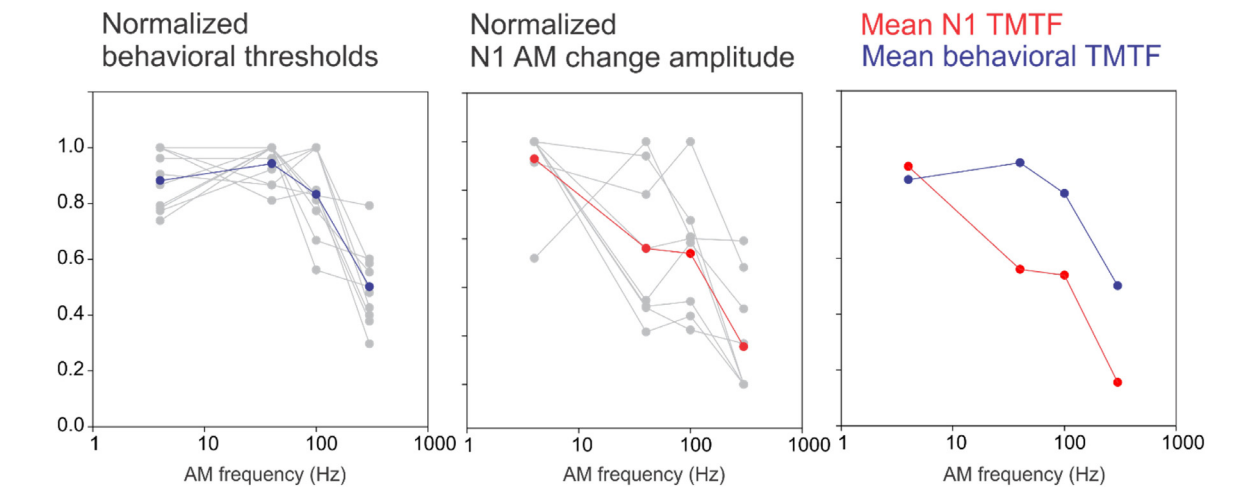

FIGURE 5 | A comparison between an N1-based TMTF and a behavioral-based TMTF in CI users. For the behavioral normalization, the smallest AM detection threshold (across the four AM rates) for each subject was used as a "reference" and all other AM depth thresholds were calculated as a ratio difference from the reference. Individual normalized behavioral AM detection thresholds are shown in gray while the mean across subjects is shown in blue. A similar process was performed for N1 amplitude except that the maximum amplitude was used a reference and all other responses (at the other AM rates) were normalized as a proportion difference from the max. The middle plot shows single subjects (gray) and mean across subjects (red). The right plot compares the mean behavioral and N1 TMTFs. Note the N1 TMTF pattern in the NH resembled a low-pass filter shape whereas in CI users, the sensitivity decreased with increasing AM rate.

temporal processing difference between ASSRs and cortical N1s. The 40 Hz N1 change response, nonetheless, by itself indexes temporal sensitivity and is related to speech perception outcomes. Another potential source of the discrepancy between the shapes of the TMTF is individual variability of 40 Hz N1 response. Inspection of the normalized N1 TMTF (**Figure 5**) suggests that 3 CI users had a low pass filter function shape while the others had decreasing functions. However, this does not relate to individual performance (i.e., two of the three low pass filter functions came from poor performers), nor does this explain why all of the behavioral TMTFs are low-pass filter shaped. Another possibility for the behavioral-N1 TMTF discrepancy is subject state sensitivity. The behavioral TMTF requires focused attention to the stimulus whereas the N1 TMTF was recorded in a passive listening paradigm. The N1 response is known to increase with attention (Hillyard et al., 1973; Picton and Hillyard, 1974) and different N1 TMTF profiles are likely to occur with attention in CI users. This interpretation would suggest that effects of attention are differentially modulated in NH versus CI users which in itself deserves further attention.

As a subtype of temporal processing, temporal resolution includes various auditory tasks such as temporal order judgment (Tallal, 1980), gap detection (Fitzgibbons and Wightman, 2005), detection of AM (Viemeister, 1979). It is well-known that the information extracted from the temporal envelope (a slow-rate temporal component among the temporal features) is necessary for speech understanding (Rosen, 1992; Drullman et al., 1994). The temporal envelope is even more important for CI users because the CI cannot extract adequate spectral information due to a limited number of frequency channels (Shannon et al., 1995; Fu, 2002), whereas low frequency temporal information is relatively well delivered through the CI. Since behavioral studies have shown that the ability to detect temporal variations has a strong correlation with speech perception (Won et al., 2011; De Ruiter et al., 2015), there has been an effort to measure how the brain processes temporal variations using auditory-evoked responses such as the ASSR and the mismatch negativity response. Using EASSRs to AM pulse trains of 4 and 40 Hz, Luke et al. (2015) found that the EASSR amplitudes at 40 Hz were related to the AM detection thresholds in five CI users and suggested the clinical significance of EASSR as an objective measure of site-specific temporal sensitivity for CIs. Very recently, Gransier et al. (2019) found that 40-Hz EASSR variability across CI electrodes was highly correlated to speech perception in CI users. In addition, Waechter et al. (2018) found that the morphology-weighted mismatch waveform evoked by a stimulus with 8-Hz modulation is positively correlated with the AM detection threshold. Their results also suggest that cortical responses strongly follow a low-rate AM. These neurophysiological results indicate that speech perception by CI users is largely dependent on temporal information and that the auditory-evoked responses elicited by AM reflect the neuronal modulation for temporal acoustic variations. In contrast to the ASSRs, in this study, we chose to study brain responses underlying detection of AM using the N1 AM-change response.

We found that the N1 responses of the CI users decreased in amplitude as the AM rate increased to a greater degree than occurred in the NH control. In addition, the N1 latency in the CI users was almost linearly modulated as a function of AM rate, a phenomenon that was not observed in the NH group. The effect of temporal variation on N1 responses has been assessed in previous studies using various temporal features, including voice onset time (Roman et al., 2004; Dimitrijevic et al., 2013; Han et al., 2016), musical/pitch matching (Timm et al., 2012; Tan, 2017), and the temporal gap (He et al., 2018). The common finding of these studies was that the N1 response was delayed according to the delay in the onset of a sound (e.g., a long duration of voice onset time). For example, using different musical onset durations, Timm et al. (2012) found that N1 latency was longer when the onset time of a musical tone was shorter; the authors suggest that N1 latency is more sensitive to temporal change than to N1 amplitude. A recent study (Han and Dimitrijevic, 2017) examined cortical responses to varied voice onset time during passive listening also showed the linear modulation of N1 latency as a function of voice onset time. Interestingly, the more linear and consistent the N1 change with increases in voice onset time, the greater the speech perception score. This suggests that in CI users, greater sensitivity to acoustic temporal fluctuation was associated with better the speech perception outcome.

In the current study, the N1 amplitudes of the NH group were larger than those of the CI group, regardless of the AM rate. Smaller and delayed peaks are distinct characteristics of cortical responses in CI users (Beynon et al., 2005; Sandmann et al., 2009), and a decreased N1 amplitude is related to the reduced neuronal population recruited to process sounds synchronously or to how the timing and frequencies are coded at the cortex (Guiraud et al., 2007; Tremblay and Ross, 2007). However, a weak response is not always the case for CI users. Previous studies on CI use have suggested that the magnitude of cortical responses is closely related to CI speech outcomes: good CI performers revealed greater cortical responses while poor CI users attained smaller or absent peaks (Groenen et al., 1996; Kelly et al., 2005). Similarly, significant N1 latency differences between good and poor CI performers were revealed in the present study. Brain plasticity associated with hearing loss has been suggested to underlie the cortical activation pattern with hearing loss and/or with CI use (Pantev et al., 2006; Stropahl et al., 2017). However, the degree of brain plasticity can be different among CI users depending on demographic factors and environmental influences, including rehabilitation.

Although we hypothesized that the N1 TMTF would resemble the behavioral TMTF in CI users, this does not appear to be the case. The N1 response decreased with increasing AM rate suggesting neural encoding progressively decreases with faster temporal modulations. More research on the reasons for the discrepancy between behavioral and neural TMTF is warranted. This could include using AM depths closer to behavioral threshold or attentive listening paradigms.

Cortical Responses to AM Change and Behavioral Performance in CI Users

We found that N1 latency for AM at 40 Hz was increased in the poor performing CI users compared to the good performing

ones and was correlated with various speech perception measures in the CI users. Previously, it has been shown that the N1 response to simple onset sounds such as a tone burst or click is poorly related to speech perception in CI users (Firszt et al., 2002; Kelly et al., 2005). One possible explanation for this is that the N1 response is related to the detection of sound rather than its discrimination. Because speech understanding needs both detection and discrimination of sounds, many studies have focused on the cortical measures for discrimination, including mismatch negativity, P300, and ACC. Among these, ACC is evoked by changes in various stimuli such as speech (Tremblay et al., 2003; Dimitrijevic et al., 2011; Small and Werker, 2012), tone (Dimitrijevic et al., 2008, 2009), and noise (Martin et al., 1999; Han and Dimitrijevic, 2015). The ACC can be modulated as a function of frequency change and is related to the behavioral threshold for frequency discrimination (Dimitrijevic et al., 2008). In CI users, the ACC can be elicited by speech (Friesen and Tremblay, 2006; Han et al., 2016), an intensity change in the CI electrodes (Kim et al., 2009), as well as a frequency change in magnetoencephalography (Pantev et al., 2006). Moreover, the cortical responses have been successfully applied to evaluate the optimization of CI fitting in single-sided deafness (Távora-Vieira et al., 2018). These results indicate that the ACC can be reliably recorded in CI users and that the magnitude of cortical response increases with an improvement in behavioral performance. In our study, we applied the AM change paradigm to evoke the N1 ACC and attempted to correlate it with behavioral measures. The results are not surprising given that AM detection thresholds have previously shown strong correlations with various speech measures such as vowel and consonant perception (Cazals et al., 1994; Fu, 2002), phoneme perception (Xu and Zheng, 2007), and word perception (Won et al., 2011). Thus, the ACC in response to AM change can effectively reflect how the central auditory system encodes a change in AM rate, which is critical for speech understanding. This is supported by the notion that poor time-locking to the detection of the temporal envelope could be related to poor discrimination of temporal variation (Joris et al., 2004). Surprisingly, in contrast to N1 responses to frequency change (Dimitrijevic et al., 2008), no significant relationships were observed between AM behavioral thresholds and N1 latency or amplitude. Further work on AM-change-related N1/ACC responses using varying degrees of AM depth may reveal stronger relationships with behavior compared to the 100% AM depth used in the current study.

Clinical Applications

In the present study, we showed that AM change stimuli can elicit robust cortical ACC responses (4 and 40 Hz) in CI users and the N1 latency to 40 Hz is related to speech perception measures. A larger sample of CI users is needed to determine if these findings generalize to more diverse CI populations. Interestingly only the 40 Hz N1 response showed a significant relationship with behavior while the other rates did not, even though the 4 Hz N1 response was robust. Given that behavioral TMTFs relate well to speech perception understanding in CI users, further research N1 TMTFs is warranted.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Cincinnati Children's Hospital Medical Center. The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AD: experimental design, analysis, and manuscript preparation. J-HH: analysis, experiment execution, and manuscript preparation.

FUNDING

This project was supported in part by the internal grant mechanism of Cincinnati Children's Hospital Medical Center and Basic Science Research Program through the National Research Foundation of South Korea (NRF) funded by the Ministry of Education (2017R1D1A1B03030613).

REFERENCES

Abbas, P. J., and Brown, C. J. (1991). Electrically evoked auditory brainstem response: refractory properties and strength-duration functions. *Hear. Res.* 51, 139–147. doi: 10.1016/0378-5955(91)90012-x

Beynon, A. J., Snik, A. F. M., Stegeman, D. F., and van den Broek, P. (2005). Discrimination of speech sound contrasts determined with behavioral tests and event-related potentials in cochlear implant recipients. *J. Am. Acad. Audiol.* 16, 42–53. doi: 10.3766/jaaa.16.1.5

Cazals, Y., Pelizzone, M., Saudan, O., and Boex, C. (1994). Low-pass filtering in amplitude modulation detection associated with vowel and consonant identification in subjects with cochlear implants. *J. Acoust. Soc. Am.* 96, 2048–2054. doi: 10.1121/1.410146

De Ruiter, A. M., Debruyne, J. A., Chenault, M. N., Francart, T., and Brokx, J. P. L. (2015). Amplitude modulation detection and speech recognition in late-implanted prelingually and postlingually deafened cochlear implant users. *Ear Hear.* 36, 557–566. doi: 10.1097/AUD.0000000000000162

Delorme, A., and Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *J. Neurosci. Methods* 134, 9–21. doi: 10.1016/j.jneumeth.2003.10.009

Dimitrijevic, A., Lolli, B., Michalewski, H. J., Pratt, H., Zeng, F. G., and Starr, A. (2009). Intensity changes in a continuous tone: auditory cortical potentials comparison with frequency changes. *Clin. Neurophysiol.* 120, 374–383. doi: 10.1016/j.clinph.2008.11.009

Dimitrijevic, A., Michalewski, H. J., Zeng, F. G., Pratt, H., and Starr, A. (2008). Frequency changes in a continuous tone: auditory cortical potentials. *Clin. Neurophysiol.* 119, 2111–2124. doi: 10.1016/j.clinph.2008.06.002

Dimitrijevic, A., Pratt, H., and Starr, A. (2013). Auditory cortical activity in normal hearing subjects to consonant vowels presented in quiet and in noise. *Clin. Neurophysiol.* 124, 1204–1215. doi: 10.1016/j.clinph.2012.11.014

Dimitrijevic, A., Starr, A., Bhatt, S., Michalewski, H. J., Zeng, F. G., and Pratt, H. (2011). Auditory cortical N100 in pre- and post-synaptic auditory neuropathy to frequency or intensity changes of continuous tones. *Clin. Neurophysiol.* 122, 594–604. doi: 10.1016/j.clinph.2010.08.005

Drullman, R., Festen, J. M., and Plomp, R. (1994). Effect of reducing slow temporal modulations on speech reception. *J. Acoust. Soc. Am.* 95, 2670–2680. doi: 10.1121/1.409836

Edwards, E., and Chang, E. F. (2013). Syllabic (˜2-5Hz) and fluctuation (˜1-10Hz) ranges in speech andauditory processing. *Hear. Res.* 305, 113–134. doi: 10.1016/j.heares.2013.08.017

Erb, J., Ludwig, A. A., Kunke, D., Fuchs, M., and Obleser, J. (2019). Temporal sensitivity measured shortly after cochlear implantation predicts 6-month speech recognition outcome. *Ear Hear.* 40, 27–33. doi: 10.1097/AUD.0000000000000588

Firszt, J. B., Chambers And, R. D., and Kraus, N. (2002). Neurophysiology of cochlear implant users II: comparison among speech perception, dynamic range, and physiological measures. *Ear Hear.* 23, 516–531. doi: 10.1097/00003446-200212000-00003

Fitzgibbons, P. J., and Wightman, F. L. (2005). Gap detection in normal and hearing-impaired listeners. *J. Acoust. Soc. Am.* 72:761. doi: 10.1121/1.388256

Friesen, L. M., and Tremblay, K. L. (2006). Acoustic change complexes recorded in adult cochlear implant listeners. *Ear Hear.* 27, 678–685. doi: 10.1097/01.aud.0000240620.63453.c3

Fu, Q.-J. (2002). Temporal processing and speech recognition in cochlear implant users. *Neuroreport* 13, 1635–1639. doi: 10.1097/00001756-200209160-00013

Gransier, R., Luke, R., van Wieringen, A, and Wouters, J. (2019). Neural modulation transmission is a marker for speech perception in noise in cochlear implant users. *Ear Hear.* doi: 10.1097/AUD.0000000000000783 [Epub ahead of print]

Groenen, P. A. P., Makhdoum, M., Van Den Brink, J. L., Stollman, M. H. P., Snik, A. F. M., and Van Den Broek, P. (1996). The relation between electric auditory brain stem and cognitive responses and speech perception in cochlear implant users. *Acta Otolaryngol.* 116, 785–790. doi: 10.3109/00016489609137926

Guiraud, J., Besle, J., Arnold, L., Boyle, P., Giard, M.-H., Bertrand, O., et al. (2007). evidence of a tonotopic organization of the auditory cortex in cochlear implant users. *J. Neurosci.* 27, 7838–7846. doi: 10.1523/JNEUROSCI.0154-07.2007

Han, J. H., and Dimitrijevic, A. (2015). Acoustic change responses to amplitude modulation: a method to quantify cortical temporal processing and hemispheric asymmetry. *Front. Neurosci.* 9:38. doi: 10.3389/fnins.2015.00038

Han, J.-H., and Dimitrijevic, A. (2017). Cortical auditory evoked potential is modulated by attention and related to speech perception abilities in cochlear implant users. *J. Hear. Sci.* 7, 54–54.

Han, J.-H., Zhang, F., Kadis, D. S., Houston, L. M., Samy, R. N., Smith, M. L., et al. (2016). Auditory cortical activity to different voice onset times in cochlear implant users. *Clin. Neurophysiol.* 127, 1603–1617. doi: 10.1016/j.clinph.2015.10.049

He, S., McFayden, T. C., Shahsavarani, B. S., Teagle, H. F. B., Ewend, M., Henderson, L., et al. (2018). The electrically evoked auditory change complex evoked by temporal gaps using cochlear implants or auditory brainstem implants in children with cochlear nerve deficiency. *Ear Hear.* 39, 482–494. doi: 10.1097/AUD.0000000000000498

Hillyard, S. A., Hink, R. F., Schwent, V. L., and Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science* 182, 177–180. doi: 10.1126/science.182.4108.177

Hine, J., and Debener, S. (2007). Late auditory evoked potentials asymmetry revisited. *Clin. Neurophysiol.* 118, 1274–1285. doi: 10.1016/j.clinph.2007.03.012

Hirschfelder, A., Gräbel, S., and Olze, H. (2012). Electrically evoked amplitude modulation following response in cochlear implant candidates: comparison with auditory nerve response telemetry, subjective electrical stimulation, and speech perception. *Otol. Neurotol.* 33, 968–975. doi: 10.1097/MAO.0b013e31825e7c5d

Hoppe, U., Rosanowski, F., Iro, H., and Eysholdt, U. (2001). Loudness perception and late auditory evoked potentials in adult cochlear implant users. *Scand. Audiol.* 30, 119–125. doi: 10.1080/010503901300112239

Joris, P. X., Schreiner, C. E., and Rees, A. (2004). Neural Processing of Amplitude-Modulated Sounds. *Physiol. Rev.* 84, 541–577. doi: 10.1152/physrev.00029.2003

Kalikow, D. N., Stevens, K. N., and Ellitt, L. L. (1977). Speech perception in noise test (SPIN). *J. Acoust. Soc. Am.* 61, 1337–1351.

Kelly, A. S., Purdy, S. C., and Thorne, P. R. (2005). Electrophysiological and speech perception measures of auditory processing in experienced adult cochlear implant users. *Clin. Neurophysiol.* 116, 1235–1246. doi: 10.1016/j.clinph.2005.02.011

Kim, J. R., Brown, C. J., Abbas, P. J., Etler, C. P., and O'Brien, S. (2009). The effect of changes in stimulus level on electrically evoked cortical auditory potentials. *Ear Hear.* 30, 320–329. doi: 10.1097/AUD.0b013e31819c42b7

Lazard, D. S., Vincent, C., Venail, F., Van de Heyning, P., Truy, E., Sterkers, O., et al. (2012). Pre-, per- and postoperative factors affecting performance of postlinguistically deaf adults using cochlear implants: a new conceptual model over time. *PLoS One* 7:e48739. doi: 10.1371/journal.pone.0048739

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *J. Acoust. Soc. Am.* 49, 467–477. doi: 10.1121/1.1912375

Luke, R., Van Deun, L., Hofmann, M., Van Wieringen, A., and Wouters, J. (2015). Assessing temporal modulation sensitivity using electrically evoked auditory steady state responses. *Hear. Res.* 324, 37–45. doi: 10.1016/j.heares.2015.02.006

Lundin, K., Stillesjo, F., and Rask-Andersen, H. (2015). Prognostic value of electrically evoked auditory brainstem responses in cochlear implantation. *Cochlear Implants Int.* 16, 254–261. doi: 10.1179/1754762815Y.0000000005

Luo, X., Fu, Q.-J., Wei, C.-G., and Cao, K.-L. (2008). Speech recognition and temporal amplitude modulation processing by mandarin-speaking cochlear implant users. *Ear Hear.* 29, 957–970. doi: 10.1097/AUD.0b013e3181888f61

Martin, B. A., Kurtzberg, D., and Stapells, D. R. (1999). The effects of decreased audibility produced by high-pass noise masking on N1 and the mismatch negativity to speech sounds /ba/and/da. *J. Speech. Lang. Hear. Res.* 42, 271–286. doi: 10.1044/jslhr.4202.271

Moberly, A. C., Bates, C., Harris, M. S., and Pisoni, D. B. (2016). The enigma of poor performance by adults with cochlear implants. *Otol. Neurotol.* 37, 1522–1528. doi: 10.1097/MAO.0000000000001211

Nie, K., Barco, A., and Zeng, F.-G. (2006). Spectral and temporal cues in cochlear implant speech perception. *Ear Hear.* 27, 208–217. doi: 10.1097/01.aud.0000202312.31837.25

Pantev, C., Dinnesen, A., Ross, B., Wollbrink, A., and Knief, A. (2006). Dynamics of auditory plasticity after cochlear implantation: a longitudinal study. *Cereb. Cortex* 16, 31–36. doi: 10.1093/cercor/bhi081

Picton, T. W., and Hillyard, S. A. (1974). Human auditory evoked potentials. II. Effects of attention. *Electroencephalogr. Clin. Neurophysiol.* 36, 191–199.

Picton, T. W., John, M. S., Dimitrijevic, A., and Purcell, D. (2003). Human auditory steady-state responses: respuestas auditivas de estado estable en humanos. *Int. J. Audiol.* 42, 177–219. doi: 10.3109/14992020309101316

Roman, S., Canévet, G., Lorenzi, C., Triglia, J. M., and Liégeois-Chauvel, C. (2004). Voice onset time encoding in patients with left and right cochlear implants. *Neuroreport* 15, 601–605. doi: 10.1097/00001756-200403220-00006

Rosen, S. (1992). Temporal information in speech: acoustic, auditory and linguistic aspects. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 336, 367–373. doi: 10.1098/rstb.1992.0070

Sandmann, P., Eichele, T., Buechler, M., Debener, S., Jäncke, L., Dillier, N., et al. (2009). Evaluation of evoked potentials to dyadic tones after cochlear implantation. *Brain* 132, 1967–1979. doi: 10.1093/brain/awp034

Shannon, R. V., Zeng, F. G., Kamath, V., Wygonski, J., and Ekelid, M. (1995). Speech recognition with primarily temporal cues. *Science* 270, 303–304. doi: 10.1126/science.270.5234.303

Small, S. A., and Werker, J. F. (2012). Does the acc have potential as an index of early speech-discrimination ability? a preliminary study in 4-month-old infants with normal hearing. *Ear Hear.* 33:e59-69. doi: 10.1097/AUD.0b013e31825f29be

Stropahl, M., Chen, L. C., and Debener, S. (2017). Cortical reorganization in postlingually deaf cochlear implant users: intra-modal and cross-modal considerations. *Hear. Res.* 343, 128–137. doi: 10.1016/j.heares.2016.07.005

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain Lang.* 9, 182–198. doi: 10.1016/0093-934X(80)90139-X

Tan, C. T. (2017). "A preliminary study to identify a neurophysiological correlate of electroacoustic pitch matching in cochlear implant users," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society, EMBS*, Piscataway, NJ.

Távora-Vieira, D., Wedekind, A., Marino, R., Purdy, S. C., and Rajan, G. P. (2018). Using aided cortical assessment as an objective tool to evaluate cochlear implant fitting in users with single-sided deafness. *PLoS One* 13:e0193081. doi: 10.1371/journal.pone.0193081

Timm, L., Agrawal, D., Viola, C. F., Sandmann, P., Debener, S., Büchner, A., et al. (2012). Temporal feature perception in cochlear implant users. *PLoS One* 7:e45375. doi: 10.1371/journal.pone.0045375

Tremblay, K., and Ross, B. (2007). Effects of age and age-related hearing loss on the brain. *J. Commun. Disord.* 40, 305–312. doi: 10.1016/j.jcomdis.2007.03.008

Tremblay, K. L., Friesen, L., Martin, B. A., and Wright, R. (2003). Test-retest reliability of cortical evoked potentials using naturally produced speech sounds. *Ear Hear.* 24, 225–232. doi: 10.1097/01.AUD.0000069229.84883.03

Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *J. Acoust. Soc. Am.* 66, 1364–1380. doi: 10.1121/1.383531

Waechter, S. M., Lopez Valdes, A., Simoes-Franklin, C., Viani, L., and Reilly, R. B. (2018). Depth matters - towards finding an objective neurophysiological measure of behavioral amplitude modulation detection based on neural threshold determination. *Hear. Res.* 359, 13–22. doi: 10.1016/j.heares.2017.12.005

Won, J. H., Drennan, W. R., Nie, K., Jameyson, E. M., and Rubinstein, J. T. (2011). Acoustic temporal modulation detection and speech perception in cochlear implant listeners. *J. Acoust. Soc. Am.* 130, 376–388. doi: 10.1121/1.3592521

Xu, L., and Zheng, Y. (2007). Spectral and temporal cues for phoneme recognition in noise. *J. Acoust. Soc. Am.* 122, 1758–1764. doi: 10.1121/1.2767000

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.

Copyright © 2020 Han and Dimitrijevic. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 18 February 2020
doi: 10.3389/fncom.2020.00011

Resonance in the Mouse Ventral Tegmental Area Dopaminergic Network Induced by Regular and Poisson Distributed Optogenetic Stimulation *in-vitro*

Luuk van der Velden[1], *Martin A. Vinck*[2] *and Wytse J. Wadman*[1]*

[1] *Swammerdam Institute for Life Sciences, University of Amsterdam, Amsterdam, Netherlands,* [2] *Ernst Strüngmann Institute (ESI) for Neuroscience in Cooperation With Max Planck Society, Frankfurt am Main, Germany*

Neurons in many brain regions exhibit spontaneous, intrinsic rhythmic firing activity. This rhythmic firing activity may determine the way in which these neurons respond to extrinsic synaptic inputs. We hypothesized that neurons should be most responsive to inputs at the frequency of the intrinsic oscillation frequency. We addressed this question in the ventral tegmental area (VTA), a dopaminergic nucleus in the midbrain. VTA neurons have a unique propensity to exhibit spontaneous intrinsic rhythmic activity in the 1–5 Hz frequency range, which persists in the *in-vitro* brain slice, and form a network of weakly coupled oscillators. Here, we combine *in-vitro* simultaneous recording of action potentials from a 60 channel multi-electrode-array with cell-type-specific optogenetic stimulation of the VTA dopamine neurons. We investigated how VTA neurons respond to wide-band stochastic (Poisson input) as well as regular laser pulses. Strong synchrony was induced between the laser input and the spike timing of the neurons, both for regular pulse trains and Poisson pulse trains. For rhythmically pulsed input, the neurons demonstrated resonant behavior with the strongest phase locking at their intrinsic oscillation frequency, but also at half and double the intrinsic oscillation frequency. Stochastic Poisson pulse stimulation provided a more effective stimulation of the entire population, yet we observed resonance at lower frequencies (approximately half the oscillation frequency) than the neurons' intrinsic oscillation frequency. The non-linear filter characteristics of dopamine neurons could allow the VTA to predict precisely timed future rewards based on past sensory inputs, a crucial component of reward prediction error signaling. In addition, these filter characteristics could contribute to a pacemaker role for the VTA in synchronizing activity with other regions like the prefrontal cortex and the hippocampus.

Keywords: dopamine, network, optogenetics, self-organization, rhythm, resonance, action potentials, multi-electrode array

OPEN ACCESS

Edited by:
Andreas Bahmer, Universitätsklinikum Würzburg, Germany

Reviewed by:
Nikolaos Karalis, Friedrich Miescher Institute for Biomedical Research, Switzerland
Lianchun Yu, Lanzhou University, China

***Correspondence:**
Wytse J. Wadman
w.j.wadman@uva.nl

Received: *19 September 2019*
Accepted: *28 January 2020*
Published: *18 February 2020*

Citation:
van der Velden L, Vinck MA and Wadman WJ (2020) Resonance in the Mouse Ventral Tegmental Area Dopaminergic Network Induced by Regular and Poisson Distributed Optogenetic Stimulation in-vitro. Front. Comput. Neurosci. 14:11. doi: 10.3389/fncom.2020.00011

1. INTRODUCTION

The ventral tegmental area (VTA) is a dopamine nucleus in the midbrain, alongside the substantia nigra and the red nucleus. VTA dopamine neurons exhibit spontaneous low-frequency rhythmic spike activity (Werkman et al., 2001; Bayer et al., 2007). The rhythm is intrinsic and is even retained in acute slices and in isolated VTA dopamine neurons (Koyama et al., 2005). Experimental research and modeling studies point to at least two distinct mechanisms that underlie this oscillatory activity (Khaliq and Bean, 2010; Drion et al., 2011): it is either based on sub-threshold calcium oscillations or on the presence of the persistent sodium current. VTA dopamine neurons have direct synaptic connections between each other but they also connect through a local interneuron network (Bayer and Pickel, 1990; Omelchenko and Sesack, 2009). In addition they are likely to interact through dopamine volume transmission (Cragg et al., 2001). This phenomenon implies that the firing neurons create an oscillating extracellular dopamine concentration that results in auto inhibition of the population and helps to synchronize all neurons (van der Velden et al., 2017). The VTA network can therefore be considered as a network of weakly coupled oscillators.

Oscillatory rhythms in the brain create the context for encoding information, specifically in the detailed timing of their firing (Buzsaki, 2009). The VTA is implicated in a low frequency local field oscillation (with spectral energy focused in the 1–5 Hz band) which synchronizes it with the prefrontal cortex and the hippocampus during memory processing (Battaglia and McNaughton, 2011; Fujisawa and Buzsáki, 2011). An additional hypothesis states that VTA pacemaker activity entrains the prefrontal cortex and the hippocampus (Fujisawa and Buzsáki, 2011). The timing information in its output plays an important role in stimulus-reward processing (Schultz, 1997) especially where activities need to be related at relatively long time scales.

Elucidating the population dynamics of the VTA dopamine neurons requires simultaneous recording of the spiking activity of as many neurons as possible, which can be accomplished in acute brain slices positioned on a multi-electrode-array (MEA). The slice is void of external input, which restricts the preparation, but also prevents complications from uncontrolled external input and thus simplifies the interpretation (Berretta et al., 2010; van der Velden et al., 2017). Specific activation of neurons at a very fast time scale (ms) can be accomplished with optogenetics. Crossing the floxed ChR2-EYFP mouse (Madisen et al., 2012) with the Pitx3 Cre driver mouse (Smidt et al., 2012) expresses light sensitive excitatory channelrhodopsin specifically in dopamine neurons. We can then use pulsed laser driven activation for precise manipulation of spike timing in individual neurons as well as at the level of the entire population (Deisseroth, 2010; Fenno et al., 2011) during long-term recordings without noticeable photo-toxicity (Cardin, 2012).

This study focuses on the lateral VTA, which mainly contains mesolimbic projecting dopamine neurons with a classic dopamine neuron phenotype (Lammel et al., 2008). The auto-oscillatory VTA dopamine neurons (van der Velden et al., 2017) and in particular their response to rhythmic and stochastic input were recorded. Their resonance characteristics were probed at the population level using regular (fixed frequency) as well as Poisson distributed (noisy) pulsed input. Driving populations of weakly coupled oscillators with such stimulation regimes has theoretically been worked out and reveals interesting dynamics (Hata et al., 2010). Experimental and modeling research showed that populations of oscillators can encode common input into their oscillatory output, even when the input is noisy (Ermentrout et al., 2008). We demonstrate here that the resonance characteristics of the VTA dopamine neuron population in response to temporally patterned stimulation reveal emergent properties at the network level.

2. METHODS

2.1. Experimental Animals

Adult male and female mice between 6 and 11 weeks old were used in the experiments. They were housed in a 12/12 light dark cycle and received water and food ad libitum. All procedures and methods were approved by the ethical committee for animal experimentation of the University of Amsterdam.

2.2. Optogenetic Expression

The Pitx3 Cre-driver mouse (Smidt et al., 2012) was crossed with the LoxP-ChR2-EYFP mouse (Madisen et al., 2012) to express channelrhodopsin-2 (ChR2) in the midbrain dopamine neurons. Both mice had a black-6 genetic background and were homozygous for their respective allele of interest. First generation offspring, heterozygous for both alleles (male and female), were used in the experiments. The Cre expression in the Pitx3-Cre mouse completely overlapped with the tyrosine hydroxylase expressing dopamine neurons in the lateral VTA (Smidt et al., 2012; Luk et al., 2013). In our F1 offspring we checked the presence of the two proteins in the lateral VTA dopamine neurons using fluorescent immuno-staining with GFP and Pitx3 anti-bodies, employing confocal and wide field imaging (see below).

2.3. Immunocytochemistry

Free floating acute brain slices (100 μm), were fixed for 1 h at room temperature with 4 % paraformaldehyde in phosphate buffered saline (PBS) pH 7.4. Subsequently washed with PBS and blocked for 4 h with 10 % normal goat serum plus 0.25 % Triton X-100 in PBS at room temperature. The slices were incubated overnight with the primary antibodies GFP (1:750; Abcam; ab13970) and Pitx3 [1:1000; Smidt et al. (2000)] diluted in 3 % normal goat serum plus 0.25 % Triton X-100 in PBS at 4°C. The following day, the slices were washed with PBS and incubated for 4 h with the secondary antibodies Goat anti-Rabbit Alexa Fluor 555 (1:1000; Molecular Probes; A-21428) and Goat anti-Chicken Alexa Fluor 488 (1:1000; Molecular Probes; A-11039) diluted in 1 % normal goat serum plus 0.025 % Triton in PBS at room temperature. Sections were washed with PBS and cover slipped with Vectashield including DAPI (Vector Laboratories, Burlingame, CA, USA). Both ChR2-EYFP and Pitx3 were expressed within dopamine nuclei in our midbrain preparation (**Figure 1A**). Confocal images showed Pitx3 expression within

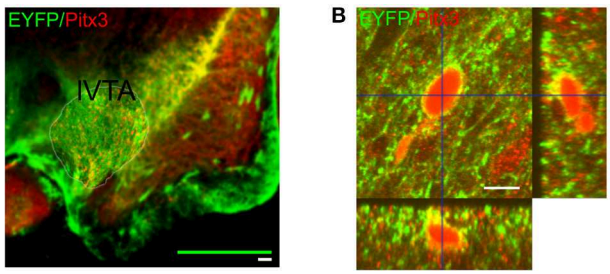

FIGURE 1 | Expression of the excitatory channelrhodpsin-2 channel in VTA dopaminergic neurons. **(A)** Representative wide-field micrograph showing the expression pattern of both EYFP (green) and Pitx3 (red) throughout the midbrain. One lateral side of a coronal midbrain slice is shown (lateral VTA). The midline is to the left side of the figure. The white scale bar corresponds to 100μm and the green bar indicates the outer size of the square MEA recording device that was optimally placed over the indicated VTA region (thin yellow contourline). The laser illumination spot fell well inside the MEA recording area as could be easily judged from the photo-electric artifact that it induced into the MEA leads. **(B)** Representative orthogonal projected confocal micrograph showing a EYFP+/Pitx3+ neuron located in the lateral VTA. Co-localization of Pitx3 expression (red) and ChR2-EYFP (green) is seen as a yellow coloring of the neuronal membrane. The orthogonal side views further illustrate the membrane expression of ChR2-EYFP. The scale bar corresponds to 10 μm.

the soma (red) combined with ChR2-EYFP expression (green) on the membrane, which lead to yellow coloring (**Figure 1B**). To characterize the co-expression, Z-stack images were acquired using a Zeiss LSM510 confocal laser scanning microscope fitted with a 63x objective. The orthogonal views enhanced the view of the co-localized expression (**Figure 1B**).

2.4. Slice Preparation and Electrophysiology

Mice were decapitated, the midbrain was dissected and kept at 4°C in artificial cerebral spinal fluid (ACSF) containing (in mM) NaCl 120, KCl 3.5, $CaCl_2$ 2.5, NaH_2PO_4 1.25, $MgSO_4$ 1.3, $NaHCO_3$ 25, D-glucose 10 which was bubbled with carbogen (95 % O_2; 5 % CO_2) to set pH at 7.4. Coronal slices were cut 300 μm thick from caudal to rostral using a VT1000S vibratome (Leica, Wetzlar, Germany). The fading of the substantia nigra during progressive slicing was a marker for the caudal-medial part of the VTA. One or two slices containing the caudal and medial part of the VTA were used for the experiments. Slices were incubated for 30 min at 32°C directly after slicing and were kept at room temperature until the start of the experiment. During recording in a MEA-1600 multi-electrode system (MultiChannel Systems, Reutlingen, Germany) the slice was kept at 32°C and continuously perfused with ACSF bubbled with carbogen. The VTA was identified in the mid-brain slice and positioned on top of the 3D MEA (Qwane Biosciences, Lausanne, Switzerland) containing 60 protruding electrodes (8*8 square layout) of 30 μm diameter and 100 μm spacing in order to record the spontaneous activity of multiple single-units (Olivier et al., 2002; van der Velden et al., 2017, 2019). All chemicals were obtained from Sigma-Aldrich (Zwijndrecht, the Netherlands).

2.5. Data Acquisition

The MEA recordings showed identifiable spikes, the extracellular representation of local action potentials, of 30–130 μV amplitude superimposed on a background noise of ∼15 μV (**Figure 2A**). The raw signal was high pass filtered at 225 Hz using a second order Butterworth filter and then sampled at 20 kHz (**Figure 2B**). Positive voltage peaks were detected, with a relatively low threshold in order not to miss spikes. Negative polarity spikes were rarely recorded and less suitable for analysis due to the negative polarity photoelectric artifact of the laser pulse stimulation on the MEA electrode leads (**Figure 2A**, green trace at the top indicates laser activation). Patches of signal containing the peak waveforms 3 ms around the peak were collected. K-means clustering was used to define the largest two principle components and the peak amplitudes of the waveforms. The auto-correlation and inter-spike-interval distribution of the peaks in the clusters were examined to identify clusters consisting of neuronal spikes. For electrodes that contained more than one neuron the most reliably recorded neuron was selected, most often based on peak amplitude (van der Velden et al., 2017).

2.6. Laser Stimulation Protocols

A 532 nm GL532T3 laser (Shanghai Laser & Optics Century Co., Shanghai, China), with an output of 320 mW, was coupled to a glass fiber, with 100 μm core radius and 0.22 NA, via a HPUC-23AF optic coupler (OZ optics, Ontario, Canada). Laser pulse stimulation to the superfused brain slice was applied from above. The fiber made contact with the perfusion liquid, but did not touch the slice. The distance between the MEA and the laser was ∼3 mm. The light intensity at the recording sites in this configuration had a lower bound of 5 mW/mm^2 (Deisseroth, Stanford). The calculation is based on light traveling through brain tissue only, but we also had travel distance through water. With this light intensity we do not expect to highly excite the neurons, but it is within the range reported in the literature (Madisen et al., 2012). The laser light induced a negative photoelectric artifact on the MEA electrodes so that visual inspection of the decay of this effect allowed us to ensure that the effective laser light spot was limited to the recording area of the MEA and covered the majority of the relevant electrode contacts. A Raspberry-Pi micro-controller (Raspberry-Pi foundation, UK) was programmed in Python to digitally control the timing of the laser stimuli (pulse duration always 10 ms) with a time resolution of around 20 μs. Two fundamentally differently timed stimulation protocols were used: The first one consisted of six segments of 240 s periods of laser stimulation (total 1440 s stimulation) interleaved with six segments of 240 s of no laser stimulation. In the on-period the laser either fired at a fixed 2 Hz rate or it produced a Poisson distributed stimulus train ($\exp(-\lambda t)$). The expectation density of the Poisson stimulation was 5 events/s and had its dominant power was distributed over the 1–5 Hz frequency range. The second protocol consisted of six 30 s periods of laser stimuli for each frequency (180 s per frequency). The laser stimulation frequency took fixed values between 0.5 and 5.5 Hz in 0.5 Hz steps. These protocols were applied in random order (repeated six times each).

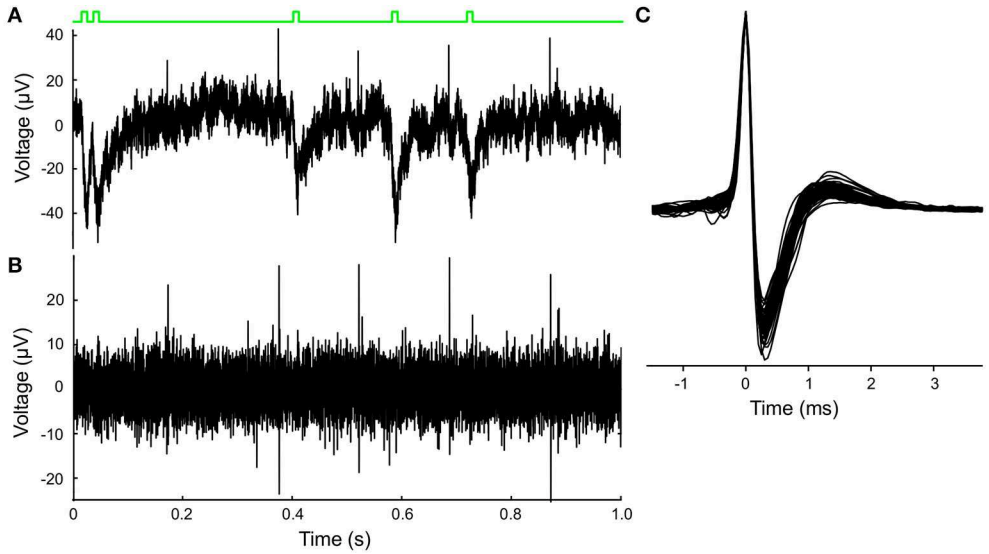

FIGURE 2 | Extracellular MEA recordings and optogenetic stimulation of VTA dopamine neuron activity. **(A)** Laser stimulation and unit recording illustrated in a 1s excerpt from a 720 s recording. Raw extracellular recording from one of the 60 MEA electrodes showing positive spikes that were clearly exceeding the noise level, and the negative photo-electrical artifacts initiated by the laser pulses. Schematic green line at the top depicts the timing of laser pulse stimuli (Poisson process). **(B)** Result of high-pass (>225 Hz) filtering of the signal in **(A)** which emphasizes the neuronal spikes and suppresses the photo-electrical artifact. **(C)** Superimposed spike-waveforms (cutout), which were averaged and normalized to the peak amplitude for each recorded VTA dopamine neuron ($n = 73$). Waveforms show the classic tri-phasic waveform of midbrain dopamine neurons.

2.7. Spike Train Analysis

The instantaneous firing rate (with 2.5 ms resolution) around the laser stimulus (at $t = 0$) was calculated from the Peri-Stimulus-Time-Histogram (PSTH) averaged over all recorded neurons (**Figure 3E**). Spike rate is expressed in spikes/s or Hz. The same graph was constructed for each neuron and we defined the efficacy of the laser stimulation for that neuron by the induced increase in firing rate above baseline level.

Rhythmic firing of VTA dopamine neurons can also be demonstrated as peaks in their auto-correlation function (**Figure 3C**). The oscillation frequency quantifies the preferred intrinsic rhythm of the neuron and was calculated as the mean interval between the side-lobes in the auto-correlation function, including the interval between the zero-lag peak and the first side lobe.

The irregularity of neuronal firing in the VTA was assessed using a measure of local variation, which quantifies the similarity between consecutive inter-spike-intervals (ISI), as in van der Velden et al. (2019). The local variation (LV, Shinomoto et al., 2005, 2009) of a spike train ranges from 0 (perfectly regular firing) to 1 (Poisson distributed firing) and above 1 for burst-like firing and is given by:

$$LV = \frac{3}{n-1}\sum_{i=1}^{n-1}\left(\frac{T_i - T_{i+1}}{T_i + T_{i+1}}\right) \quad (1)$$

where T_i is the i-th interval in the spike train that contains n spikes.

The Phase-Response-Curve (PRC) measures the laser-pulse-induced perturbation of a neurons' spike cycle. The phase of the spike cycle was normalized between zero and one, taking the oscillation frequency as a reference. A laser pulse could either delay the time to the next spike (negative phase shift) or advance it (positive phase shift). Averaging the PRC across neurons was done by binning the stimulus phase. The strength of the synchrony between a neuron's spike train and the laser pulses was quantified with the Pairwise-Phase-Consistency (PPC), as previously defined (Vinck et al., 2010) and validated (van der Velden et al., 2019). The PPC estimates the similarity of the relative phases of the point processes with respect to a chosen reference frequency and estimates the square of the classic Phase-Lock-Value (Lachaux et al., 1999). The PPC is an unbiased metric of phase-synchronization that scales with the square rather than the square root of the coherence and phase locking value (Vinck et al., 2010). To compute the PPC, spike and stimulation pulse trains were binned at 1 ms bins and a (Hanning) windowed Fourier Transform was computed on a series of time segments of the spike train. The length of the time segments was set to contain a fixed number of cycles of the reference frequency of interest (e.g., 5 cycles). The relative phase is defined as the complex argument of the classic spectral coherence (Lachaux et al., 1999; Vinck et al., 2010). From these relative phases the PPC was computed:

$$PPC = \frac{2}{N(N-1)}\sum_{j=1}^{N-1}\sum_{k=j+1}^{N} cos(\theta_j - \theta_k) \quad (2)$$

where there are N time segments, segment j has relative spike phase θ_j and segment k has relative spike phase θ_k, computed in respect to a chosen reference frequency. For the interleaved 2 Hz and the Poisson protocol we computed the PPC at selected frequencies independent of the laser pulse frequency. In the

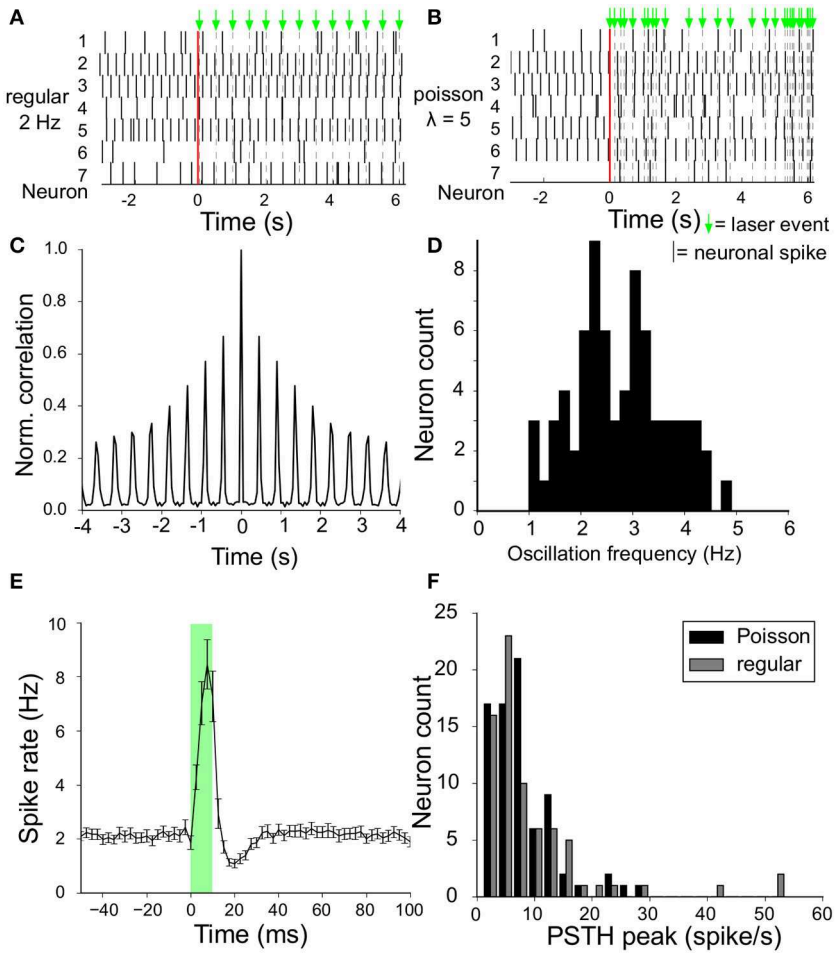

FIGURE 3 | Spontaneous activity and modulation by laser stimuli. **(A)** Raster plot of simultaneously recorded spike activity of a sample of seven VTA dopamine neurons in the same slice, before and during laser stimulation (onset at $t = 0$, vertical red line). Laser stimuli were delivered at a frequency of (2Hz), and are depicted by the green arrows above the top panel. **(B)** Same as **(A)** but now with laser pulses with a Poisson interval distribution (having an expectation value of 5 events/s). **(C)** Auto-correlation of the spontaneous activity of an example VTA dopamine neuron. The oscillation frequency was derived from the side-lobes in the auto-correlation. **(D)** Distribution of oscillation frequencies of all recorded VTA dopamine neurons ($n = 73$) from 7 experiments (=slices/animals). **(E)** Instantaneous firing rate around the laser stimulus, calculated from the Peri-Stimulus-Time-Histogram (2.5ms bin size) accumulated from all neurons ($n = 73$) for the 2Hz regular stimulation protocol. A sharp peak is followed by a trough in the spike rate during and after the laser stimulus given at ($t = 0$, transparent green) with stimulus duration of 10ms). **(F)** The maximum increase in instantaneous firing rate (peak minus baseline value) of the curve in **(E)** for all neurons is aggregated into a histogram to show the distribution of laser stimulation efficacy ($n = 73$). Results shown for both regular (gray) and Poisson stimulation (black).

stepped frequency protocol we computed the PPC only at the frequencies of the laser stimuli. The PPC was also used to assess the synchrony between spike trains simultaneously recorded from pairs of neurons. To disentangle the synchrony between neurons driven by common laser input and the part due to direct interactions between the neurons, we computed the partial coherence (Rosenberg et al., 1989, 1998; Brett et al., 2009) as defined in Sun et al. (2004), Brett et al. (2009). The spectral estimation for the coherence computation was performed using Welch's algorithm with overlapping windows of 5 s (Welch, 1967; Brett et al., 2009).

2.8. Experimental Design and Statistical Analysis

Independent experiments were performed on 7 slices during which in total 73 neurons were recorded, each slice was obtained from a different animal. Differences between experiments (=slices/animals) were tested with an ANOVA on two indicator parameters: the neuronal oscillation frequency and the laser induced spike probability. Detailed statistics were performed across all recorded neurons with $n = 73$ degrees of freedom. Exact p-values are given in the results section. The main statistical test used across neurons is the Spearman's rank correlation between the parameter of interest and the oscillation frequency of the neurons. Unless otherwise mentioned, all values reported in this study are given as mean and standard error of the mean (mean $\pm$ sem). In graphs the shaded area around a line indicates $\pm$ sem. Unless otherwise mentioned, direct comparison of two means was performed with Student's t-test, after checking for normality. Comparison of multiple groups was performed with ANOVA and *post-hoc* testing. The variance of the PPC for individual neurons was estimated with a jackknife method for each frequency.

The variable n indicates the number of observations, usually the number of neurons $n = 73$; $p < 0.05$ is assumed to reject the null hypothesis. Numpy/Scipy (Oliphant, 2007), Pandas (McKinney, 2010) and NIPY (Brett et al., 2009) packages for the Python programming language were used to perform the analysis. Data visualization was performed with Matplotlib (Hunter, 2007).

3. RESULTS

3.1. Spontaneous Activity Modulation

VTA dopamine neurons spontaneously fire rhythmic action potentials. A raw (**Figure 2A**) and filtered (**Figure 2B**) trace recorded from a single MEA electrode illustrate the activity of a neuron and the laser-pulse induced artifacts (photoelectric effect observed in the MEA leads). In the high pass filtered signal the neuronal spikes are easily detected above the (high-frequency) noise and the artifacts can be suppressed. Neuronal firing activity is also present in the absence of the laser stimulation, as reported previously (van der Velden et al., 2017, 2019). We calculated the amplitude-normalized averaged spike-waveform for each individual neuron and superimposed them in **Figure 2C** to demonstrate the highly characteristic classic tri-phasic waveform of midbrain dopamine neurons (Werkman et al., 2001; van der Velden et al., 2017, 2019). All neurons displayed the classic tri-phasic waveform of midbrain dopamine neurons (Werkman et al., 2001; van der Velden et al., 2017, 2019). The intrinsic oscillation frequency in the spike activity was computed from the time interval between peaks in the auto-correlations of the individual neurons as in van der Velden et al. (2019) (example in **Figure 3C**). **Figure 3D** shows the distribution of the intrinsic oscillation frequencies of all the neurons (mean ± standard deviation: 2.9 ± 0.9. The oscillation frequencies varied between 1 and 5 Hz, which matches well with the Local Field Potential power spectra of the VTA *in vivo* Fujisawa and Buzsáki (2011). This variation in oscillation frequency was observed between the neurons in the same slice and did not differ between experiments (=slices) (ANOVA, $n = 7$, $p = 0.56$). In previous work we have shown that the spontaneous activity of the lateral dopamine neurons is highly rhythmical and variation in regularity did not lead to separable neuronal populations (van der Velden et al., 2019). The raster plots in **Figures 3A,B** show 8 s spike trains (excerpts out of the six recording periods of 240 s) simultaneously recorded from seven VTA dopamine neurons in the same slice under regular 2 Hz laser stimulation (**Figure 3A**) and Poisson distributed laser stimulation ($\lambda = 5$, **Figure 3B**). The average PSTH of all neurons ($n = 73$) during 2 Hz regular stimulation (**Figure 3E**) showed a strong increase in spike rate after the onset of the laser pulse, immediately followed by a trough. Nonetheless, modulation by the laser stimulation was not equally effective for all neurons and showed considerable variation (**Figure 3F**). This shows that the optogenetic stimulation was, as expected from the expression patterns of ChR2 (**Figure 1**), strongly enhancing neuronal firing rates directly after the laser pulse onset.

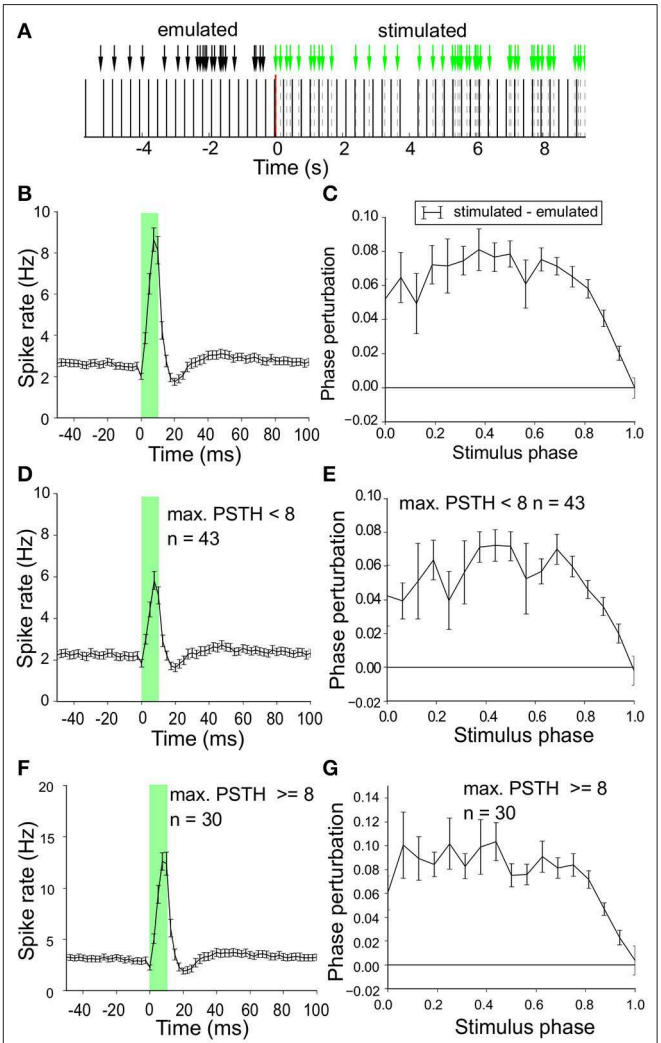

FIGURE 4 | Phase shift induced by laser stimulation. **(A)** Poisson stimulation protocol ($\lambda = 5$) during a stimulus off to on transition at $t = 0$. In the off condition we recorded baseline neuronal activity and for our analysis we emulated surrogate laser pulses (black arrows) with exactly the same time pattern but zero intensity. This allows us to perform identical assessment of the phase shift during laser on and off (control). **(B)** Mean instantaneous firing rate (spike/s) around the laser pulse (details as in **Figure 3E**) for all neurons in response to Poisson stimulation with a sharp peak during laser stimulation (green overlay) followed by a trough. **(C)** Phase-Response-Curve (PRC) (mean ± sem) for all neurons in response to Poisson stimulation. The PRC for the surrogate laser stimulation was subtracted to correct for potential bias.The x-axis represents the cycle duration, which is proportional to the neurons' intrinsic oscillation frequency. Spike phase advance (positive values) was largest around halfway the spikes (phase ∼ 0.5). **(D)** Mean instantaneous firing rate (spike/s) around the laser pulse (details as in **Figure 3E**) for neurons with a relatively low optogenetic efficacy (relative amplitude of peak above baseline < 8 spike/s). A sharp peak is followed by a trough, similar to the population result in **(B)**. **(E)** PRC [similar to **(C)**] of neurons with low optogenetic efficacy (relative amplitude of peak above baseline < 8 spike/s). The phase advance due to laser stimulation is largest halfway the spikes (phase ∼ 0.5). **(F)** Mean instantaneous firing rate (spike/s) around the laser pulse (details as in **Figure 3E**) for neurons with high optogenetic efficacy (relative amplitude of peak above baseline ≥ 8 spike/s). A sharp peak is seen followed by a trough, similar to the population result in **(B)**. **(G)** PRC [similar to **(C)**] of neurons with strong optogenetic efficacy (PSTH peak ≥ 8 spike/s).

3.2. Phase Response and Optogenetic Stimulation Strength

The spike time modulation by the optogenetic pulses was first studied during Poisson laser stimulation ($\lambda = 5$) with Phase-Response-Curves (PRCs). Phase was normalized to values between 0 and 1 using the neuron's intrinsic oscillation frequency. The phase-shift of the next spike induced by the laser stimulus was determined as a function of the phase of the stimulus (**Figure 4**). To control for possible random bias in our procedure we ran an identical PRC analysis on emulated laser pulses (i.e., on a surrogate laser pulse train with exactly the same timing, but no laser activation) during interleaved baseline recordings (black arrows in **Figure 4A**). The PRC for the surrogate train was then subtracted from the PRC of the actual laser train. The PRC, averaged over all neurons, showed an ∼8 % phase advance due to laser stimulation (**Figure 4C**). The maximal phase perturbation was observed halfway the spike cycle (phase 0.5), but this peak was not very pronounced. The mean instantaneous spike rate of this group of neurons around the stimulus is illustrated in **Figure 4B** and in essence hardly different from the one in **Figure 3E** for regular stimulation.

The shape of the PRC could well be related to the efficacy of the laser stimulation as measured in **Figure 4B**. Therefore we split the recorded neuron population based on their stimulation efficacy around the mean (8 spikes/s, **Figure 3F**) and obtained two groups (efficacy < 8 Hz, $n = 43$, efficacy ≥ 8 Hz, $n = 30$). **Figures 4D,F** show the instantaneous spike rate around the laser pulse averaged for each group of neurons. The shape of the curves is similar with a sharp peak during laser stimulation (green overlay) followed by a trough in spike rate. The PRCs, mean over all neurons in the group, are given in **Figures 4E,G** and both show phase-advancing induced by the laser reaching its maximal value at a stimulation phase halfway the spike cycle (around phase 0.5). The phase advance induced by the laser is slightly larger in the high efficacy group but the two PRCs did not show qualitative differences in the timing response to optogenetic pulse stimulation. These findings indicate that laser stimulation directly after a neuron has fired (start of the cycle) or directly before a neuron is expected to fire (end of the intrinsic oscillation cycle) have a relatively small effect on the timing of the next spike, but laser stimulation in the middle of the cycle can induce up to 10 % phase advance in the timing of the next spike. Similar results were obtained with regular 2 Hz stimulation (not shown), but Poisson stimulation sampled the phase relations between neuronal spikes and laser pulses more effectively, without the incidental risk of very strong phase locking.

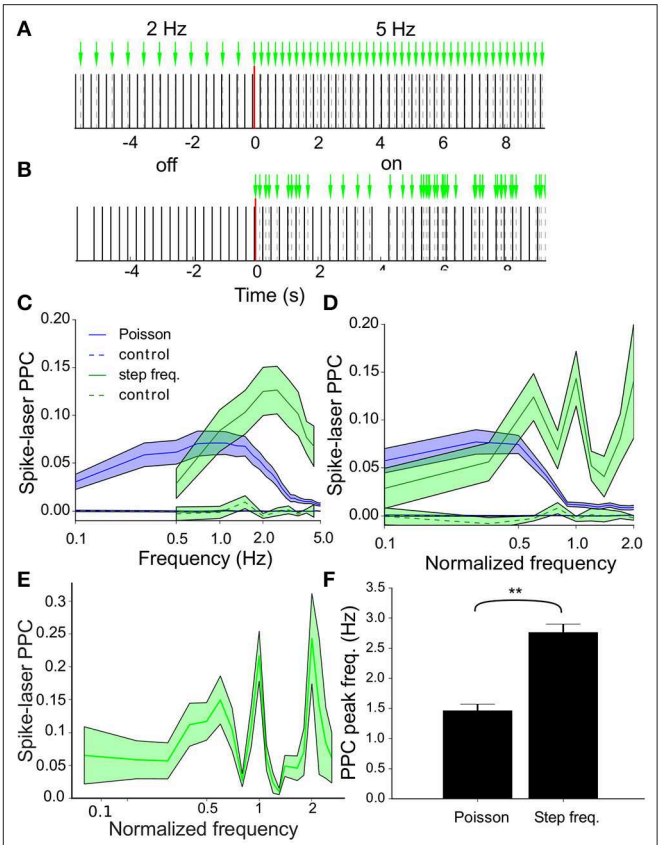

FIGURE 5 | Resonance of the VTA dopamine neuron population. **(A)** Step frequency protocol with frequency transition between two 30s recording periods transitioning from 2Hz to 5Hz stimulation. In addition to the actual protocol a control analysis was performed on an emulated (surrogate) stimulation protocol applied to baseline recordings of the neurons. **(B)** Laser stimulation as a Poisson process with 5 events/s ($\lambda = 5$) with laser-off and laser-on periods interleaved. **(C)** Pairwise phase consistency (PPC) spectra averaged across all neurons ($n = 73$) during stimulation [regular (green) and Poisson (blue), (mean ± sem)] measuring the phase synchronization between neuronal spikes and laser pulses. During regular stimulation a peak in resonance between neurons and laser was seen around 2Hz, whereas the highest resonance was observed around 1Hz during Poisson stimulation. The controls (dashed lines in the same color scheme) show that the correlation was not due to biasing factors such as incidental synchrony between laser and neuron. **(D)** PPC spectra averaged across all neurons ($n = 73$) during stimulation [regular (green) and Poisson (blue), (mean ± sem)]. Before averaging, the frequency values (x-axis) were normalized by dividing by the oscillation frequency for each individual neuron. During regular stimulation, we observed speaks in resonance around the oscillation frequency (corresponding to frequency value of 1.0) and its harmonics. Poisson stimulation led to resonance at relatively lower frequencies as compared to stimulation with regular pulse trains. **(E)** Same as **(D)** but now at a higher resolution, only for regular stimulation. **(F)** Bar plots of the peak spike-laser PPC frequency. The peak resonance-frequency is higher during step frequency (i.e., regular) stimulation as compared to Poisson stimulation. By contrast, the oscillation frequency (not shown) did not change between protocols.

3.3. Resonance to Pulse Stimulation Under Two Stimulus Regimes

The influence on the spike timing quantified by the PRCs in **Figure 4** is quite moderate, but these effects can have a cumulative effect on the synchrony between neuronal spikes and laser pulses. We used the Pairwise-Phase-Consistency (PPC, Vinck et al., 2010) to quantify this synchrony between spiking and laser (spike-laser PPC) and studied resonance at various frequencies. Two experimental laser stimulation protocols were used. The step frequency protocol sampled regular stimulation frequencies between 0.5 and 5.5 Hz with 0.5 Hz resolution (**Figure 5A**). Each frequency was applied six times for 30 s and for each neuron the spike-laser PPC was computed between neuron and laser at these frequencies. The Poisson protocol consistent

of randomly timed pulses during 6 periods of 240 s interleaved with as many periods of baseline recordings. As controls for both protocols, identical spike-laser PPC analyses were performed between emulated laser pulses (i.e., a surrogate laser stimulation train) and baseline firing activity.

Spike-laser PPC spectra were obtained for all recorded neurons during both protocols and averaged across neurons (Poisson: blue and regular: green, **Figure 5C**). Strong resonance between the neurons and laser was found for both protocols, but interestingly at different dominant frequencies, namely at around 2 Hz for the step frequency stimulation and around 1 Hz for Poisson stimulation. By contrast, the surrogate laser controls for both protocols showed PPC values very close or indistinguishable from zero at all frequencies. The relation between these spike-laser synchrony curves and the intrinsic oscillation frequency (mean ± sd: 2.9 ± 0.9, **Figure 3D**) was further studied. To this end, the spike-laser PPC spectrum frequency axis of each neuron was divided by its oscillation frequency, before averaging the PPC spectra across neurons (**Figures 5D,E**; D at lower resolution than E). For regular stimulation we observed peaks at the oscillation frequency (1.0) and peaks at integer fractions or multiples of the intrinsic oscillation frequency at 0.5 and 2.0 times (**Figures 5D,E**). Poisson stimulation shows a different pattern where resonance between laser and neuron decreases above the even sub-harmonic of the oscillation frequency (0.5) (**Figures 5D,E**). The oscillation frequencies did not differ between the baseline measurements of both protocols (delta osc. freq.: -0.013 ± 0.08, t-test: $p = 0.87$, $n = 73$), while the peak spike-laser PPC frequency increased from Poisson to step frequency (delta PPC peak freq.: 1.3 ± 0.18, t-test: $p < 0.001$, $n = 73$) (**Figure 5F**).

Example PPC spectra for three neurons are shown in **Figure 6**. The intrinsic oscillation frequencies are plotted as circles for the two protocols (Poisson blue and regular 2 Hz green). The PPC spectra for the controls (averaged over 100 randomized runs of surrogate laser runs) are shown as dashed lines in the same color scheme. The three neurons show different resonance behaviors; **Figure 6A** shows resonance to low frequencies during Poisson stimulation; **Figure 6B** shows high resonance at and above its oscillation frequency during regular stimulation and **Figure 6C** shows resonance for both protocols below its oscillation frequency. The oscillation frequency of the neuron shown in **Figure 6B** coincides with the laser stimulation frequency (1:1), which leads to strong phase locking. The other two neurons could contain other modi and beat tones (based on the difference between the laser pulse and oscillation frequency), which leads to these more elaborate spectra.

To further quantify the difference in resonance between the two protocols, we measured the strongest resonance frequency as the frequency with the highest spike-laser PPC value between neuronal spikes and laser pulses (peak PPC frequency) for each neuron. The oscillation frequency of the neuron correlated with the peak spike-laser PPC frequency (spearman $\rho = 0.5$, $p < 0.001$, **Figure 7A**) for the Poisson protocol, with maximum spike-laser PPC values around 50% of the intrinsic oscillation frequency. For the regular stimulation protocol, we computed correlations between intrinsic oscillation frequency for three different ranges separately, because neurons showed three resonant peaks at 0.5, 1, and 2 times the oscillation frequency (**Figure 5E**). These ranges were 0–0.75 the intrinsic oscillation frequency, 0.75–1.5 the intrinsic oscillation frequency, and 1.5–2.75 times the intrinsic oscillation frequency. We found that in these three ranges, there was a very strong correlation between the intrinsic oscillation frequency and the peak frequency of spike-LFP PPC. In the 0–0.75 range (times the oscillation frequency), we observed that neurons showed maximum spike-laser PPC values around 0.5 the oscillation frequency, and that neurons with a higher intrinsic oscillation frequency also had a higher peak spike-laser PPC frequency ($R = 0.64$, $p < 0.001$) (**Figure 7B**). In the 0.75–1.5 range (times the oscillation frequency), we observed that neurons

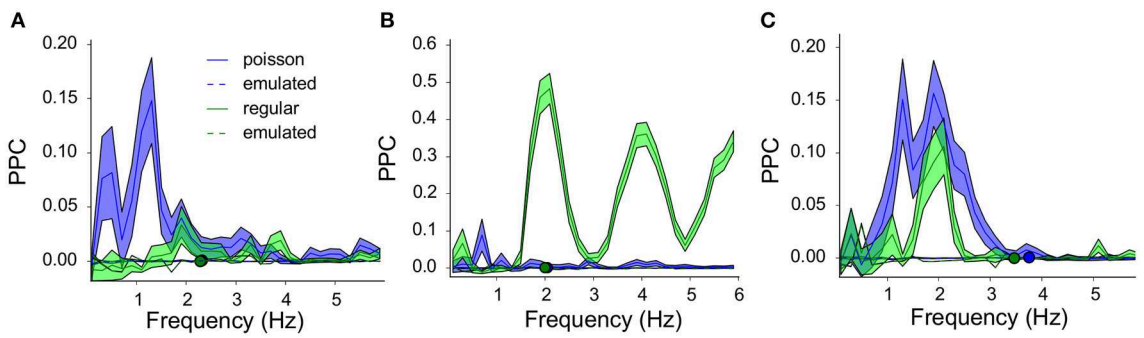

FIGURE 6 | Three example neurons and their resonance spectra illustrate non-linear, band-pass filtering by the VTA dopamine neurons: The response to all frequencies (Poisson stimulation) is not a simple sum of the response to specific frequencies (regular stimulation). **(A)** PPC spectra for a neuron oscillating at around 2.3Hz (green and blue overlapping circles), during regular (green, 2Hz) and Poisson stimulation [blue, $\lambda = 5$, (mean ± sem)]. The average of the 100 emulated (surrogate) controls showed PPC values very close to zero. Regular stimulation showed some structure around 2Hz, but synchrony between neuron and laser was stronger for the Poisson stimulation. A visible harmonic structure is related to the neurons' oscillation frequency (2.3Hz) and the laser frequency (2Hz). **(B)** As in **(A)** for a neuron oscillating at 2Hz. During regular stimulation (green) strong synchrony was seen at 2Hz and its even harmonics (4, 6Hz), which were absent during Poisson stimulation. **(C)** As in **(A)** for a neuron oscillating around 3.5–4Hz (green and blue circles). During both protocols resonance was seen at 2Hz).

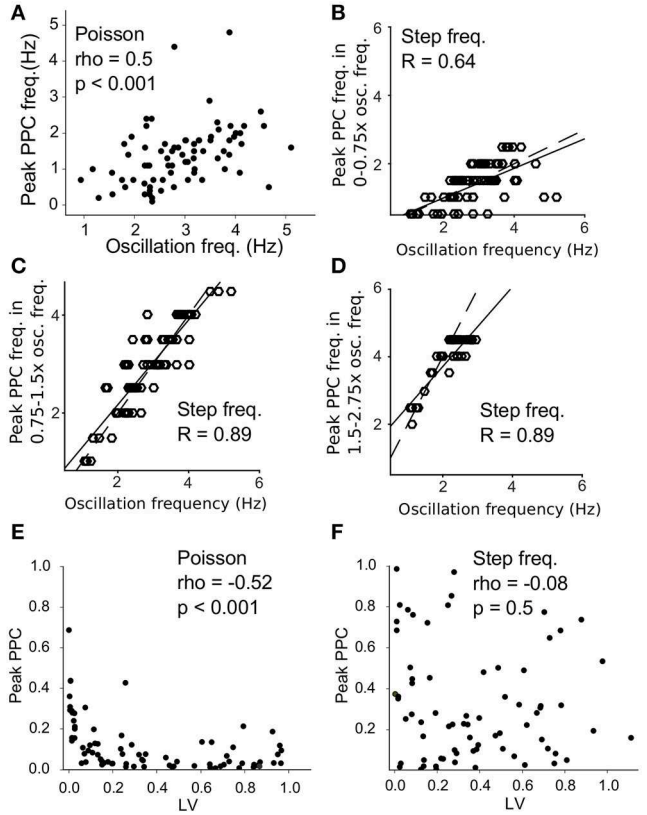

FIGURE 7 | Resonance at different ratios of the intrinsic oscillation frequency for Poisson and regular stimulation. **(A)** Intrinsic oscillation frequency (x-axis) vs. the peak frequency of spike-laser coupling (PPC) for Poisson stimulation. **(B)** Intrinsic oscillation frequency (x-axis) vs. the peak frequency of spike-laser coupling (PPC) for regular stimulation, computed in the frequency range of 0-0.75x the intrinsic oscillation frequency. The dashed line shows the 1:2 equality line, and the solid line the regression fit (R value computed with Pearson). **(C)** Same as **(B)** but now computed for the 0.75–1.5x the intrinsic oscillation frequency range. **(D)** Same as **(B,C)** but now computed for 1.5–2.75x the intrinsic oscillation frequency range. Note that for neurons with a high intrinsic oscillation frequency range, we likely underestimated the peak frequency of spike-laser coupling because of the range of laser stimulation. **(E)** Local Variance (LV) vs. peak PPC during Poisson stimulation. Highly regular firing neurons resonated stronger to the laser pulses as seen by a higher peak PPC. **(F)** LV vs. peak PPC during step frequency stimulation. High resonance to the laser stimulation (high peak PPC) was not correlated to the intrinsic firing regularity of the neurons.

showed maximum spike-laser PPC values around 1 times the oscillation frequency, with again a strong positive relationship between intrinsic oscillation frequency and peak spike-laser PPC frequency ($R = 0.89$, $p < 0.001$) (**Figure 7C**). In the 1.5–2.75 range (times the oscillation frequency), we observed that neurons showed maximum spike-laser PPC values around 2 times the oscillation frequency, with again a strong positive relationship between intrinsic oscillation frequency and peak spike-laser PPC frequency ($R = 0.89$, $p < 0.001$) (**Figure 7D**). Note that because we stimulated with regular laser pulses only up to 4.5 Hz, we likely underestimated the peak spike-laser PPC frequency for neurons with the highest intrinsic oscillation frequency, which causes a departure from the 1:2 trend for neurons with the highest oscillation frequency.

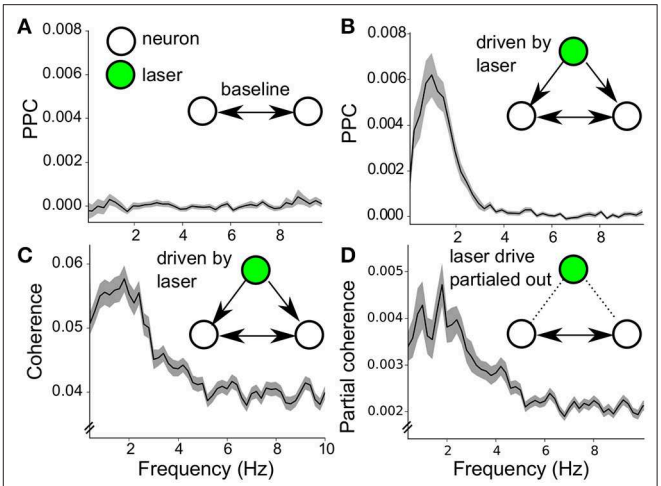

FIGURE 8 | Noise induced synchrony during Poisson stimulation. **(A)** Baseline PPC spectrum (stimulation off) of pairwise neuron-neuron interactions (neuron pairs = 347) within the 7 recorded populations (mean ± sem). **(B)** Same calculation as in **(A)** during laser stimulation (mean ± sem). The peak around 1Hz indicates synchrony among pairs of VTA dopamine neurons (neuron pairs = 347). **(C)** Coherence spectrum of neuron-neuron interactions during stimulation on (mean ± sem, neuron pairs = 347), confirming the results in **(B)**. **(D)** Partial coherence spectrum (mind the y-axis, mean ± sem, neuron pairs = 347). Around 10% of the coherence in **(C)** is not directly related to the common drive, as it is taken into account during the coherence computation.

We further wondered how the electrophysiological properties of neurons during baseline (i.e., outside laser stimulation periods) correlated with entrainment by the laser. We found that during Poisson stimulation, the most highly regular firing neurons (low LV, as measured during baseline) showed the strongest resonance (high peak PPC, **Figure 7E**). This relation was not seen for stepped frequency stimulation; high PPC values did not correlate with high regularity (**Figure 7F**). To conclude, we found that VTA dopamine neurons show resonant filtering of extrinsic inputs around 0.5, 1, and 2x their intrinsic oscillation frequency when these extrinsic (laser) inputs where regular, but responded more strongly to a 0.5x sub harmonic of their oscillation frequency when the extrinsic (laser) inputs were noisy (wide-band; Poisson). Furthermore, firing regularity determined the strength of the resonance with noisy extrinsic inputs, but not for regular inputs. These findings (difference Poisson vs. regular, and multiple resonant modes) suggest that the filtering exhibited by VTA dopamine neurons is not merely linear, but exhibits a non-linearity.

3.4. Neuron-Neuron Interactions Induced by Poisson Stimulation

The dopamine neurons in the VTA population were able to respond collectively to Poisson stimulation, due to the latter's broadband frequency content (**Figure 5C**). Can such common noise stimulation induce a structured network state? To this end, the PPC was computed between all pairs of neurons within the population during stimulation off as well as stimulation on periods. The mean PPC spectrum during stimulation off (**Figure 8A**) showed no prominent PPC peak in the baseline pairwise neuron interactions, the values were never different

from zero [Note that during baseline, only about 10–20% of neurons showed significant spike-spike synchronization, and only when the oscillation frequencies of the neurons match, van der Velden et al. (2019)]. During stimulation, the PPC spectrum has a large peak around 1 Hz (**Figure 8B**), which suggests neuron-neuron synchrony driven by the Poisson stimulation at selective frequencies.

The results on noise-induced-synchrony can be further refined because we have full knowledge of the timing of the Poisson laser pulses and thus can factor out its driving influence on the neuron-neuron synchrony. The component of the neuronal interactions due to common drive by the laser pulses was separated by computing the coherence and the partial coherence (see methods) (Rosenberg et al., 1989, 1998; Sun et al., 2004; Brett et al., 2009). The coherence (**Figure 8C**) corroborated the results of the PPC for neuron-neuron synchrony (**Figure 8B**). The partial coherence showed that about one-tenth of the coherence between pairs of neurons was not a direct result of the common drive by the laser pulses (**Figure 8D**, note the different scale of the y-axis). Thus the wide-band Poisson regime was capable of inducing a synchronous state across the VTA network, mainly due to common drive of the neurons (noise-induced-synchrony Galán et al., 2006; Ermentrout et al., 2008). Interestingly, part of the neuron-neuron interactions was not directly attributable to common drive, which points to emergent network properties during noise stimulation.

4. DISCUSSION

VTA dopamine neuron exhibited spontaneous rhythmic firing with oscillation frequencies between 1 and 5 Hz in agreement with our earlier reports (van der Velden et al., 2017, 2019). Optogenetic stimulation using laser pulses affected the timing of the spikes. This spike timing modulation induced synchrony between the spikes and the laser pulses, which was used to study input resonance properties of the dopamine neurons and their network. VTA dopamine neurons resonated to the laser pulses in both the regular and the Poisson stimulation regime. The resonance spectra showed varying strength of phase locking between neuronal spikes and laser pulses related to the intrinsic oscillation frequency of the neurons. The VTA dopamine neurons had a different resonance response to Poisson than to regular input. The dominant resonance frequency to regular stimulation was almost 100 % higher than the one to Poisson stimulation. During regular stimulation the resonance frequencies were similar to the intrinsic oscillation frequency, as well as 0.5 and 2x this intrinsic oscillation frequency. For Poisson stimulation the peak resonance (at 1.5 Hz) scaled with the sub harmonic of the oscillation frequency (at 2.9 Hz). Additionally, for Poisson stimulation, the strength of the resonance also correlated with the intrinsic regularity of the auto oscillator. The separation of the resonance behavior into two frequency domains for regular vs. noise stimulation could mediate selective information exchange with different brain areas. Information could be processed independently between regular direct drive and noisy background signals, representing respectively local and long range communications. In all, this dynamic resonance behavior means that lateral VTA dopamine neurons behave as non-linear filters; at the same input frequencies they respond differently to narrow (regular) and broad-band (Poisson) input, through altered filter characteristics based on spike timing synchrony.

The broad-band nature of Poisson stimulation allowed us to drive simultaneously recorded neurons effectively, as each neuron can resonate to its preferred frequency. This common drive induced neuron-neuron synchrony (pairwise PPC), which was below the intrinsic oscillation frequency. Such noise-induced-synchrony has rarely been recorded in a neuronal system and is a recent addition to our understanding of the brains' function in the presence of noise (Galán et al., 2006; Ermentrout et al., 2008; Hata et al., 2010). Our study suggests that the lateral VTA with its weakly-coupled intrinsic oscillators and resonance characteristics is a suitable system for studying emergent network properties beyond the noise-induced-synchrony that is indicated by our results.

Fujisawa and Buzsáki (2011) hypothesized that the VTA is the pacemaker of low frequency local field rhythm (their data shows it between 1 and 5 Hz with a peak around 3 Hz), which entrains the prefrontal cortex and the hippocampus (Lisman and Grace, 2005; Fujisawa and Buzsáki, 2011; Kim et al., 2012). In our study the intrinsic oscillation frequencies of VTA dopamine neurons were between 1 and 5 Hz *in vitro*. The intrinsic oscillation frequency can be increased with extracellular glutamate (van der Velden et al., 2019). Under regular stimulation dopamine neurons resonate most strongly around 2.8 Hz, but also showed resonance at 2x the intrinsic oscillation frequency. Poisson stimulation induced noise-induced-synchrony among VTA dopamine neurons at 1.5 Hz. These results provide mechanisms underlying a pacemaker role for the VTA. Additionally, our data show that the lateral VTA has self organizing properties, as it exhibits neuron-neuron synchrony at selective frequencies while being driven by broad-band noise. The selective frequency during noisy stimulation is a 0.5x sub harmonic of its mean oscillation frequency. VTA dopamine neurons that fire with higher intrinsic regularity are able to resonate stronger at this sub harmonic. The VTA's self-organizing frequency-selective output during noise-like input thus represents a mechanism for low-frequency pacemaker activity.

We did not study the cellular mechanisms behind the sub oscillation frequency resonance to Poisson pulses, which likely involves sub-threshold oscillations (Lampl and Yarom, 1997) as described in several models of the dopamine neuron (Wilson and Callaway, 2000; Medvedev et al., 2003; Kuznetsov et al., 2006). We hypothesize that the observed resonance can be understood as the stochastic drive of a non-linear oscillator, which has a preferred stimulus phase within the spike cycle (i.e., halfway the spike cycle) (Tiesinga, 2002; Hata et al., 2010). Following this, Poisson stimuli are less likely to fall within the right time window each spike cycle, than every other cycle, every third or fourth cycle and thus will prefer to push the neuron with sub-harmonics of its intrinsic oscillation frequency.

Our findings have broad functional implications. The VTA has long been implicated in stimulus-reward coupling, with reward prediction based on an internal time code. However, the neural substrate of such a time code in the range of seconds is as yet unknown (Schultz, 1997; Lak et al., 2016). VTA dopamine neurons' phasic activity is proposed to act as a reward prediction signal, which is then modulated by GABAergic VTA neurons that receive external input (Cohen et al., 2012). Moghaddam et al. (2017) indeed showed that VTA network properties are critical to understanding the encoding of information about serial behavior. An underlying timing structure of the reward prediction is needed for such encoding and behavior. Our results suggest that the lateral VTA acts like a filter bank, whose neuron sub-populations resonate to different frequencies in the laser driven input. The neurons in these sub-populations are organized by having a similar or harmonically related intrinsic oscillation frequency. The output of these oscillating sub-populations encode the timing information in the input and could be used as a time code by down-stream target areas of the lateral VTA and as a reward prediction error code in general.

Clinically the VTA has been discovered as a target for deep brain stimulation (DBS) in the treatment of pharmacoresistant cluster headache (Akram et al., 2016, 2017). The frequency used for electrical stimulation of the VTA in this form of DBS therapy is 180 Hz, but no optimization has as yet been undertaken. This frequency is so high in comparison to the intrinsic VTA firing rates that if might best be considered as noisy stimulation. DBS is also a therapy of last resort in pharmacoresistant epilepsy, mostly at frequencies around 120 Hz, but here it could be demonstrated in an animal model that Poisson distributed stimulation was more effective than fixed frequency stimulation with the same energy (Wyckhuys et al., 2010).

Our research indicates that the VTA network is a non-linear filter bank when responding to external input, which can decode timing information related to stimulus-reward processing. These network states are self-organizing as they impose structure on noisy input and could underlie a pacemaker role of the VTA, linked to entrainment of the hippocampus and the prefrontal cortex during cognitive tasks.

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 Ethics Committee of the Swammerdam Institute for life sciences of the University of Amsterdam.

AUTHOR CONTRIBUTIONS

LV, MV, and WW designed the study, contributed to the data analysis, and wrote the manuscript. LV performed the experiments. LV and MV performed the data analysis.

FUNDING

This work was supported by grant 114000091 to WW from the Platform alternatief voor dierproeven of the Dutch organization for fundamental research.

ACKNOWLEDGMENTS

The authors thank Hans Lodder for early assistance with the optical stimulation. This paper was based on chapter 4 of the Ph.D. thesis of LV (van der Velden, 2018).

REFERENCES

Akram, H., Miller, S., Lagrata, S., Hariz, M., Ashburner, J., Behrens, T., et al. (2017). Optimal deep brain stimulation site and target connectivity for chronic cluster headache. *Neurology* 89, 2083–2091. doi: 10.1212/WNL.0000000000004646

Akram, H., Miller, S., Lagrata, S., Hyam, J., Jahanshahi, M., Hariz, M., et al. (2016). Ventral tegmental area deep brain stimulation for refractory chronic cluster headache. *Neurology* 86, 1676–1682. doi: 10.1212/WNL.0000000000002632

Battaglia, F. P., and McNaughton, B. L. (2011). Polyrhythms of the brain. *Neuron* 72, 6–8. doi: 10.1016/j.neuron.2011.09.019

Bayer, H. M., Lau, B., and Glimcher, P. W. (2007). Statistics of midbrain dopamine neuron spike trains in the awake primate. *J. Neurophysiol.* 98, 1428–1439. doi: 10.1152/jn.01140.2006

Bayer, V. E., and Pickel, V. M. (1990). Ultrastructural localization of tyrosine hydroxylase in the rat ventral tegmental area: relationship between immunolabeling density and neuronal associations. *J. Neurosci.* 10, 2996–3013.

Berretta, N., Bernardi, G., and Mercuri, N. B. (2010). Firing properties and functional connectivity of substantia nigra pars compacta neurones recorded with a multi-electrode array *in vitro*. *J. Physiol.* 588(Pt 10), 1719–1735. doi: 10.1113/jphysiol.2010.189415

Brett, M., Taylor, J., Burns, C., Millman, K. J., Perez, F., Roche, A., et al. (2009). NIPY: an open library and development framework for FMRI data analysis. *NeuroImage* 47:S196. doi: 10.1016/S1053-8119(09)72223-2

Buzsaki, G. (2009). *Rhythms of the Brain*. Oxford: Oxford University Press.

Cardin, J. A. (2012). Dissecting local circuits *in vivo*: Integrated optogenetic and electrophysiology approaches for exploring inhibitory regulation of cortical activity. *J. Physiol. Paris* 106, 104–111. doi: 10.1016/j.jphysparis.2011.09.005

Cohen, J. Y., Haesler, S., Vong, L., Lowell, B. B., and Uchida, N. (2012). Neuron-type-specific signals for reward and punishment in the ventral tegmental area. *Nature* 482, 85–88. doi: 10.1038/nature10754

Cragg, S. J., Nicholson, C., Kume-Kick, J., Tao, L., and Rice, M. E. (2001). Dopamine-mediated volume transmission in midbrain is regulated by distinct extracellular geometry and uptake. *J. Neurophysiol.* 85, 1761–1771. doi: 10.1152/jn.2001.85.4.1761

Deisseroth, K. (2010). Optogenetics. *Nat. Methods* 8, 26–29. doi: 10.1038/nmeth.f.324

Drion, G., Massotte, L., Sepulchre, R., and Seutin, V. (2011). How modeling can reconcile apparently discrepant experimental results: the case of pacemaking in dopaminergic neurons. *PLoS omput. Biol.* 7:e1002050. doi: 10.1371/journal.pcbi.1002050

Ermentrout, G. B., Galán, R. F., and Urban, N. N. (2008). Reliability, synchrony and noise. *Trends Neurosci.* 31, 428–434. doi: 10.1016/j.tins.2008.06.002

Fenno, L., Yizhar, O., and Deisseroth, K. (2011). The development and application of optogenetics. *Ann. Rev. Neurosci.* 34, 389–412. doi: 10.1146/annurev-neuro-061010-113817

Fujisawa, S., and Buzsáki, G. (2011). A 4 Hz oscillation adaptively synchronizes prefrontal, VTA, and hippocampal activities. *Neuron* 72, 153–165. doi: 10.1016/j.neuron.2011.08.018

Galán, R. F., Fourcaud-Trocmé, N., Ermentrout, G. B., and Urban, N. N. (2006). Correlation-induced synchronization of oscillations in olfactory bulb neurons. *J. Neurosci.* 26, 3646–3655. doi: 10.1523/JNEUROSCI.4605-05.2006

Hata, S., Shimokawa, T., Arai, K., and Nakao, H. (2010). Synchronization of uncoupled oscillators by common gamma impulses: from phase locking to noise-induced synchronization. *Phys. Rev. E Statist. Nonlinear Soft Matt. Phys.* 82:036206. doi: 10.1103/PhysRevE.82.036206

Hunter, J. D. (2007). Matplotlib: a 2D graphics environment. *Comput. Sci. Eng.* 9, 90–95. doi: 10.1109/MCSE.2007.55

Khaliq, Z. M., and Bean, B. P. (2010). Pacemaking in dopaminergic ventral tegmental area neurons: depolarizing drive from background and voltage-dependent sodium conductances. *J. Neurosci.* 30, 7401–7413. doi: 10.1523/JNEUROSCI.0143-10.2010

Kim, Y., Wood, J., and Moghaddam, B. (2012). Coordinated activity of ventral tegmental neurons adapts to appetitive and aversive learning. *PLoS ONE* 7:e29766. doi: 10.1371/journal.pone.0029766

Koyama, S., Kanemitsu, Y., and Weight, F. F. (2005). Spontaneous activity and properties of two types of principal neurons from the ventral tegmental area of rat. *J. Neurophysiol.* 93, 3282–3293. doi: 10.1152/jn.00776.2004

Kuznetsov, A. S., Kopell, N. J., and Wilson, C. J. (2006). Transient high-frequency firing in a coupled-oscillator model of the mesencephalic dopaminergic neuron. *J. Neurophysiol.* 95, 932–947. doi: 10.1152/jn.00691.2004

Lachaux, J. P., Rodriguez, E., Martinerie, J., and Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* 8, 194–208.

Lak, A., Stauffer, W. R., and Schultz, W. (2016). Dopamine neurons learn relative chosen value from probabilistic rewards. *eLife* 5:e18044. doi: 10.7554/eLife.18044

Lammel, S., Hetzel, A., Häckel, O., Jones, I., Liss, B., and Roeper, J. (2008). Unique properties of mesoprefrontal neurons within a dual mesocorticolimbic dopamine system. *Neuron* 57, 760–773. doi: 10.1016/j.neuron.2008.01.022

Lampl, I., and Yarom, Y. (1997). Subthreshold oscillations and resonant behavior: two manifestations of the same mechanism. *Neuroscience* 78, 325–341.

Lisman, J. E., and Grace, A. A. (2005). The hippocampal-VTA loop: controlling the entry of information into long-term memory. *Neuron* 46, 703–713. doi: 10.1016/j.neuron.2005.05.002

Luk, K. C., Rymar, V. V., van den Munckhof, P., Nicolau, S., Steriade, C., Bifsha, P., et al. (2013). The transcription factor Pitx3 is expressed selectively in midbrain dopaminergic neurons susceptible to neurodegenerative stress. *J. Neurochemist.* 125, 932–943. doi: 10.1111/jnc.12160

Madisen, L., Mao, T., Koch, H., Zhuo, J.-M., Berenyi, A., Fujisawa, S., et al. (2012). A toolbox of Cre-dependent optogenetic transgenic mice for light-induced activation and silencing. *Nat. Neurosci.* 15, 793–802. doi: 10.1038/nn.3078

McKinney, W. (2010). "Data structures for statistical computing in Python," in *Proceedings of the 9th Python in Science Conference (SciPy 2010)*, eds S. van der Walt and J. Millman (Austin, TX), 51. Available online at: https://conference.scipy.org/proceedings/scipy2010/

Medvedev, G. S., Wilson, C. J., Callaway, J. C., and Kopell, N. (2003). Dendritic synchrony and transient dynamics in a coupled oscillator model of the dopaminergic neuron. *J. Comput. Neurosci.* 15, 53–69. doi: 10.1023/A:1024422802673

Moghaddam, B., Koerner, F. S., Kass, R. E., Wood, J., and Simon, N. W. (2017). Networks of VTA neurons encode real-time information about uncertain numbers of actions executed to earn a reward. *Front. Behav. Neurosci.* 11, 1–16. doi: 10.3389/fnbeh.2017.00140

Oliphant, T. E. (2007). Python for Scientific Computing. *Comput. Sci. Eng.* 9, 10–20. doi: 10.1109/MCSE.2007.58

Olivier, M., Fejtl, M., Raggenbass, M., Bertrand, D., Renaud, P., Heuschkel, M. O., et al. (2002). A three-dimensional multi-electrode array for multi-site stimulation and recording in acute brain slices. *J. Neurosci. Methods* 114, 135–148. doi: 10.1016/S0165-0270(01)00514-3

Omelchenko, N., and Sesack, S. R. (2009). Ultrastructural analysis of local collaterals of rat ventral tegmental area neurons: GABA phenotype and synapses onto dopamine and GABA cells. *Synapse* 63, 895–906. doi: 10.1002/syn.20668

Rosenberg, J., a.M. Amjad, Breeze, P., Brillinger, D., and Halliday, D. (1989). The Fourier approach to the identification of functional coupling between neuronal spike trains. *Prog. Biophys. Mol. Biol.* 53, 1–31. doi: 10.1016/0079-6107(89)90004-7

Rosenberg, J. R., Halliday, D. M., Breeze, P., and Conway, B. a. (1998). Identification of patterns of neuronal connectivity - partial spectra, partial coherence, and neuronal interactions. *J. Neurosci. Methods* 83, 57–72.

Schultz, W. (1997). A neural substrate of prediction and reward. *Science* 275, 1593–1599.

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

Shinomoto, S., Miura, K., and Koyama, S. (2005). A measure of local variation of inter-spike intervals. *Bio Systems* 79, 67–72. doi: 10.1016/j.biosystems.2004.09.023

Smidt, M. P., Asbreuk, C. H., Cox, J. J., Chen, H., Johnson, R. L., and Burbach, J. P. (2000). A second independent pathway for development of mesencephalic dopaminergic neurons requires Lmx1b. *Nat. Neurosci.* 3, 337–341. doi: 10.1038/73902

Smidt, M. P., Oerthel, L. V., Hoekstra, E. J., Schellevis, R. D., and Hoekman, M. F. M. (2012). Spatial and temporal lineage analysis of a Pitx3-Driven cre-recombinase knock-in mouse model. *PLoS ONE* 7:e42641. doi: 10.1371/journal.pone.0042641

Sun, F. T., Miller, L. M., and D'Esposito, M. (2004). Measuring interregional functional connectivity using coherence and partial coherence analyses of fMRI data. *NeuroImage* 21, 647–658. doi: 10.1016/j.neuroimage.2003.09.056

Tiesinga, P. H. E. (2002). Precision and reliability of periodically and quasiperiodically driven integrate-and-fire neurons. *Phys. Rev. E Statist. Nonlinear. Soft Matter Phys.* 65, 1–14. doi: 10.1103/PhysRevE.65.041913

van der Velden, L., Vinck, M., Werkman, T. R., and Wadman, W. J. (2017). Tuning of neuronal interactions in the lateral ventral tegmental area by dopamine sensitivity. *Neuroscience* 366, 62–69. doi: 10.1016/j.neuroscience.2017.10.009

van der Velden, L., Vinck, M. A., Werkman, T. R., and Wadman, W. J. (2019). Modulation of functional connectivity between dopamine neurons of the rat ventral tegmental area *in vitro*. *Front. Int. Neurosci.* 13:20. doi: 10.3389/fnint.2019.00020

van der Velden, L. J. J. (2018). *Drumming with dopaminic neurons: resonance and synchronization in the Ventral Tegmental Area* (Doctoral thesis). University of Amsterdam, Amsterdam, Netherlands.

Vinck, M., van Wingerden, M., Womelsdorf, T., Fries, P., and Pennartz, C. M. A. (2010). The pairwise phase consistency: a bias-free measure of rhythmic neuronal synchronization. *NeuroImage* 51, 112–122. doi: 10.1016/j.neuroimage.2010.01.073

Welch, P. D. (1967). The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms. *IEEE Trans. Audio Electroacoust.* 15, 70–73.

Werkman, T. R., Kruse, C. G., Nievelstein, H., Long, S. K., and Wadman, W. J. (2001). *In vitro* modulation of the firing rate of dopamine neurons in the rat substantia nigra pars compacta and the ventral tegmental area by antipsychotic drugs. *Neuropharmacology* 40, 927–936. doi: 10.1016/S0028-3908(01)00015-6

Wilson, C. J., and Callaway, J. C. (2000). Coupled Oscillator Model of the Dopaminergic Neuron of the Substantia Nigra. *J. Neurophysiol.* 83, 3084–3100. doi: 10.1152/jn.2000.83.5.3084

Wyckhuys, T., Boon, P., Raedt, R., Van Nieuwenhuyse, B., Vonck, K., and Wadman, W. J. (2010). Suppression of hippocampal epileptic seizures in the kainate rat by Poisson distributed stimulation. *Epilepsia* 51, 2297–2304. doi: 10.1111/j.1528-1167.2010.02750.x

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.

Copyright © 2020 van der Velden, Vinck and Wadman. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Neuroscience

ORIGINAL RESEARCH
published: 03 April 2020
doi: 10.3389/fnins.2020.00230

Decoding Adaptive Visuomotor Behavior Mediated by Non-linear Phase Coupling in Macaque Area MT

Mohammad Bagher Khamechian[1] *and Mohammad Reza Daliri*[1,2]*

[1] *Neuroscience and Neuroengineering Research Lab, Biomedical Engineering Department, School of Electrical Engineering, Iran University of Science & Technology, Tehran, Iran,* [2] *Cognitive Neurobiology Laboratory, School of Cognitive Sciences (SCS), Institute for Research in Fundamental Sciences (IPM), Tehran, Iran*

The idea that a flexible behavior relies on synchronous neural activity within intra- and inter-associated cortical areas has been a matter of intense research in human and animal neuroscience. The neurophysiological mechanisms underlying this behavioral correlate of the synchronous activity are still unknown. It has been suggested that the strength of neural synchrony at the level of population is an important neural code to guide an efficient transformation of the sensory input into the behavioral action. In this study, we have examined the non-linear synchronization between neural ensembles in area MT of the macaque visual cortex by employing a non-linear cross-frequency coupling technique, namely bicoherence. We trained a macaque monkey to detect a brief change in the cued stimulus during a visuomotor detection task. The results show that the non-linear phase synchronization in the high-gamma frequency band (100–250 Hz) predicts the animal's reaction time. The strength of non-linear phase synchronization is selective to the target stimulus location. In addition, the non-linearity characteristics of neural synchronization are selectively modulated when the monkey covertly attends to the stimulus inside the neuron's receptive field. This additional evidence indicates that non-linear neuronal synchronization may be affected by a cognitive function like spatial attention. Our neural and behavioral observations reflect that two crucial processes may be involved in processing of visuomotor information in area MT: (I) a non-linear cortical process (measured by the bicoherence) and (II) a linear process (measured by the spectral power).

Keywords: bicoherence, quadratic phase coupling, non-linear phase synchronization, visual area MT, spatial attention

OPEN ACCESS

Edited by:
Daya Shankar Gupta, Camden County College, United States

Reviewed by:
Geoffrey Mohon Ghose, University of Minnesota Twin Cities, United States
Takahiro Osada, Juntendo University, Japan
Hamid Reza Marateb, Universitat Politecnica de Catalunya, Spain

***Correspondence:**
Mohammad Reza Daliri
daliri@iust.ac.ir

Specialty section:
This article was submitted to Perception Science, a section of the journal Frontiers in Neuroscience

Received: *12 September 2019*
Accepted: *02 March 2020*
Published: *03 April 2020*

Citation:
Khamechian MB and Daliri MR (2020) Decoding Adaptive Visuomotor Behavior Mediated by Non-linear Phase Coupling in Macaque Area MT. Front. Neurosci. 14:230. doi: 10.3389/fnins.2020.00230

INTRODUCTION

Neural oscillations are frequently observed in cortical activities. Notably, it has been widely asserted that neural oscillations are involved in many cortical computations, including sensory coding (Siegel et al., 2007; Belitski et al., 2008; Schroeder and Lakatos, 2009) and information transmission (Hipp et al., 2011; van Kerkoerle et al., 2014; Rohenkohl et al., 2018). Brain networks can communicate through frequency-specific oscillations. These oscillatory activities can play a functional role in brain networks to flexibly integrate, process, and transmit neural information among cortical circuitries (Moore and Armstrong, 2003; Buschman and Miller, 2007;

Saalmann et al., 2007; Siegel et al., 2008; Hipp et al., 2011). However, recent studies have suggested that brain oscillations could interdependently interact, forming so-called cross-frequency coupling (CFC) (Buzsáki, 2006; Jensen and Colgin, 2007). This form of interactive computation has been observed in several brain areas of different species (Canolty et al., 2010; Igarashi et al., 2014; Esghaei et al., 2015). The CFC has an important role in many cortical functions, including sensory processing (Saleh et al., 2010), learning (Tort et al., 2009; Igarashi et al., 2014), memory (Axmacher et al., 2010), and attention (Esghaei et al., 2015; Spyropoulos et al., 2018). It is believed that CFC can functionally facilitate information coordination between neurons, simultaneously in time and space (Aru et al., 2015). Furthermore, recent studies in human and non-human primates have shown that CFC may serve as a potential physiological mechanism underlying intra-areal communication in the brain (Darvas et al., 2009; Canolty and Knight, 2010; Holz et al., 2010; Fiebelkorn et al., 2018). For example, a study on the visuospatial working memory in human indicated that CFC between oscillatory phases of theta (4–8 Hz) and gamma (50–70Hz) activities can regulate an effective communication between occipital and parietal brain regions (Holz et al., 2010). Another investigation in macaque monkey suggested that coupling between the phase of theta oscillations (3–8 Hz) and the power of high frequencies (9–45 Hz) during spatial attention potentially facilitates an interregional communication between the frontal eye field (FEF) area, lateral intraparietal area (LIP), and visual cortex (Fiebelkorn et al., 2018).

Contemporary investigations into visual areas have shown that oscillatory components of local field potential (LFP) (Liu and Newsome, 2006; Womelsdorf et al., 2006; Smith et al., 2015; Khamechian et al., 2019) and neural spiking activity (Liu and Newsome, 2005; Smith et al., 2015; Parto Dezfouli et al., 2018) could provide useful information about how neural activities are linked to visuomotor behavior. These studies have reported a trial-by-trial correlation between the power of beta (10–30 Hz) (Smith et al., 2015), gamma, and high-gamma (50–200 Hz) (Liu and Newsome, 2006) LFPs and behavioral output. Moreover, they have shown that the strength of gamma (Womelsdorf et al., 2006) and high-gamma synchronization (Khamechian et al., 2019) between sensory neurons in the dorsal and ventral visual pathway, respectively, predict the speed of behavioral responses. Despite these promising observations on neural-behavior correlation in the sensory visual areas, the contributive role of non-linear neuronal synchronization in guiding visuomotor behavior has not been studied in the visual cortex.

Bicoherence is an advanced signal processing technique capable of tracking the neuronal non-linearity and non-Gaussian signals underlying brain functions (Bullock et al., 1997; Darvas et al., 2009; Li et al., 2009, 2013). Many studies has shown that this technique can quantify the strength of non-linear phase-phase CFC [i.e., quadratic phase coupling (QPC)] between frequency components of the LFP signal (von Stein et al., 2000; Wang et al., 2007; Darvas et al., 2009; Sheremet et al., 2019). The neural generators of QPC have been reported for object coding in single neuron, in which different features of an object (e.g., size and angular speed) are encoded by a multiplicative process (Gabbiani et al., 2002). QPC has also been found in neuronal control circuits underlying sensorimotor control (Ahissar and Kleinfeld, 2003). Furthermore, QPC can effectively facilitate transmission of selective information between cortical networks (Darvas et al., 2009; Akam and Kullmann, 2014). On the other hand, it has been shown that the QPC plays a key role in multiplexing neural signals, which improves neural transmission (Akam and Kullmann, 2014).

Here, we have studied LFP signals by employing the bicoherence method to examine how non-linear neuronal synchronization in the MT area is involved in the processing of visuomotor information. For this purpose, we trained a macaque monkey to perform a visuomotor detection task. The animal had to detect a brief change in the target stimulus. Results have indicated that the strength of non-linear phase synchronization among MT neurons predicts the animal's reaction time on a trial-by-trial basis. Importantly, we observed that the non-linear phase synchronization mostly occurs in the high-gamma frequency band (100–250 Hz) of LFPs, in line with a recent study (Khamechian et al., 2019). Moreover, the result demonstrated that non-linear characteristics of neuronal synchronization are modulated when the monkey covertly attends to the stimulus inside the neuron's receptive field. Furthermore, we observed that the non-linear and the linear neuronal synchronizations potentially play a functional role in processing visuomotor information in the MT area of the visual cortex.

MATERIALS AND METHODS

Animal Welfare

All animal procedures in this study were performed at the German Primate Center in Göttingen, Germany, and were approved by the responsible regional government office [Niedersaechsisches Landesamt fuer Verbraucherschutz und Lebensmittelsicherheit (LAVES)], under the permit numbers 33.42502/08-07.02 and 33.14.42502-04-064/07. For more details on the non-human primate facilities, training facilities, and surgical techniques in this laboratory, please see Roelfsema and Treue (2014), Calapai et al. (2017), Berger et al. (2018), Pfefferle et al. (2018).

Experimental Task and Recording

A male macaque monkey was trained to fixate on a central fixation point and covertly attend to one of two coherently moving random dot patterns (RDP). Each trial was initiated by pressing a lever while maintaining the gaze on a central fixation point for 130 ms (**Figure 1**). Next, a static RDP appeared for 455 ms to cue the upcoming target's location. Following a short blank period (325 ms), two moving RDPs were shown inside and outside the receptive field (RF) of the recorded neurons for a random period of 680–4250 ms. The monkey had to release the lever immediately after the target underwent a brief change in direction of motion. The RDP's direction for target and non-target (distractor) stimuli were the same, chosen randomly from eight possible directions (0–360° with

steps of 45°). The monkey was rewarded if he correctly released the lever within 150–650 ms after the target change occurred. Trials were terminated without a reward when the monkey (i) broke the maintenance of his gaze on the fixation spot, (ii) released the lever in response to a distractor change, or (iii) responded too late after the target change. The monkey correctly detected the target changes in 86% of the trials without fixation breaks. He incorrectly terminated 3 and 11% of trials by responding to a non-target change (false alarm) and ended the trial without performing any response (miss trial), respectively.

Single-unit neural activities (SUAs) and local field potentials (LFPs) were recorded extracellularly from MT neurons using a multi-channel recording system (Mini-Matrix, Thomas Recording, and Plexon data acquisition system, Plexon Inc.). The signals were split into SUA and LFP by hardware filters. Moreover, the LFPs and the SUAs were amplified and digitized at 1 and 40 kHz, respectively. The 50 Hz noise of the power line was eliminated from the LFPs using a non-causal 4th-order Butterworth notch filter. Action potentials of recorded units were sorted online using a Plexon MAP data acquisition system (Plexon, Dallas, TX, United States). Single units were isolated online using a window discrimination procedure. The data were collected from 111 sites with five parallel electrodes, advanced separately into brain tissue to isolate direction-tuned MT neurons with overlapping RF. These electrodes were not implanted chronically but were inserted simultaneously in each experimental session. MT sites were identified by their anatomical location in cortex (using structural MRI imaging) and by the physiological properties of recorded neurons: neurons were direction-selective and the average diameter of the neuron's RF was almost equal to the RF eccentricity. The RF centers of MT neurons at different locations were predictable along the superior temporal sulcus in cortex. For more details on the experimental procedure, behavioral task, and recording details see Esghaei and Daliri (2014).

Data Analysis Procedure

In the following sections, the analyses and quantitative procedure are discussed. All analyses were implemented using MATLAB software (R2017b; MathWorks, Natick, MA, United States).

Trial Selection Procedure

We only analyzed the hit trials in which the monkey correctly detected the target change. The hit trials were sorted based on reaction times (RTs) and sub-divided into four quartiles. An equal number of these trials were selected from the first and the last quartiles and labeled as the fast and the slow trials, respectively. Through this process, there were 725 trials at each fast and slow group. We used single-unit spiking activity and the LFP of chosen trials to predict the animal's reaction time (RT). All analyses were carried out for stimulus presentation period, for a time window of 1500 ms before the target change occurred (see "Analysis window" in **Figure 1**). We chose the trials in which the target stimulus was changed 3000 ms after the trial onset. The rationale for this selection was to be ensured that the analysis window was far enough from the stimulus-evoked activities induced by the stimulus onset. We employed a built-in MATLAB function to perform digital filtering with zero-phase distortion (the *filtfilt* function).

Analysis of Bicoherence

General harmonic wavelet transform (GHWT)-based wavelet bicoherence (WBIC) (Li et al., 2009, 2011) was used to measure the quadratic phase coupling (QPC) in LFP signals. A segment-averaging approach (Hagihira et al., 2001; Li et al., 2009) was employed for calculating WBIC in order to obtain a reliable estimate of bicoherence. We used a time window of 500 ms with a 375 ms overlap to divide the LFP signal into eight time epochs. For each epoch, the GHWT-based WBIC algorithm was run to calculate bicoherence in all frequency pairs from 1 to 250 Hz, with a step of 1 Hz and bandwidth of 2 Hz. The implementation of this algorithm is briefly explained in the following (for more details on the GHWT-based WBIC algorithm, see Li et al. (2009, 2011). First, we conducted the GWHT for each epoch of a trial's LFP ($X_k(t)$, where k denotes the kth epoch of a given LFP signal) to calculate the wavelet coefficient $a_k(f, t)$ in a frequency component f. This frequency component varied from 1 to 250 Hz (as mentioned previously). Next, the normalized squared WBIC was calculated for each possible pair of frequency component as given in eq. 1:

$$b_k(f_1, f_2) = \frac{\left|B_k(f_1, f_2)\right|^2}{\sum_{t=1}^{N}\left|a_k(f_1, t)a_k(f_2, t)\right|^2 \sum_{t=1}^{N}\left|a_k(f_1 + f_2, t)\right|^2}, \quad 1 \le f_1, f_2 \le 250 \text{ Hz} \tag{1}$$

where N represents the time length of the epoch, (f_1, f_2) indicates a frequency pair (bifrequency), and B_k denotes the phase-randomized wavelet bicoherence, which is calculated as indicated in eq. 2:

$$B_k(f_1, f_2) = \sum_{t=1}^{N} a_k(f_1, t)a_k(f_2, t)a_k^*(f_1 + f_2, t)\, e^{iR\varphi_k(f_1, f_2, t)} \tag{2}$$

where $R \in [-\pi, \pi]$ is a random variable and $\varphi_k(f_1, f_2, t)$ denotes instantaneous biphase, which is calculated using the function provided in eq. 3:

$$\varphi_k(f_1, f_2, t) = \varphi_k(f_1, t) + \varphi_k(f_2, t) - \varphi_k(f_1 + f_2, t) \tag{3}$$

Next, we made use of a surrogate method to eliminate all spurious QPCs and obtained a reliable estimate for the wavelet bicoherence (Li et al., 2009, 2011). To this end, the biphase function $\varphi_k(f_1, f_2, t)$ was replaced with a new biphase $\varphi_k^{'}(f_1, f_2, t) = \varphi_k(f_1, f_2, t) + \theta$ in Eq. 2 in order to calculate a surrogated bicoherence for a given bifrequency (f_1, f_2). θ is a random variable chosen from $(-\pi, \pi]$. We generated a hundred samples of surrogated bicoherence for the bifrequency (f_1, f_2) and computed their mean (μ) and standard deviation (σ). The original bicoherence was preserved if it exceeded $\mu + 1.6\,\sigma$ of the surrogate bicoherence (as a 95% statistical threshold value); otherwise, it was set to zero. The GHWT-based WBIC method

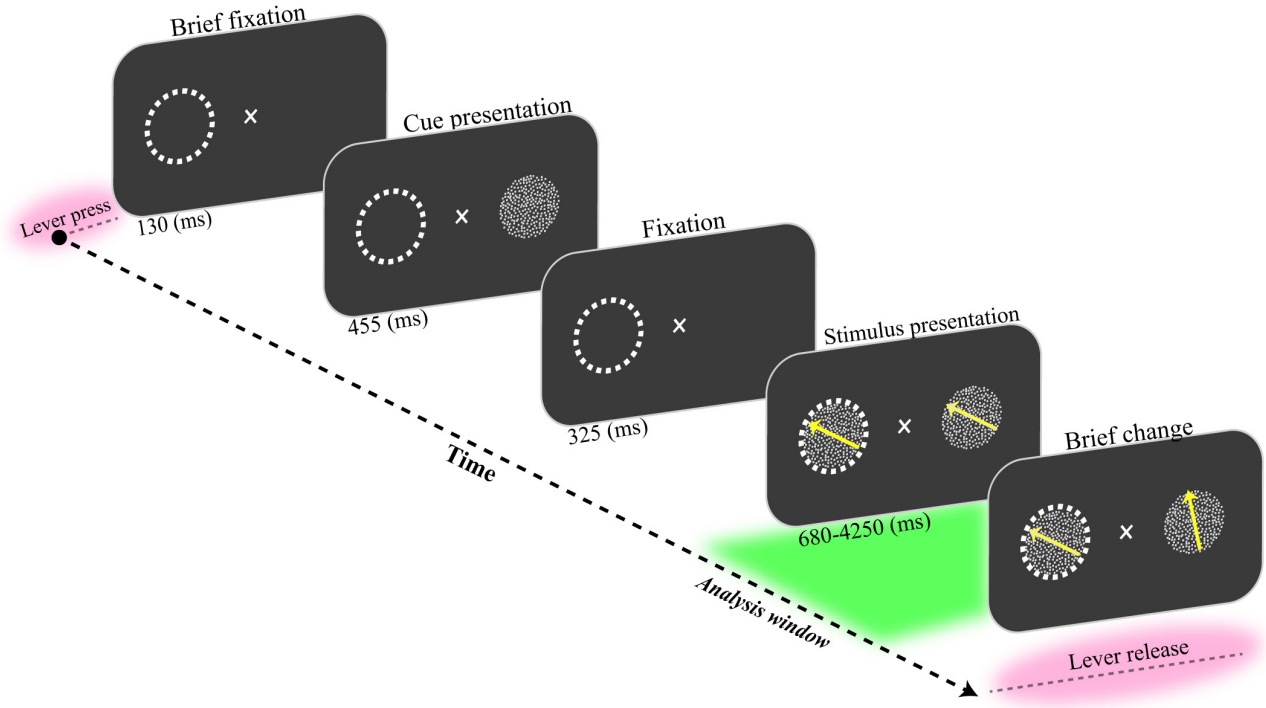

FIGURE 1 | Behavioral paradigm. The monkey had to press a lever while maintaining his gaze on a fixation spot for 130 ms. Then, a static random dot pattern (RDP) appeared for 455 ms to indicate the upcoming target location on the screen. The screen was blanked for the next 325 ms. Next, two RDPs were presented inside and outside the receptive field (marked by a dashed circle here) for a random period of 680 to 4250 ms. The monkey received a drop of juice if it correctly detected a short change in the direction of the target RDP and released the lever within a short time window (150–650 ms). The analysis window was 1500 ms preceding the direction change in the target stimulus (delineated as a green area in the figure).

provided a two-dimensional bicoherence matrix with 250 × 250 bifrequency components for each trial's LFP signal.

Quantitative Analysis of the Bicoherence Matrix Using Bicoherence Indices

We calculated four indices using the bicoherence matrix obtained for each trial. These indices were computed in WBIC studies to quantify the bicoherence matrix (Li et al., 2009, 2011, 2013; Wang et al., 2017). They were computed for each trial's bicoherence matrix as follows:

(i) Total amount of the wavelet bicoherence across all bifrequency pairs of (f_1, f_2);

$$Total\ Bic\ =\ \sum\sum b\ (f_1, f_2) \tag{4}$$

where $1 \leq f_1, f_2 \leq 250$ Hz and b is the bicoherence matrix.

(ii - iii) Eigen-decomposition for b; since bicoherence matrix is a symmetric matrix with respect to the main diagonal ($f_1 = f_2$), Eigen-decomposition can be conducted as follows:

$$b\ v_i = \lambda_i v_i, \quad \lambda_i \in \{\lambda_1\ \leq \lambda_2 \leq \ldots \leq\ \lambda_M, \quad M = 250\} \tag{5}$$

where λ_i, v_i are the eigenvalue and eigenvector, respectively. M denotes the number of frequency components (f). The maximum eigenvalue (Li et al., 2009) and Shannon entropy of the eigenvalue distribution (Cui et al., 2010; Dauwels et al., 2010; Li et al., 2011) were considered as the next bicoherence indices. The Shannon entropy of the eigenvalue distribution is computed with the following function:

$$Entropy\ of\ eigenvalues\ =\ -\ \frac{\sum_{i=1}^{M}\ \lambda_i'\ \ \log(\lambda_i')}{\log(M)} \tag{6}$$

where $\lambda_i^{'} = |\lambda_i\ |\ /\ \sum_{i=1}^{M} |\lambda_i\ |$ is the normalized absolute eigenvalue.

(iv) Average diagonal elements of the bicoherence matrix ($f_1 = f_2$):

$$Diagonal\ Bic\ =\ \frac{\sum b(f_1 = f_2, f_2 = f_1)}{M} \tag{7}$$

We computed the bicoherence indices for the fast and the slow trials in both target position conditions (target-in and target-out, see **Figure 2**). To ensure that the bicoherence indices were independent of the spectral power, a subset of the fast and the slow trials with no significant differences in their spectral power in a wide frequency band (0–500 Hz) were selected. We calculated the LFP band-power for each fast and slow trial in a wide frequency band (0–500 Hz). Consequently, the same number of trials were sub-selected from individual histogram bins of the fast and the slow band-power. This procedure provided two subsets of the fast and the slow trials that had no significant differences in

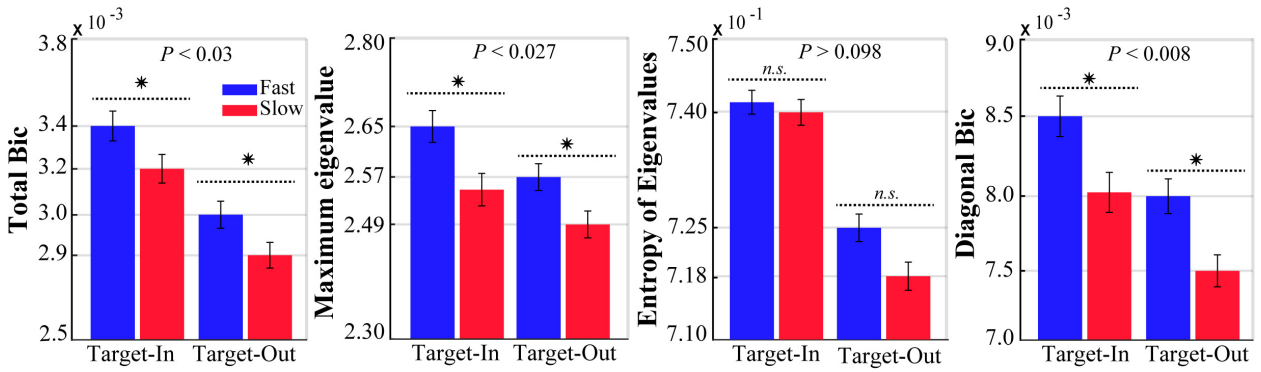

FIGURE 2 | The bicoherence indices for the fast and the slow trials at each target location condition (Target-In\Target-Out). Stars show a significant difference between the bicoherence indexes calculated for the fast (blue) and the slow (red) trials (two-sided Wilcoxon rank-sum test, *P*-values have been reported on top of the bars). Error bars indicate SEM.

power spectrums ($p > 0.98$, for both target stimulus conditions, using a two-sided Wilcoxon rank-sum test).

Selection of the Bifrequency Components by Using a Feature-Ranking Method

We compared the bifrequency components of the bicoherence matrix between the fast and the slow trials using a two-sided Wilcoxon rank-sum test. We then chose the bifrequency components with significance levels of $p < 0.01$. With this, 965 and 610 bifrequency components were sub-selected from 250×250 components in the bicoherence matrix for the target-in and the target-out conditions, respectively. The bicoherence at each bifrequency component was z-score-normalized across trials. We further employed a feature-ranking method to exclude bifrequency components yielding low performance in decoding the fast and slow trials. Firstly, a repeated holdout method was conducted for 100 independent repetitions to segregate trials into training and test subsets. At each repetition, 70% of trials (1015 trials) were randomly selected for training, and the remaining trials (435 trials) were used for the test. Then, the sub-selected bifrequency components were sorted based on their performance in decoding the fast and the slow trials in descending order. We utilized a built-in MATLAB function (*rankfeatures*, using receiver-operator-characteristic (ROC) criteria) to sort the bifrequency components. A *k*-Nearest Neighbor classifier ($k = 1$) was employed to evaluate the sorting process. This classifier assigned a query sample to the class of the single sample in the training subset that was nearest to it. We used the metric of Euclidean distance (Ed) to measure the dissimilarity between samples. The classifier was trained several times, equal to the number of bifrequency components sub-selected for each target position condition. We used the first F bifrequency components (features) for training the classifier, which had the top ranks in the sorting analysis. F was varied from 1 (the best feature) to the number of sub-selected bifrequency components (see **Figures 3A,C** x-axis). The accuracy of the classifier was assessed using the test trials. We repeated the feature-ranking method 100 times to measure the average accuracy of each F value. To extract the features that had better decoding performance, we set F based on a trade-off of maximizing two factors; (1) the ratio between the number of selected features to the total number of features (which are shown by the x-axes in **Figures 3A,C**) and (2) the decoding performance of the classifier. We extracted features for which the rank numbers were lower than $F = 140$ at each repetition. Considering this procedure, we ensured that we selected the features that provided classification accuracy above 90% in decoding the fast and the slow trials (see **Figures 3B,D**). Since the rank of a feature that had a moderate F was not consistent across different repetitions, we adopted a selection routine. This routine extracted the feature that was repeated between $F = [1-140]$ across all algorithm repetitions. We applied the feature-ranking method for each target position condition (target-in\target-out, see **Figure 3**) and obtained 85 and 89 features (bifrequency components) for the target-in and the target-out conditions, respectively.

Analysis of Spectral Power

We implemented a power spectrum analysis using a built-in MATLAB function (*pwelch function*). Briefly, the trial's LFP was sub-divided into eight segments using a 500 ms time window with a 375 ms overlap. Individual segments were windowed with a Hamming window. Then, spectral density was calculated for each segment using discrete Fourier transform. The power spectrum was calculated by the average squared magnitude of spectral densities across all segments. We calculated normalized power for each trial's LFP using the following equation:

$$Power = \frac{\sum_{f_l}^{f_h} PSD\,(f)}{\sum_{0 \le f \le 250} PSD\,(f)} \tag{8}$$

where f_l, f_h denote the lower and upper frequency bounds of the power spectral density (PSD), respectively. This equation calculates the normalized power by dividing PSD in the narrow frequency band by the total PSD. We computed PSD for frequency bands ranging between 2 and 250 Hz, with a step of 2 Hz and bandwidth of 4 Hz. This provided 124 components of

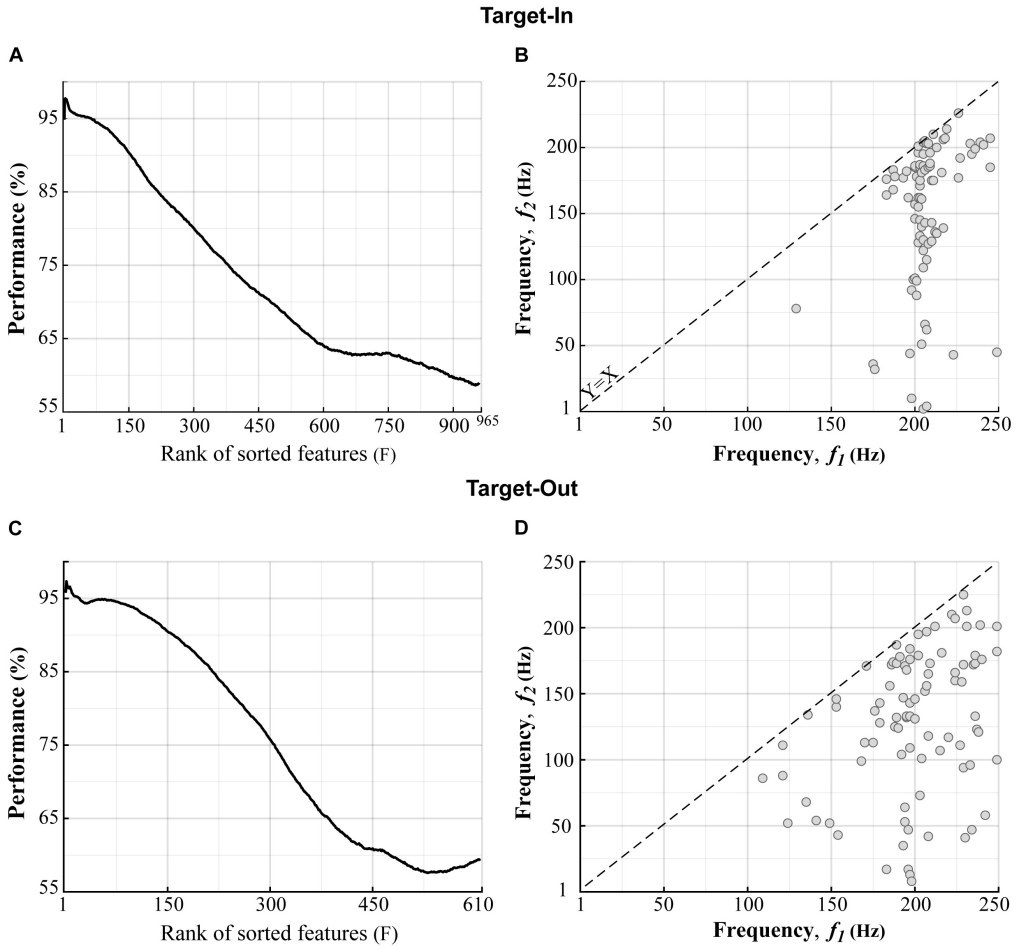

FIGURE 3 | Distributions of the bifrequency components provided high classification performance in decoding the fast and the slow trials. **(A,C)** The bifrequency components (features) in the Target-In and the Target-Out conditions, respectively, which were sorted based on the decoding performance in descending order. **(B,D)** Distributions of the bifrequency components in panels **(A,C)**, respectively, which provided performance of over 90% in decoding the fast and the slow trials. **(B,D)** show 85 and 89 bifrequency components for the Target-In and the Target-Out conditions, respectively. We did not analyze the upper bound of the $X = Y$-axis in panels **(B,D)** because of the symmetric property of the bicoherence. X-Y axes indicate the center bound of frequency bands in panels **(B,D)**.

normalized power for each trial's LFP. We applied this analysis on individual trials, including all the target position conditions (target-in\ target-out).

Analysis of Feature Extraction

We extracted three types of features from each trial's LFP, namely: (1) the bifrequency components that were sub-selected from the bicoherence matrix, (2) the bicoherence indices, and (3) the normalized power. We used the feature-ranking method to extract the best features provided a high decoding performance for classifying the fast and the slow trials, but here, we only extracted the 60 first features (by setting $F = 60$) instead of the 140 features ($F = 140$) extracted in the original algorithm. Furthermore, we selected features that repeated across 90% of algorithm repetitions within the first 60 features (the algorithm was repeated 100 times). The classifier could reach an accuracy of over 95% in decoding the fast and the slow trials. However, despite choosing different F values for the analyses shown in **Figures 3**, **4**, the number of selected features in both analyses were comparable. In more detail, about 15–30% of the total number of features were sub-selected in each analysis.

Categorization of the Bifrequency Components

We categorized the bifrequency components sub-selected from a trial's feature vector into the Bic_{fast} or Bic_{slow} group based on the average bicoherence in the fast and the slow trials. To this end, a bifrequency component was labeled as Bic_{fast} or Bic_{slow} if the average bicoherence for that bifrequency component was larger in the fast than the slow trials or vice versa, respectively (**Figure 5**). We further calculated the median and the median absolute deviation (MAD) for each Bic_{fast} and Bic_{slow} group. The MAD is calculated using the following equation:

$$MAD = median\ \left|X - median\ (X)\right| \tag{9}$$

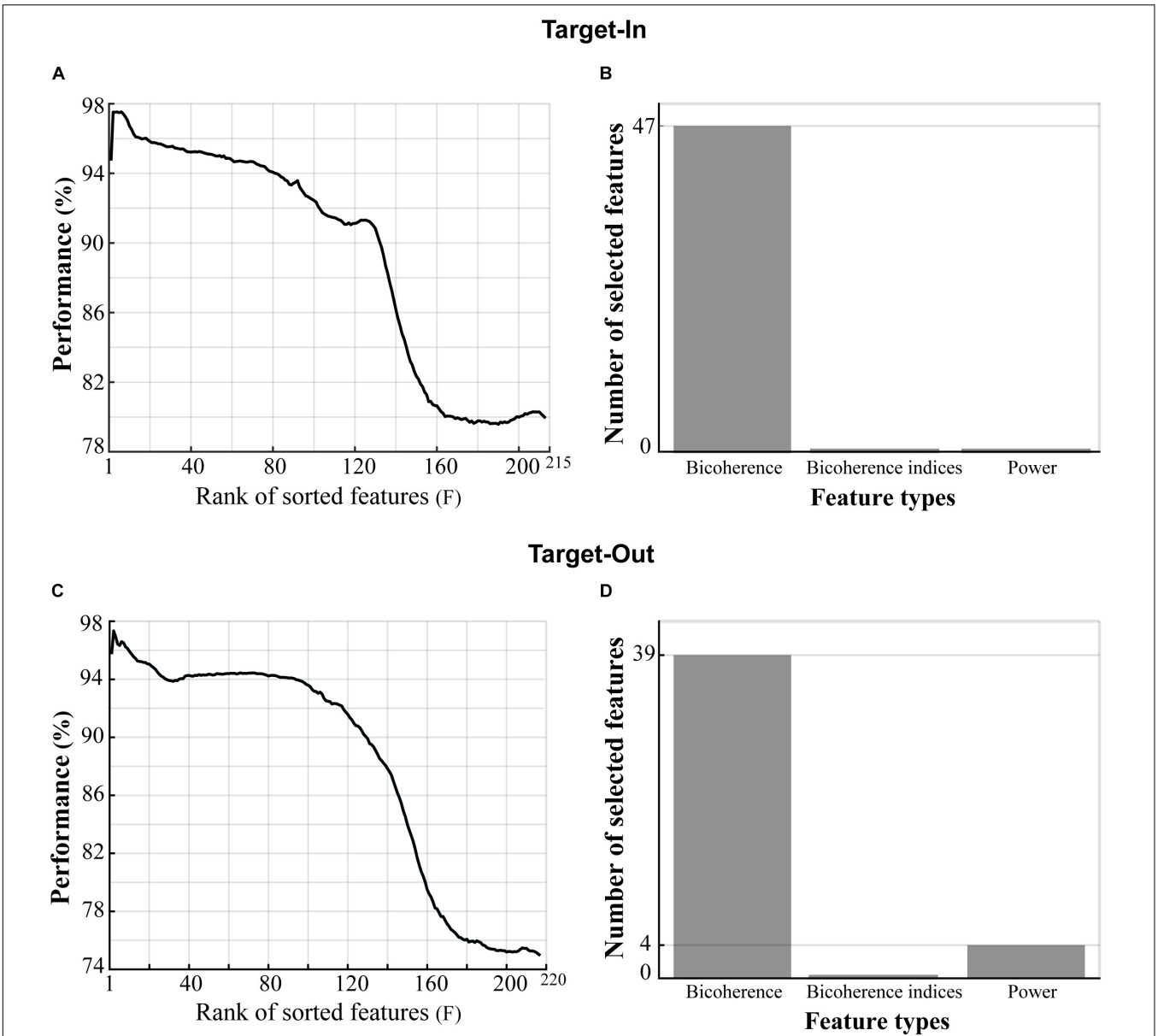

FIGURE 4 | Type of features providing high classification performance in decoding the fast and the slow trials. **(A,C)** Three types of features (shown on the *x*-axis of panels **B,D**) in the Target-In and the Target-Out conditions, respectively, which were sorted based on the decoding performance in descending order. **(B,D)** Types of features in panels **(A,C)**, respectively, which provided classification performance of over 95% in decoding the fast and the slow trials. The selected bicoherence in panels **(B,D)** lies on the high-gamma frequency band (150–250 Hz, see **Figure 5**). The selected powers in panel **(D)** lie between 190 and 200 Hz.

Where X is a vector of the bifrequency components at each Bic_{fast} or Bic_{slow} group. We conducted a permutation test to analyze significant difference between bifrequency distributions of Bic_{fast} and Bic_{slow} groups. The Ed was calculated between the bifrequency components and the corresponding median at the Bic_{fast} and Bic_{slow} groups. Then, the Eds of Bic_{fast} and Bic_{slow} were randomly shuffled between these groups 100,000 times. For each repetition, we calculated an absolute difference between the average Ed in the Pseudo-Bic_{fast} and Pseudo-Bic_{slow} group. Then, the proportion of repetitions with absolute differences larger than the original absolute difference was calculated. The proportion showed a significant difference between the bifrequency distribution of the Bic_{fast} and Bic_{slow}group if it was smaller than 0.05.

Analysis of QPC Temporal Dynamic for the Fast and the Slow Trials

We employed a time window of 150 ms with a 125 ms overlap to calculate the temporal dynamic of the quadratic phase coupling (QPC) in the same analysis time window used for the original bicoherence analyses (**Figure 7**). We ensured that the analysis

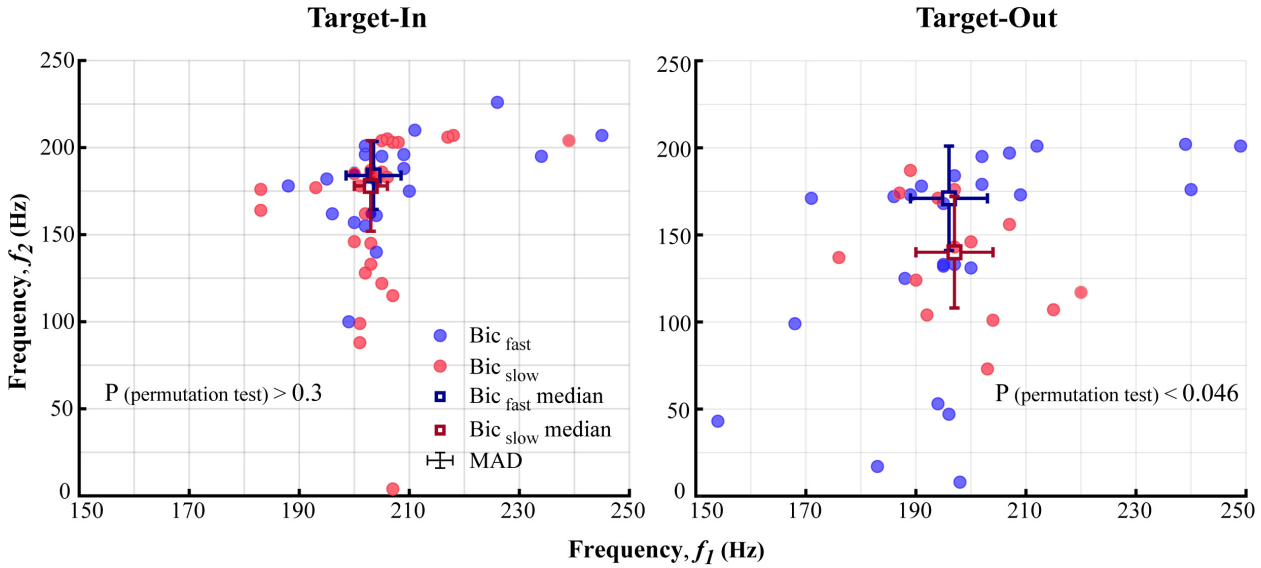

FIGURE 5 | The characteristic bifrequency of QPC in fast and slow trials for each target position condition (Target-In\Target-Out). The blue (Bic_{fast}) and the red (Bic_{slow}) circles show the bifrequency components in which the average bicoherence was larger in fast rather than slow trials, and vice versa, respectively. Blue and red squares represent medians of Bic_{fast} and Bic_{slow} groups, respectively. Error-bars demonstrate the median absolute deviation (MAD) calculated for each frequency axis. Distributions of the bifrequency component in Bic_{fast} and Bic_{slow} groups are significantly different in the Target-Out condition ($p < 0.046$; using a permutation test). This is not the case for the Target-In condition ($p > 0.3$, using a permutation test).

time window was at least 590 ms after the stimulus onset. Firstly, we filtered the trial's LFP with the GWHT to calculate wavelet coefficients $a\ (f_1, t)$ and $a\ (f_2, t)$ in a given frequency pair at the Bic_{fast} or Bic_{slow} group. Second, we took advantage of the GHWT-based WBIC algorithm (Eq. 1) to compute the bicoherence in each frequency pair for each time epoch. Next, the bicoherences were averaged across frequency pairs in each Bic_{fast} and Bic_{slow} group. Eventually, we averaged the bicoherence at each time epoch for the fast and the slow trials. The following equation computes the bicoherence for each time epoch in the analysis window:

$$Bic(t_i) = \frac{1}{N_F \times N_{Tr}} \sum_{j=1}^{N_{Tr}} \sum_{l=1}^{N_f} b_i^j(f_1^l, f_2^l) \quad i = 1\ ,\ 2\ ,\ \ldots\ ,\ N_{ep} \tag{10}$$

where $b_i^j(\cdot)$ indicates the WBIC of the ith time epoch in the jth trial, (f_1^l, f_2^l) shows the lth frequency component in the Bic_{fast} and Bic_{slow} group, N_{Tr} denotes the total number of trials, N_f represents the total number of bifrequency components at each Bic_{fast} and Bic_{slow} group, and $N_{ep} = 55$ is the total number of time epochs in the analysis window. The $Bic(\cdot)$ lies between [0–1] in which zero indicates no QPC and 1 reflects perfect QPC, respectively. We used a permutation test analysis to characterize the time epochs with a significant difference between the QPC of the fast and the slow trials. The QPC of trials at each time epoch were randomly shuffled between the fast and the slow groups 1000 times. For each repetition, we calculated the difference between average QPCs in the shuffled fast and the shuffled slow trials. Then, the proportion of repetitions with larger absolute differences compared to the original absolute difference was calculated. The time epochs that had a proportion smaller than 0.05 were considered the time epoch with a significant difference between QPCs in the fast and the slow trials. We next used false discovery rate (FDR) for multiple comparisons.

Analysis of Correlation Between Single-Unit Spiking Activity and Bicoherence

We pooled the fast and the slow trials, regardless of their behavioral outcomes. First, we calculated the bicoherence at each bifrequency component in the Bic_{fast} and Bic_{slow} group to measure correlation between single-unit spiking activity and bicoherence. Second, the bicoherence at each bifrequency component was z-score-normalized across trials. Third, the trial's bicoherence was averaged across all bifrequency components at each Bic_{fast} and Bic_{slow} group. Fourth, we calculated the trial's spike-rate using single-unit spiking activity. The analysis of spike-rate was conducted for the same time window used in the bicoherence analyses. Fifth, the Spearman's correlation was employed to calculate the correlation between the trial's spike-rate and the trial's bicoherence at each Bic_{fast} and Bic_{slow} group (see **Figures 6**, **8**).

RESULTS

To study the functional interaction of neural circuits underlying behavior, we trained a monkey to perform a change detection task. In brief, the monkey had to covertly attend to one (target) of two coherently random dot patterns (RDP). The monkey

was rewarded with a drop of juice if it correctly detected a short direction change in the target RDP (**Figure 1**, see also section "Materials and Methods"). The monkey correctly reported the target change in 86% of trials without breaking its eye fixation. We recorded the local field potentials (LFP) and the single-unit spiking activity from the MT area while the monkey performed the task. To study the neural process underlying behavior, we analyzed the hit trials in which the monkey correctly detected the target change. The hit trials were subdivided into fast and slow trials based on the animal's reaction time (see section "Materials and Methods"). Next, we calculated the bicoherence for LFPs to investigate how the non-linear neuronal synchronization likely leads to fast or slow behavior. We employed general harmonic wavelet transform (GHWT)-based wavelet bicoherence (WBIC) (Li et al., 2009, 2011) to measure the strength of QPC in LFP signals. We applied GHWT-based WBIC for a time window of 1500 ms before the target change occurred. Our analyses indicated that the QPC, especially in high-gamma frequencies (150–250 Hz), can reliably decode the animal's reaction time. Moreover, we observed that the characteristic frequency pair of QPC are selective to the target position condition (target-in\target-out) and the speed of visuomotor behavior (fast\ slow).

QPC Influences Behavior Systematically

Analysis of QPC allows us to measure the phase synchrony between three signals with different frequencies. We applied the GHWT-based WBIC method on LFPs to calculate the bicoherence for each trial. We calculated the bicoherence for all frequency pairs [e.g., (f_1, f_2)] between 1 and 250 Hz, with a step of 1 Hz and bandwidth of 2 Hz. This provided a two-dimensional matrix (bicoherence matrix) with 250×250 components for each trial. Each element in the bicoherence matrix represents the strength of QPC in a pair of frequency components (bifrequency) in the LFP spectrum. To analyze the bicoherence matrix, we calculated four indices (Li et al., 2009, 2011, 2013; Wang et al., 2017), namely (I) total bicoherence (Total Bic), (II) average diagonal elements (Diagonal Bic), (III) maximum eigenvalue, and (IV) Shannon entropy of the eigenvalue distribution (see section "Materials and Methods"). **Figure 2** shows the bicoherence indices in the fast and the slow trials (blue and red, respectively) for the target-in and target-out condition. The result clearly demonstrates that the strength of QPC in fast trials is significantly larger than in slow trials for the three bicoherence indices and both target position conditions ($p < 0.03$, using two-sided Wilcoxon rank-sum test, excluding the significance level in the Shannon entropy of eigenvalues). In addition, these differences in bicoherence indices are not due to the difference between the length of stimulus presentation in the fast and the slow trials (see **Supplementary Figure S1**). Moreover, it is visually evident that the strength of bicoherence indices is clearly enhanced in the target-in condition compared with the target-out condition, irrespective of the animal's reaction time. This observation suggests that a cognitive process like selective attention probably modulates the QPCs in the target-in condition. However, the QPC enhancement in fast trials among both target position conditions reflects that the QPC is potentially not a cortical function that is preferably processing only that stimulus placed inside the neuron's RF.

QPC in High-Gamma Frequencies Plays a Crucial Role in Guiding Behavior

We next examined how neuronal oscillatory activities at different frequencies can individually or interactively contribute to the processing of visuomotor information in MT cortex. For this purpose, we analyzed the bicoherence matrix to find the bifrequency component that provided maximum discrimination between the fast and the slow trials. Firstly, we extracted the bifrequency components that showed a significant difference between the fast and the slow trials ($p < 0.01$using two-sided Wilcoxon rank-sum test). Then, a feature ranking method was employed for sub-selecting the bifrequency components provided the classification performance above 90% in decoding the fast and the slow trials (see section "Materials and Methods" for more details). **Figures 3A,C** show the bifrequency components (features) sorted based on their decoding performances in descending order in the target-in and the target-out condition, respectively. **Figures 3B,D** demonstrate the distributions of the bifrequency components for the features in **Figures 3A,C**, respectively, showing a decoding performance above 90%. There are 85 and 89 bifrequency components in **Figures 3B,D**, respectively. It is visually evident that the selected bifrequency components are distributed in a broadband high-gamma frequency range for each target position condition (100–250 Hz). In addition, the result clearly indicates that the selected bifrequency components in the target-out condition are more widely distributed than the target-in condition in the high-gamma frequency range.

We next examined the contributive role of the spectral power to the bicoherence in decoding the animal's RT. The idea is that the power spectrum does not retain the phase information of the signal but captures the statistical property of the signal's Gaussianity. In contrast, the bicoherence can extract information relevant to the signal's non-Gaussianity and signal phase spectrum (Nikias and Mendel, 1993). We computed LFP power in narrow frequency bands between 2 to 250 Hz, with a step of 2 Hz and bandwidth of 4 Hz. Then, the LFP powers were normalized to the total power in 2–250 Hz *(see materials and methods)*. Next, we defined a feature vector for each trial including three types of features: (I) the selected bifrequency components shown in **Figures 3B,D**, (II) the normalized spectral powers, and (III) the bicoherence indices (see section "Materials and Methods"). We employed the same selection routine used for **Figures 3B,D** to select the best feature from the feature vector. In brief, we sorted the features based on decoding performances in descending order. We then extracted features yielding classification performance of over 95% in decoding the fast and the slow trials (see section "Materials and Methods" for more details). **Figures 4A,C** show the sorted features in the target-in and the target-out conditions, respectively. **Figures 4B,D** demonstrate the types of selected features extracted from **Figures 4A,C**, respectively. The result

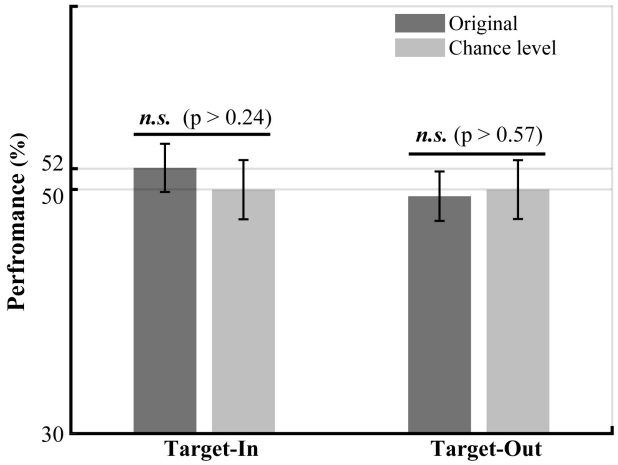

FIGURE 6 | The performance of LFP power in the high-gamma frequency band (190–210 Hz) in decoding the fast and the slow trials. The LFP power in the high-gamma frequency band was calculated for the fast and the slow trials using the Welch method (see Materials and methods). The bounds of high-gamma frequencies (i.e., 190 to 210 Hz) were defined based on the frequency ranges in **Figure 5** with a high concentration of bifrequency components. We normalized the high-gamma power to the average high-gamma power at each site. The performance of the high-gamma power for decoding the fast and the slow trials was calculated by employing a *k*-Nearest Neighbor classifier. We used a repeated holdout method (100 times) to subdivide the trials into the training and the test subsets. In each repetition, we used 70% of the trials for training and the remaining trials for the test. The chance level was calculated for each target position condition by repeatedly (100 times) shuffling the trials between the fast and the slow groups. The result clearly indicates that the decoding performance of high-gamma power is not significantly different between the original and the shuffled high-gamma power in each target position condition (permutation test, $p > 0.05$). The *p*-values have been shown on top of the bars.

is clearly evident that bicoherence is the most frequent type of feature that was selected for each target position condition. This observation suggests that QPC in the high-gamma frequency band functionally plays a key role in guiding the fast and the slow behavioral responses.

Switching Toward the Neuron's RF Modulates the Characteristic Bifrequency of the Fast and the Slow Trials

To further investigate how the selected bifrequency components in **Figures 4B,D** are distributed across the bifrequency map, we subdivided the bifrequency component into Bic_{fast} and Bic_{fast} groups based on average bicoherence in the fast and the slow trials. In more detail, a bifrequency component was labeled as Bic_{fast} or Bic_{slow} if the its average bicoherence was larger in the fast compared to the slow trials or vice versa, respectively. We further calculated the median and the median absolute distance (MAD) for the bifrequency component of the Bic_{fast} and Bic_{fast} group per target position condition (see section "Materials and Methods"). **Figures 5** and **Supplementary Figure S2** show the distribution of bifrequency components in Bic_{fast} and Bic_{slow} groups for each position condition. The result indicates that the QPC in a narrower band of high-gamma frequencies (i.e., 150–250 Hz, instead of 100–250 Hz in **Figures 3B,D**) is more implicated in guiding visuomotor behavior. In addition, the high-gamma QPC is disassociated from potential differences between high-gamma powers in the fast and the slow trials (see **Figure 6**). The result demonstrates that distributions of the bifrequency in Bic_{fast} and Bic_{slow} groups are not significantly different for the target-in condition ($p > 0.3$, using a permutation test, see section "Materials and Methods") and show a significant difference for the target-out condition ($p < 0.046$, using a permutation test, see section "Materials and Methods"). Given that Bic_{fast} and Bic_{slow} potentially represent the characteristic bifrequency of QPC in the fast and the slow trials, respectively, the result visually indicates that the distributions of characteristic bifrequencies in the fast and the slow trials are clearly different across target position conditions. In more detail, we observe that the medians of characteristic bifrequency in Bic_{fast} and Bic_{slow} groups increases for (7 Hz, 7 Hz) and (6 Hz, 38 Hz) for each (f_1, f_2) dimension, respectively, when the monkey performed the target-in condition. In addition, the result clearly illustrates that the median of characteristic bifrequency in slow trials is strongly modulated by switching toward the neuron's RF (i.e., target-in condition). But this is not the case for the fast trials. Despite this observation, the increase of medians in the target-in condition suggests that the frequency of non-linear coupling increases when the monkey attends to the stimulus inside the neuron's RF. Despite the different distributions of the characteristic bifrequencies in fast and slow trials across target position conditions, the decoding performance of QPC is similar across the target-in and the target-out conditions (see **Figures 3B,D**; see section "Materials and Methods").

QPC of Characteristic Bifrequencies in the Fast and the Slow Trials Follow Different Temporal Dynamics Among Target Position Conditions

Our analyses highlighted that the QPC in the high-gamma frequency band plays a crucial role in processing visuomotor information. In addition, we observed that the fast and the slow trials are discriminated by the distinct characteristic bifrequency of QPC in the MT area. To study the time dynamics of QPC in the fast and the slow trials, we used a time window of 150 ms with a 125 ms overlap. We calculated the bicoherence at each characteristic bifrequency of Bic_{fast} and Bic_{slow} for a given time window. Then, the bicoherence was averaged across all characteristic bifrequencies and trials for each time window (see section "Materials and Methods" for more details). **Figure 7** shows the temporal dynamics of QPC in the fast and the slow trials using characteristic bifrequncies of Bic_{fast} and Bic_{slow} groups in each target position condition. We observed that the QPC in the fast and the slow trials follows a distinct temporal pattern at each target position condition using characteristic bifrequencies of Bic_{fast} and Bic_{slow} groups. In addition, it is visually evident that the magnitude of the significant QPC difference between the fast and the slow trials in **Figure 7C** is

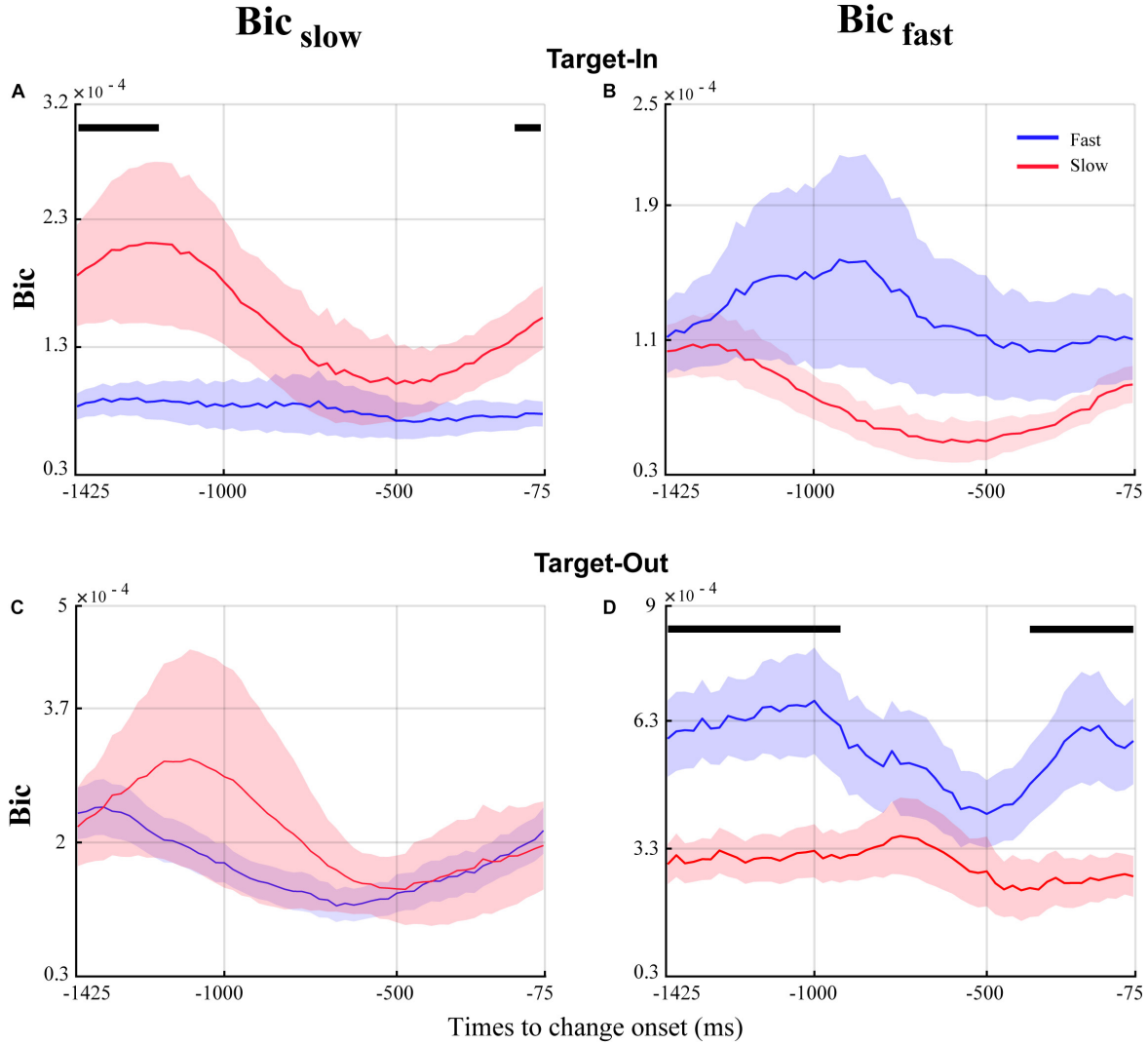

FIGURE 7 | Temporal dynamics of QPC in the fast and the slow trials for the time window before the target change. The blue and red curves show the temporal dynamics of QPC for the fast and the slow trials, respectively. The curves show the average bicoherence for the bifrequency component of the Bic_{fast} (right column) and Bic_{slow} (left column) groups shown in **Figure 5**. The black lines on top of the traces mark the times that the QPC of the fast and the slow trials are significantly different ($p < 0.05$, permutation test, FDR corrected for multiple comparisons).

somewhat enhanced in **Figure 7A** (comparing the times showing a significant difference between red and blue curves in **Figure 7A** with the corresponding times in **Figure 7C**). In other words, the significant QPC difference between fast and slow trials in characteristic bifrequencies of Bic_{slow} is modulated when the monkey attends to the stimulus inside the neuron's RF. In contrast, we observe that the magnitude of the significant QPC difference between the fast and the slow trials in **Figure 7D** strongly decreases in **Figure 7B** (comparing the times showing a significant difference between the red and blue curves in **Figure 7D** with the corresponding times in **Figure 7B**). In other words, the significant QPC difference between fast and slow trials in the characteristic bifrequencies of Bic_{fast} strongly decreases when the animal attends to the stimulus inside the neuron's RF. We hypothesize that these contrary observations in the characteristic bifrequency of the fast and the slow behavior (i.e., Bic_{fast} and Bic_{slow}, respectively) along target position conditions are potentially due to the influence of a cognitive function like attention. Our hypothesis is in line with previous studies that have shown that attention could decouple sensory neurons and thereby enhance the neural representation of relevant stimuli to effectively guide a fast behavioral reaction (Esghaei et al., 2015, 2018; Spyropoulos et al., 2018).

QPC Is Anti-correlated With the Neuronal Spike Rate Exclusively for the Target-in Condition

We calculated the correlation between the QPC and the single-unit spiking activities to study how neuronal non-linear coupling

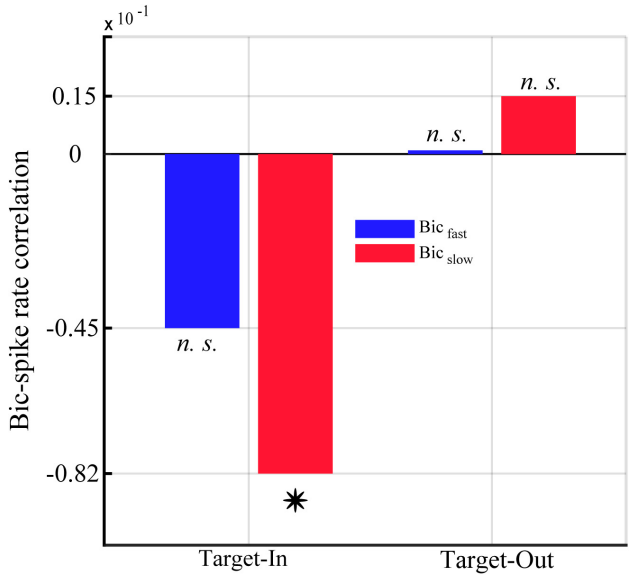

FIGURE 8 | Correlation between the single-unit spike rate and the QPC of characteristic bifrequencies in Bic_{fast} and Bic_{slow} groups for the Target-In and the Target-Out condition. *Y*-axis represents the strength of correlation between the QPC and the spike rate using Spearman's correlation method. The star shows a significant correlation in Bic_{slow} groups at the Target-In condition ($p < 0.002$, Spearman correlation).

potentially influences the neuronal output (see section "Materials and Methods"). We computed the single-unit spike rate as well as the average normalized QPC for each trial using characteristic bifrequencies of Bic_{fast} and Bic_{slow} groups (shown in **Figure 5**). **Figure 8** shows the Spearman correlation between spike rates and QPCs in each target position condition. The result indicates that the QPC and the spike rate are strongly anti-correlated in the target-in condition. Notably, we observe that the spike rates and the QPC of the Bic_{slow} group are significantly anti-correlated in the target-in condition ($p < 0.002$, using Spearman correlation). In contrast, there is a negligible non-significant positive correlation between QPCs and the spike-rates for the target-out condition.

DISCUSSION

Many studies have highlighted that oscillatory activity plays a mediating role in the neuronal coupling underlying cognitive functions (Lisman and Jensen, 2013; Khodagholy et al., 2017; Rohenkohl et al., 2018). However, the relationship between this neuronal coupling and behavior has not been studied in the visual cortex.

In this study, we recorded the LFP and the single-unit spiking activity from the visual area MT of a behaving monkey. The animal had to covertly attend to one of two RDPs placed inside or outside the RF of recorded neurons and detect a short direction change in the target stimulus. We examined how linear and non-linear neural synchronization could influence the animal's RT. For this purpose, the spectral representation of the second-order statistics (i.e., the power spectrum) and the third-order statistics (i.e., bicoherence) were calculated for LFPs on a trial-by-trial basis.

We measured the strength of non-linear coupling between all frequency pairs in the LFP spectrum (1–250 Hz) using four bicoherence indices. The bicoherence indices were: (i) total Bic, which reflects the strength of QPC between different low-frequency oscillations and one of the high-frequency oscillations, which is useful for investigating the strength of rhythmic synchronization between neuronal populations oscillating at different frequencies (Li and Li, 2016), (ii) maximum eigenvalue, (iii) Shannon entropy of eigenvalues, which measures information on the synchronization between oscillatory activities in the neuronal population (Li and Li, 2016), and (iv) diagonal Bic, which reveals the presence of self-frequency and self-phase coupling in neural circuits (Muthuswamy et al., 1999). We selected the trials which had no significant difference between average spectral powers to prevent dependence of our analyses to the different levels of $1/f$ noise (Bédard et al., 2006; Lombardi et al., 2017). With this approach, we ensured that the signal-to-noise ratio (SNR) was not significantly different between the chosen fast and slow trials (see section "Materials and Methods"). In addition, we ensured that the change in the bicoherence indices was potentially due to the change in underlying neuronal non-linear coupling (Pesaran et al., 2018). Our analysis revealed that the strength of the non-linear coupling between the oscillatory activities of MT neurons is strongly increased in the fast rather than the slow trials (see **Figure 2**). Furthermore, we observed that switching toward the neuron's RF increases the strength of non-linear coupling between neural oscillations. We speculate that this finding is possibly due to the influence of a cognitive function like attention that enhances the non-linear synchronization between local neurons. Our hypothesis is in line with previous studies that have shown that spatial attention selectively increases the strength of synchronization between neurons processing the target stimulus (Womelsdorf et al., 2006; Zareian et al., 2018; Khamechian et al., 2019).

To further study the non-linear neuronal synchronization underlying behavior, we implemented a machine learning approach to extract the bifrequency component that accurately discriminates the fast and the slow trials (see section "Materials and Methods"). The result showed that oscillatory activities in the high-gamma frequency band (100–250 Hz) are quadratically phase-coupled in the fast and the slow trials (**Figures 3B,D**). This observation is in line with a recent study showing that the strength of neural synchronization in the high-gamma frequency band (180 to 220 Hz) predicts the animal's RT (Khamechian et al., 2019). In addition, this study also showed that the difference between high-gamma synchronizations in the fast and the slow trials cannot be attributed to the difference between the magnitude of the spike leakage onto LFPs (Khamechian et al., 2019). Many studies have suggested that interneurons contribute to the generation of high-gamma oscillations in the LFPs (Brunel and Wang, 2003; Buzsáki and Draguhn, 2004;

Henrie and Shapley, 2005; Gieselmann and Thiele, 2008; Stark et al., 2014; Suffczynski et al., 2014).

We next examined the contributive role of cortical Gaussian and non-Gaussian processes (activities) in guiding visuomotor behavior. We computed the power spectrum (as a measure of Gaussianity) and the bicoherence (as a measure of non-Gaussianity) for the fast and slow the trials. We then adopted a machine learning method (see section "Materials and Methods") to examine the potential role of these processes in predicting the animal's behavior. The results illustrated that the neural non-Gaussian process (in addition to the Gaussian process) plays a key role in coding behavioral RTs in the macaque area MT (see **Figures 4B,D**). The result is consistent with a recent study indicating that bicoherence is a biomarker candidate for identifying neurodevelopmental-behavioral disorders like attention deficit hyperactivity disorder (ADHD) (Chen et al., 2019).

We further examined the QPC to understand which specific bifrequency of the bicoherence at the broadband high-gamma frequency range (100–250 Hz, see **Figure 3**) might orchestrate the fast and the slow behaviors. The result indicated that the distribution of the characteristic bifrequency is significantly different between the fast and the slow trials, particularly for the target-out condition (see **Figure 5**). In addition, we observed that switching toward the neuron's RF enhances the characteristic bifrequency of the QPC explicitly in the slow trials. We speculate that a top-down cognitive function like attention probably modulates the characteristic bifrequency of the QPC in slow trials in the target-in condition. In addition, our observations suggest that this modulatory effect mostly occurs between the MT neurons that selectively process the target stimulus. Moreover, we observed that the characteristic bifrequency of the fast trials has similar distribution medians in the target-in and the target-out conditions. The result suggests that entire neurons in the MT area can be synchronized in the high-gamma band to efficiently process the behavioral information and facilitate a fast behavioral action. Given that spatial attention can effectively shorten RT (Posner, 1980; Womelsdorf et al., 2006) and modulate neuronal synchronization (Womelsdorf et al., 2006; Hoogenboom et al., 2010; Khamechian et al., 2019), we hypothesize that such synchronization in the fast trial can also be attributed to top-down attention. In addition, our hypothesis is in line with previous studies that suggested that attention could improve neuronal communication and thereby route the most relevant information into associative areas in the brain (Gregoriou et al., 2009; Morishima et al., 2009; Briggs et al., 2013).

We next examined the temporal dynamics of QPC in the fast and the slow trials based on the characteristic bifrequency obtained for each target position condition. The result demonstrated that switching to the neuron's RF enhances the QPC difference between the fast and the slow trials using the characteristic bifrequency of the slow behavior (i.e., Bic_{slow}). In contrast, we observed that the QPC difference between the fast and the slow trials strongly decreases in the characteristic bifrequency of the fast behavior (i.e., Bic_{fast}) when the monkey covertly attends to the stimulus inside the neuron's RF. We hypothesize that this contrary observation for the QPC difference in the characteristic bifrequency of the fast and the slow behavior is due to a cognitive function like attention. Our speculation is based on previous studies suggesting that attention can suppress the strength of coupling between oscillatory activities in the visual cortex (Esghaei et al., 2015; Spyropoulos et al., 2018).

Some physiological models have shown that decisions are formed based on accumulating sensory evidence over time to a bound (Gold and Shadlen, 2001; Palmer et al., 2005; Ratcliff and McKoon, 2008). In addition, they have indicated that these computations could shape the RT distribution and the speed of behavior. The accumulation of evidence has been observed in several electrophysiological studies at different cortical areas of monkeys (Roitman and Shadlen, 2002; Purcell et al., 2012; de Lafuente et al., 2015), rodents (Hanks et al., 2015), and humans (Kelly and O'Connell, 2013; Twomey et al., 2016). For example, some of these studies reported that oscillatory activities underlying the accumulation process follow different accumulation-to-bound dynamics that predict the behavioral RTs (Kelly and O'Connell, 2013; Twomey et al., 2016). However, it is unclear how these oscillatory activities transmit sensory information from upstream to downstream cortical areas to shape the accumulation process. Previous investigations have shown that neural oscillatory activities can interact via CFC to facilitate communication of information between brain regions (Darvas et al., 2009; Canolty and Knight, 2010; Holz et al., 2010; Fiebelkorn et al., 2018). Based on these studies, we speculate that the QPC (as a non-linear form of CFC measured by the bicoherence) could play a functional role in the transmission of the relevant information between associative neurons in the intra- or inter-areal of the cortex.

In summary, we employed bicoherence and spectral power to examine non-linear and linear neuronal coupling underlying visuomotor behavior. Our results show that: (I) the non-linear phase coupling between oscillatory activities of sensory neurons is a good candidate for predicting the speed of the animal's behavior, (II) the non-linear neuronal coupling is expressed in a broad band of high-gamma frequencies (100–250 Hz) in area MT of the macaque visual cortex, (III) the non-Gaussian cortical process (measured by the bicoherence) and the Gaussian process (measured by the spectral power) are both involved in the processing of visuomotor information, and (IV) the non-linear characteristic of neuronal synchronization among MT neurons is probably controlled by a cognitive function like selective attention.

DATA AVAILABILITY STATEMENT

The datasets analyzed in this manuscript are not publicly available. Requests to access the datasets should be directed to MD, daliri@iust.ac.ir (upon reasonable request).

ETHICS STATEMENT

The animal study was reviewed and approved by the responsible regional government office [Niedersaechsisches Landesamt fuer

Verbraucherschutz und Lebensmittelsicherheit (LAVES)] under the permit numbers 33.42502/08-07.02 and 33.14.425 02-04-064/07.

AUTHOR CONTRIBUTIONS

MK and MD designed the study, interpreted the data and wrote the manuscript. MD recorded the data. MK performed the data analyses.

FUNDING

This work was supported by grants from the Iranian Cognitive Sciences and Technologies Council (Grant 4225 to MK).

ACKNOWLEDGMENTS

We would like to thank Prof. Dr. Stefan Treue and the German Primate Center – Leibniz Institute for Primate Research (DPZ) for providing the infrastructure, intellectual, and financial support for recording the data. We also thank Mr. Morteza Saraf and Miss. Mahsa Behroozi and her colleagues for their gracious assistance in proofreading the manuscript and correcting grammatical and typographical errors.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fnins.2020.00230/full#supplementary-material

REFERENCES

Ahissar, E., and Kleinfeld, D. (2003). Closed-loop neuronal computations: focus on vibrissa somatosensation in rat. *Cereb. Cortex* 13, 53–62. doi: 10.1093/cercor/13.1.53

Akam, T., and Kullmann, D. M. (2014). Oscillatory multiplexing of population codes for selective communication in the mammalian brain. *Nat. Rev. Neurosci.* 15, 111–122. doi: 10.1038/nrn3668

Aru, J., Aru, J., Priesemann, V., Wibral, M., Lana, L., Pipa, G., et al. (2015). Untangling cross-frequency coupling in neuroscience. *Curr. Opin. Neurobiol.* 31, 51–61. doi: 10.1016/j.conb.2014.08.002

Axmacher, N., Henseler, M. M., Jensen, O., Weinreich, I., Elger, C. E., and Fell, J. (2010). Cross-frequency coupling supports multi-item working memory in the human hippocampus. *Proc. Natl. Acad. Sci. U. S. A.* 107, 3228–3233. doi: 10.1073/pnas.0911531107

Bédard, C., Kröger, H., and Destexhe, A. (2006). Does the 1/f frequency scaling of brain signals reflect self-organized critical states? *Phys. Rev. Lett.* 97:118102. doi: 10.1103/PhysRevLett.97.118102

Belitski, A., Gretton, A., Magri, C., Murayama, Y., Montemurro, M. A., Logothetis, N. K., et al. (2008). Low-frequency local field potentials and spikes in primary visual cortex convey independent visual information. *J. Neurosci.* 28, 5696–5709. doi: 10.1523/JNEUROSCI.0009-08.2008

Berger, M., Calapai, A., Stephan, V., Niessing, M., Burchardt, L., Gail, A., et al. (2018). Standardized automated training of rhesus monkeys for neuroscience research in their housing environment. *J. Neurophysiol.* 119, 796–807. doi: 10.1152/jn.00614.2017

Briggs, F., Mangun, G. R., and Usrey, W. M. (2013). Attention enhances synaptic efficacy and the signal-to-noise ratio in neural circuits. *Nature* 499, 476–480. doi: 10.1038/nature12276

Brunel, N., and Wang, X.-J. (2003). What determines the frequency of fast network oscillations with irregular neural discharges? I. Synaptic dynamics and excitation-inhibition balance. *J. Neurophysiol.* 90, 415–430. doi: 10.1152/jn.01095.2002

Bullock, T. H., Achimowicz, J. Z., Duckrow, R. B., Spencer, S. S., and Iragui-Madoz, V. J. (1997). Bicoherence of intracranial EEG in sleep, wakefulness and seizures. *Electroencephalogr. Clin. Neurophysiol.* 103, 661–678. doi: 10.1016/s0013-4694(97)00087-4

Buschman, T. J., and Miller, E. K. (2007). Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315, 1860–1862. doi: 10.1126/science.1138071

Buzsáki, G. (2006). *Rhythms of the Brain*. Oxford: Oxford University Press, 136–160.

Buzsáki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929. doi: 10.1126/science.1099745

Calapai, A., Berger, M., Niessing, M., Heisig, K., Brockhausen, R., Treue, S., et al. (2017). A cage-based training, cognitive testing and enrichment system optimized for rhesus macaques in neuroscience research. *Behav. Res. Methods* 49, 35–45. doi: 10.3758/s13428-016-0707-3

Canolty, R. T., Ganguly, K., Kennerley, S. W., Cadieu, C. F., Koepsell, K., Wallis, J. D., et al. (2010). Oscillatory phase coupling coordinates anatomically dispersed functional cell assemblies. *Proc. Natl. Acad. Sci.* 107, 17356–17361. doi: 10.1073/pnas.1008306107

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

Chen, H., Chen, W., Song, Y., Sun, L., and Li, X. (2019). EEG characteristics of children with attention-deficit/hyperactivity disorder. *Neuroscience* 406, 444–456. doi: 10.1016/j.neuroscience.2019.03.048

Cui, D., Liu, X., Wan, Y., and Li, X. (2010). Estimation of genuine and random synchronization in multivariate neural series. *Neural. Netw.* 23, 698–704. doi: 10.1016/j.neunet.2010.04.003

Darvas, F., Miller, K. J., Rao, R. P. N., and Ojemann, J. G. (2009). Nonlinear phase-phase cross-frequency coupling mediates communication between distant sites in human neocortex. *J. Neurosci.* 29, 426–435. doi: 10.1523/JNEUROSCI.3688-08.2009

Dauwels, J., Vialatte, F., Musha, T., and Cichocki, A. (2010). A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG. *Neuroimage* 49, 668–693. doi: 10.1016/j.neuroimage.2009.06.056

de Lafuente, V., Jazayeri, M., and Shadlen, M. N. (2015). Representation of accumulating evidence for a decision in two parietal areas. *J. Neurosci.* 35, 4306–4318. doi: 10.1523/JNEUROSCI.2451-14.2015

Esghaei, M., and Daliri, M. R. (2014). Decoding of visual attention from LFP signals of macaque MT. *PLoS One* 9:e100381. doi: 10.1371/journal.pone.0100381

Esghaei, M., Daliri, M. R., and Treue, S. (2015). Attention decreases phase-amplitude coupling, enhancing stimulus discriminability in cortical area MT. *Front. Neural Circuits* 9:82. doi: 10.3389/fncir.2015.00082

Esghaei, M., Daliri, M. R., and Treue, S. (2018). Attention decouples action potentials from the phase of local field potentials in macaque visual cortical area MT. *BMC Biol.* 16:86. doi: 10.1186/s12915-018-0551-2

Fiebelkorn, I. C., Pinsk, M. A., and Kastner, S. (2018). A dynamic interplay within the frontoparietal network underlies rhythmic spatial attention. *Neuron* 99, 842.e8–853.e8. doi: 10.1016/j.neuron.2018.07.038

Gabbiani, F., Krapp, H. G., Koch, C., and Laurent, G. (2002). Multiplicative computation in a visual neuron sensitive to looming. *Nature* 420, 320–324. doi: 10.1038/nature01190

Gieselmann, M. A., and Thiele, A. (2008). Comparison of spatial integration and surround suppression characteristics in spiking activity and the local field potential in macaque V1. *Eur. J. Neurosci.* 28, 447–459. doi: 10.1111/j.1460-9568.2008.06358.x

Gold, J. I., and Shadlen, M. N. (2001). Neural computations that underlie decisions about sensory stimuli. *Trends Cogn. Sci.* 5, 10–16. doi: 10.1016/s1364-6613(00)01567-9

Gregoriou, G., Gotts, S. J., Zhou, H., and Desimone, R. (2009). High-frequency, long-range coupling between prefrontal and visual cortex during attention. *Science* 324, 1207–1210. doi: 10.1126/science.1171402

Hagihira, S., Takashina, M., Mori, T., Mashimo, T., and Yoshiya, I. (2001). Practical issues in bispectral analysis of electroencephalographic signals. *Anesth. Analg* 93, 966–970. doi: 10.1097/00000539-200110000-00032

Hanks, T. D., Kopec, C. D., Brunton, B. W., Duan, C. A., Erlich, J. C., and Brody, C. D. (2015). Distinct relationships of parietal and prefrontal cortices to evidence accumulation. *Nature* 520, 220–223. doi: 10.1038/nature14066

Henrie, J. A., and Shapley, R. (2005). LFP power spectra in V1 cortex: the graded effect of stimulus contrast. *J. Neurophysiol.* 94, 479–490. doi: 10.1152/jn.00919.2004

Hipp, J. F., Engel, A. K., and Siegel, M. (2011). Oscillatory synchronization in large-scale cortical networks predicts perception. *Neuron* 69, 387–396. doi: 10.1016/j.neuron.2010.12.027

Holz, E. M., Glennon, M., Prendergast, K., and Sauseng, P. (2010). Theta-gamma phase synchronization during memory matching in visual working memory. *Neuroimage* 52, 326–335. doi: 10.1016/j.neuroimage.2010.04.003

Hoogenboom, N., Schoffelen, J.-M., Oostenveld, R., and Fries, P. (2010). Visually induced gamma-band activity predicts speed of change detection in humans. *Neuroimage* 51, 1162–1167. doi: 10.1016/j.neuroimage.2010.03.041

Igarashi, K. M., Lu, L., Colgin, L. L., Moser, M.-B., and Moser, E. I. (2014). Coordination of entorhinal-hippocampal ensemble activity during associative learning. *Nature* 510, 143–147. doi: 10.1038/nature13162

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

Kelly, S. P., and O'Connell, R. G. (2013). Internal and external influences on the rate of sensory evidence accumulation in the human brain. *J. Neurosci.* 33, 19434–19441. doi: 10.1523/JNEUROSCI.3355-13.2013

Khamechian, M. B., Kozyrev, V., Treue, S., Esghaei, M., and Daliri, M. R. (2019). Routing information flow by separate neural synchrony frequencies allows for "functionally labeled lines" in higher primate cortex. *Proc. Natl. Acad. Sci. U.S.A* 116, 12506–12515. doi: 10.1073/pnas.1819827116

Khodagholy, D., Gelinas, J. N., and Buzsáki, G. (2017). Learning-enhanced coupling between ripple oscillations in association cortices and hippocampus. *Science* 358, 369–372. doi: 10.1126/science.aan6203

Li, D., Li, X., Hagihira, S., and Sleigh, J. W. (2011). The effect of isoflurane anesthesia on the electroencephalogram assessed by harmonic wavelet bicoherence-based indices. *J. Neural Eng.* 8:56011. doi: 10.1088/1741-2560/8/5/056011

Li, D., Li, X., Hagihira, S., and Sleigh, J. W. (2013). Cross-frequency coupling during isoflurane anaesthesia as revealed by electroencephalographic harmonic wavelet bicoherence. *Br. J. Anaesth.* 110, 409–419. doi: 10.1093/bja/aes397

Li, X., Li, D., Voss, L. J., and Sleigh, J. W. (2009). The comodulation measure of neuronal oscillations with general harmonic wavelet bicoherence and application to sleep analysis. *Neuroimage* 48, 501–514. doi: 10.1016/j.neuroimage.2009.07.008

Li, D., and Li, X. (ed.). (2016). "The comodulation measure of neuronal oscillations," in *Signal Processing in Neuroscience*, (Singapore: Springer Singapore), 215–233. doi: 10.1007/978-981-10-1822-0_11

Lisman, J. E., and Jensen, O. (2013). The θ-γ neural code. *Neuron* 77, 1002–1016. doi: 10.1016/j.neuron.2013.03.007

Liu, J., and Newsome, W. T. (2005). Correlation between speed perception and neural activity in the middle temporal visual area. *J. Neurosci.* 25, 711–722. doi: 10.1523/JNEUROSCI.4034-04.2005

Liu, J., and Newsome, W. T. (2006). Local field potential in cortical area MT: stimulus tuning and behavioral correlations. *J. Neurosci.* 26, 7779–7790. doi: 10.1523/JNEUROSCI.5052-05.2006

Lombardi, F., Herrmann, H. J., and de Arcangelis, L. (2017). Balance of excitation and inhibition determines 1/f power spectrum in neuronal networks. *Chaos* 27:47402. doi: 10.1063/1.4979043

Moore, T., and Armstrong, K. M. (2003). Selective gating of visual signals by microstimulation of frontal cortex. *Nature* 421, 370–373. doi: 10.1038/nature01341

Morishima, Y., Akaishi, R., Yamada, Y., Okuda, J., Toma, K., and Sakai, K. (2009). Task-specific signal transmission from prefrontal cortex in visual selective attention. *Nat. Neurosci.* 12, 85–91. doi: 10.1038/nn.2237

Muthuswamy, J., Sherman, D. L., and Thakor, N. V. (1999). Higher-order spectral analysis of burst patterns in EEG. *IEEE Trans. Biomed. Eng* 46, 92–99. doi: 10.1109/10.736762

Nikias, C. L., and Mendel, J. M. (1993). Signal processing with higher-order spectra. *IEEE Signal Process. Mag.* 10, 10–37. doi: 10.1109/79.221324

Palmer, J., Huk, A. C., and Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *J. Vis.* 5, 376–404. doi: 10.1167/5.5.1

Parto Dezfouli, M., Khamechian, M. B., Treue, S., Esghaei, M., and Daliri, M. R. (2018). Neural activity predicts reaction in primates long before a behavioral response. *Front. Behav. Neurosci.* 12:207. doi: 10.3389/fnbeh.2018.00207

Pesaran, B., Vinck, M., Einevoll, G. T., Sirota, A., Fries, P., Siegel, M., et al. (2018). Investigating large-scale brain dynamics using field potential recordings: analysis and interpretation. *Nat. Neurosci.* 21, 903–919. doi: 10.1038/s41593-018-0171-8

Pfefferle, D., Plümer, S., Burchardt, L., Treue, S., and Gail, A. (2018). Assessment of stress responses in rhesus macaques (*Macaca mulatta*) to daily routine procedures in system neuroscience based on salivary cortisol concentrations. *PLoS One* 13:e0190190. doi: 10.1371/journal.pone.0190190

Posner, M. I. (1980). Orienting of attention. *Q. J. Exp. Psychol.* 32, 3–25. doi: 10.1080/00335558008248231

Purcell, B. A., Schall, J. D., Logan, G. D., and Palmeri, T. J. (2012). From salience to saccades: multiple-alternative gated stochastic accumulator model of visual search. *J. Neurosci.* 32, 3433–3446. doi: 10.1523/JNEUROSCI.4622-11.2012

Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922. doi: 10.1162/neco.2008.12-06-420

Roelfsema, P. R., and Treue, S. (2014). Basic neuroscience research with nonhuman primates: a small but indispensable component of biomedical research. *Neuron* 82, 1200–1204. doi: 10.1016/j.neuron.2014.06.003

Rohenkohl, G., Bosman, C. A., and Fries, P. (2018). Gamma synchronization between V1 and V4 improves behavioral performance. *Neuron* 100, 953.e3–963.e3. doi: 10.1016/j.neuron.2018.09.019

Roitman, J. D., and Shadlen, M. N. (2002). Response of neurons in the lateral intraparietal area during a combined visual discrimination reaction time task. *J. Neurosci.* 22, 9475–9489. doi: 10.1523/JNEUROSCI.22-21-09475.2002

Saalmann, Y. B., Pigarev, I. N., and Vidyasagar, T. R. (2007). Neural mechanisms of visual attention: how top-down feedback highlights relevant locations. *Science* 316, 1612–1615. doi: 10.1126/science.1139140

Saleh, M., Reimer, J., Penn, R., Ojakangas, C. L., and Hatsopoulos, N. G. (2010). Fast and slow oscillations in human primary motor cortex predict oncoming behaviorally relevant cues. *Neuron* 65, 461–471. doi: 10.1016/j.neuron.2010.02.001

Schroeder, C. E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* 32, 9–18. doi: 10.1016/j.tins.2008.09.012

Sheremet, A., Kennedy, J. P., Qin, Y., Zhou, Y., Lovett, S. D., Burke, S. N., et al. (2019). Theta-gamma cascades and running speed. *J. Neurophysiol.* 121, 444–458. doi: 10.1152/jn.00636.2018

Siegel, M., Donner, T. H., Oostenveld, R., Fries, P., and Engel, A. K. (2007). High-frequency activity in human visual cortex is modulated by visual motion strength. *Cereb. Cortex* 17, 732–741. doi: 10.1093/cercor/bhk025

Siegel, M., Donner, T. H., Oostenveld, R., Fries, P., and Engel, A. K. (2008). Neuronal synchronization along the dorsal visual pathway reflects the focus of spatial attention. *Neuron* 60, 709–719. doi: 10.1016/j.neuron.2008.09.010

Smith, J. E. T., Beliveau, V., Schoen, A., Remz, J., Zhan, C. A., and Cook, E. P. (2015). Dynamics of the functional link between area MT LFPs and motion detection. *J. Neurophysiol.* 114, 80–98. doi: 10.1152/jn.00058.2015

Spyropoulos, G., Bosman, C. A., and Fries, P. (2018). A theta rhythm in macaque visual cortex and its attentional modulation. *Proc. Natl. Acad. Sci. U.S.A.* 115, E5614–E5623. doi: 10.1073/pnas.1719433115

Stark, E., Roux, L., Eichler, R., Senzai, Y., Royer, S., and Buzsáki, G. (2014). Pyramidal cell-interneuron interactions underlie hippocampal ripple oscillations. *Neuron* 83, 467–480. doi: 10.1016/j.neuron.2014.06.023

Suffczynski, P., Crone, N. E., and Franaszczuk, P. J. (2014). Afferent inputs to cortical fast-spiking interneurons organize pyramidal cell network oscillations at high-gamma frequencies (60-200 Hz). *J. Neurophysiol.* 112, 3001–3011. doi: 10.1152/jn.00844.2013

Tort, A. B. L., Komorowski, R. W., Manns, J. R., Kopell, N. J., and Eichenbaum, H. (2009). Theta-gamma coupling increases during the learning of item-context associations. *Proc. Natl. Acad. Sci. U. .A.* 106, 20942–20947. doi: 10.1073/pnas.0911331106

Twomey, D. M., Kelly, S. P., and O'Connell, R. G. (2016). Abstract and effector-selective decision signals exhibit qualitatively distinct dynamics before delayed perceptual reports. *J. Neurosci.* 36, 7346–7352. doi: 10.1523/JNEUROSCI.4162-15.2016

van Kerkoerle, T., Self, M. W., Dagnino, B., Gariel-Mathis, M.-A., Poort, J., van der Togt, C., et al. (2014). Alpha and gamma oscillations characterize feedback and feedforward processing in monkey visual cortex. *Proc. Natl. Acad. Sci. U.S.A.* 111, 14332–14341. doi: 10.1073/pnas.1402773111

von Stein, A., Chiang, C., and König, P. (2000). Top-down processing mediated by interareal synchronization. *Proc. Natl. Acad. Sci. U.S.A.* 97, 14748–14753. doi: 10.1073/pnas.97.26.14748

Wang, J., Fang, Y., Wang, X., Yang, H., Yu, X., and Wang, H. (2017). Enhanced gamma activity and cross-frequency interaction of resting-state electroencephalographic oscillations in patients with Alzheimer's Disease. *Front. Aging Neurosci.* 9:243. doi: 10.3389/fnagi.2017.00243

Wang, X., Chen, Y., and Ding, M. (2007). Testing for statistical significance in bispectra: a surrogate data approach and application to neuroscience. *IEEE Trans. Biomed. Eng.* 54, 1974–1982. doi: 10.1109/TBME.2007.895751

Womelsdorf, T., Fries, P., Mitra, P. P., and Desimone, R. (2006). Gamma-band synchronization in visual cortex predicts speed of change detection. *Nature* 439, 733–736. doi: 10.1038/nature04258

Zareian, B., Daliri, M. R., Maboudi, K., Abrishami Moghaddam, H., Treue, S., and Esghaei, M. (2018). Attention enhances LFP phase coherence in macaque visual cortex, improving sensory processing. [preprint]. *bioRxiv* doi: 10.1101/499756

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.

Copyright © 2020 Khamechian and Daliri. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 08 July 2020
doi: 10.3389/fncom.2020.00060

A Model of Memory Linking Time to Space

Hubert Löffler[1]* *and Daya Shankar Gupta*[2]

[1] Independent Researcher, Bregenz, Austria, [2] Biology Department, Camden County College, Blackwood, NJ, United States

The storage of temporally precise spike patterns can be realized by a single neuron. A spiking neural network (SNN) model is utilized to demonstrate the ability to precisely recall a spike pattern after presenting a single input. We show by using a simulation study that the temporal properties of input patterns can be transformed into spatial patterns of local dendritic spikes. The localization of time-points of spikes is facilitated by phase-shift of the subthreshold membrane potential oscillations (SMO) in the dendritic branches, which modifies their excitability. In reference to the points in time of the arriving input, the dendritic spikes are triggered in different branches. To store spatially distributed patterns, two unsupervised learning mechanisms are utilized. Either synaptic weights to the branches, spatial representation of the temporal input pattern, are enhanced by spike-timing-dependent plasticity (STDP) or the oscillation power of SMOs in spiking branches is increased by dendritic spikes. For retrieval, spike bursts activate stored spatiotemporal patterns in dendritic branches, which reactivate the original somatic spike patterns. The simulation of the prototypical model demonstrates the principle, how linking time to space enables the storage of temporal features of an input. Plausibility, advantages, and some variations of the proposed model are also discussed.

Keywords: neural memory, temporally precise spike trains, subthreshold membrane potential oscillations, phase coding, gamma-theta code, working memory, spiking neural networks

OPEN ACCESS

Edited by:
Joaquín J. Torres,
University of Granada, Spain

Reviewed by:
Ausra Saudargiene,
Lithuanian University of Health Sciences, Lithuania
Hans Albert Braun,
University of Marburg, Germany

***Correspondence:**
Hubert Löffler
loeffler.hubert@outlook.com

Received: *15 October 2019*
Accepted: *26 May 2020*
Published: *08 July 2020*

Citation:
Löffler H and Gupta DS (2020) A Model of Memory Linking Time to Space.
Front. Comput. Neurosci. 14:60.
doi: 10.3389/fncom.2020.00060

INTRODUCTION

In daily life, we can distinguish between temporal and spatial properties of our world. In the brain, the temporal as well as spatial properties of the world is largely represented by spatiotemporal patterns of neural spikes. However, an important question currently facing scientists is: How the temporal dimension of the physical world is represented in the brain? The representation of time-dimension is required for a successful interaction of the brain with the four-dimensional physical world (Gupta and Merchant, 2017).

All hitherto presented ideas use spatial properties related to temporal ones even if their relationship is not directly addressed. Larson et al. (2010) explicitly spread time components into spatial components of an input as they investigated the question how sensory systems recognize time varying stimuli by spiking activity. Their model consisted of a succession of end-to-end excitatory neurons (neuronal chain) in combination with STDP to preserve the temporal features of spike patterns via their spatial distribution. In the neuronal chain model, the sensory input activates the first neuron in a chain of neurons, following which the neighboring neurons were activated sequentially with a delay of 2 ms.

Several past studies of spatiotemporal patterns are based on the processing the temporal features of an input. Many of these studies address the association between a precise input spike train

and a desired output spike train in a temporally specific manner (Gütig and Sompolinsky, 2006; Ponulak and Kasinski, 2010; Florian, 2012; Mohemmed et al., 2012; Sporea and Grüning, 2013; Memmesheimer et al., 2014; Albers et al., 2015; Guo et al., 2015). These studies (known as ReSuMe—remote supervised method) assume that the temporal features of the input come from distinct neurons. The localization of the input time-points at different neurons is a precondition.

Data from several recent publications (Düzel et al., 2010; Lisman and Jensen, 2013; Bosman et al., 2014; Axmacher, 2016; Häusser et al., 2016) have provided some empirical evidence that memory for sequences of events is supported by the precise timing of item-related gamma activity with respect to underlying theta oscillations of membrane or field potentials. This neural interaction, referred as "cross-frequency coupling," between lower (e.g., in the theta range) and higher (e.g., in the gamma range) frequency oscillations is also involved in sensory, motor and cognitive brain processes (e.g., Masquelier et al., 2008; Lisman and Jensen, 2013; Gupta and Chen, 2016; Maris and Fries, 2016). Furthermore, Remme et al. (2009) showed that dendritic potential oscillations enable dendritic inputs to be globally integrated on spatiotemporal scale, which can help to control somatic spike (action potential). Empirical data have also shown that oscillations and their combinations play an important role in neural memory of temporal events (Düzel et al., 2010; Headley and Paré, 2017). Thus, it is noteworthy that a dominant network pattern in the hippocampus is a slow oscillation in theta-alpha frequency band (Buzsáki, 2002).

There is deepening interest in the temporal processing of information in auditory system involving interaction between slow and fast oscillations. Different oscillation frequencies, corresponding to different pitches of auditory inputs, generate different spatial patterns in the auditory brain stem chopper neurons (Schreiner and Langer, 1988; Bahmer and Langner, 2006; Bahmer and Gupta, 2018). Modeling studies further suggest that chopper neurons are involved in the transformation of a temporal pitch code into a place code (Wiegrebe and Meddis, 2004). The topographic organization of temporal response characteristics in the auditory system suggests that the transformation of temporal properties, namely frequencies, into spatial patterns is beneficial in implementing neural code of pitch and harmony (Langer, 2015).

On the basis of the above research, the question arises how oscillations enable the transformation of temporal features into spatial ones.

Especially the interaction between two frequencies of oscillations enables a localization process, resulting in time-points according to phase-shifts. Phase coding refers to the process of encoding spike timing in relation to the oscillation phase of SMOs and has been empirically established (Nadasdy, 2009, 2010; Lundqvist et al., 2011; Hasselmo and Stern, 2014; Maris and Fries, 2016).

In fact, there is growing evidence for phase-shifted oscillations in neuronal units or ensembles. Sinha and Narayanan (2015) have shown that the differences in spike phase, due to modulation in either ionic channels or the synaptic conductance within the same neuron may be significant and vary by much as by 100°. In this study, the phases varied considerably as a function of radial distance from the soma, enabling spatial localization. In another study, Stiefel et al. (2010) reported that inhibitory postsynaptic potentials in cortical neurons can considerably shift the oscillatory phase. Cholinergic modulations change the power of oscillation as well as the magnitude of phase shifts. At least two distinct types of models of network activity have been proposed: intrinsic resonance property-based models and circuit-based models (Lee et al., 2018). The modulations can be caused by intrinsic SMOs or [e.g., by a rhythmic inhibition Fries, 2005, 2009, 2015]. Theoretical work, by simulating dendritic oscillations as weakly coupled oscillators with cable, shows that stable phase differences can be maintained between SMOs at different dendritic branches (Remme et al., 2009).

Based on phase differences between SMOs in different dendritic branches, in the following section a model is described linking time to space by encoding the temporal pattern of a spike train from an input neuron into a spatial pattern of a memory neuron, where it is stored by STDP, thus allowing recall and retransformation from spatial into temporal pattern, everything happening in a single neuron following a single input train.

In the next chapter the structure of the model is described, illustrating the mechanisms of encoding, learning and recall, also showing the main equations, which underlie the simulations. The Results section presents detailed simulation data with particular focus on the alterations of dendritic and somatic processes, again during encoding, learning, and recall. The biological plausibility of the model and diverse specific points will be discussed, followed by a conclusion.

MATERIALS AND METHODS

The prototypical model—as outlined below—consists of an input and an output-level (**Figure 1**, inserts—green and orange). The input level comprises two neurons, the input neuron I and the attention neuron A. The activity of neuron A is considered a correlate of "attention" serving storage and recall. The output level consists of a single neuron, the extension neuron which comprises several dendritic branches with the particular feature that they all exhibit subthreshold membrane potential oscillations of the same oscillation period but of different phases.

Spike activities of I and A (**Figure 1**, left) are assumed to be synchronized by LFPs in the gamma band range (Rodriguez et al., 1999). Neuron I fires randomly but temporally precisely aligned to the spikes of the regular firing neuron A. Here their common basic rhythm is 100 Hz. The dendritic SMOs are in the theta band range (here 8.33 Hz, see inset of **Figure 1**, right, bottom). These values are in a biologically plausible range. Gamma frequency in our model can vary between 100 and 30 Hz and theta frequency can vary between 10 and 5 Hz without affecting the principle simulation outcomes. Only the number of storable time-points (12 in the example of **Figure 1** simulations) depends on the relationship of theta to gamma frequencies. Low theta oscillations

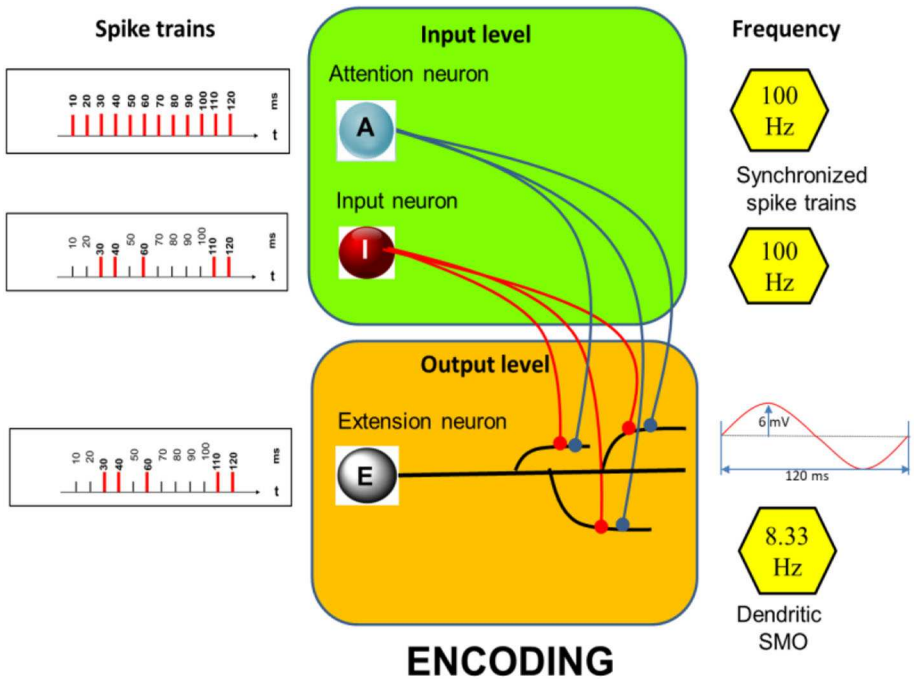

FIGURE 1 | Encoding: Structure of the model with examples of spike trains **(Left)** aligned to a basic rhythm of 100 Hz (indicated on the right) including an example of a dendritic SMO of 8.33 Hz **(Right, Bottom)**. Here only three dendritic branches are shown while the simulations are run with 12 branches, each exhibiting subthreshold oscillations of the same frequency but different phases (which are equally distributed across the oscillation period). Neuron A and I are synaptically connected (filled circles) to all 12 dendrites.

combined with high gamma oscillations allow most spikes to be stored.

Encoding

The core of the presented memory model is the time to space extension process during encoding (**Figure 2**). It takes place in neuron E. Dendritic branches of E represent the spatial dimension. The transformation of input times into a dendritic location depends on the coupling of the gamma input-frequency with the theta SMO-frequency in the branches of neuron E. The dendritic branches serve as the coincidence detectors. Axons of neuron I and A connect to all branches of E and propagate spikes to them. Only a specific combination of synchronized inputs from I and A generates dendritic spikes in E. Additionally, a third component is necessary for a dendritic spike generation. The intrinsic sinusoidal SMO of a branch must be near the peak so that it is sufficiently excitable. Only the coincident inputs from A and I at the peak of dendritic membrane excitability at a branch lead to a dendritic spike.

Finally, the soma of neuron E accumulates the scaled potentials of dendritic spikes with scaled potentials of arriving inputs from I and A. All three membrane potential changes sum together to produce a somatic spike in neuron E. It is believed that the impact of dendritic oscillations on the soma is balanced out because of the distribution of phases between branches. Therefore, the effect of dendritic oscillations on somatic membrane potential are excluded.

Learning

For the storage of inputs, a learning mechanism is required. In the proposed model, we consider two "unsupervised" mechanisms for learning, occurring whenever a dendritic spike in E is generated due to the coincidence of spikes from neurons I and A with a depolarization maximum of SMOs in the dendrites of extension neuron E. Gamma-range coincidence detection corresponds to the findings of Das et al. (2017), Das and Narayanan (2017) in hippocampal CA1 neurons, enabling to decode synchronous gamma-frequency inputs. Storage and recall can be achieved either by STDP and also by activity dependent alterations of the amplitude of the SMOs. For simplicity, the different dendrites of neuron E, are considered as functionally independent units not interfering with each other and not being affected by somatic spikes although each spike in any one of the dendrites generates a spike in the neuron's soma.

Learning by Spike Timing Dependent Plasticity (STDP)

The first strategy to store the spike pattern from input neuron I is based on the enhancement of synaptic weights. The choice of the initial synaptic weight is arbitrary, but it is interrelated to other neuronal parameters.

The synaptic weights between input neuron I and the dendrites of neuron E, set to $W(IE) = 0.15$, are kept constant, unaffected by the learning process. The relevant changes have to take place at the synapses from neuron A to neuron E because

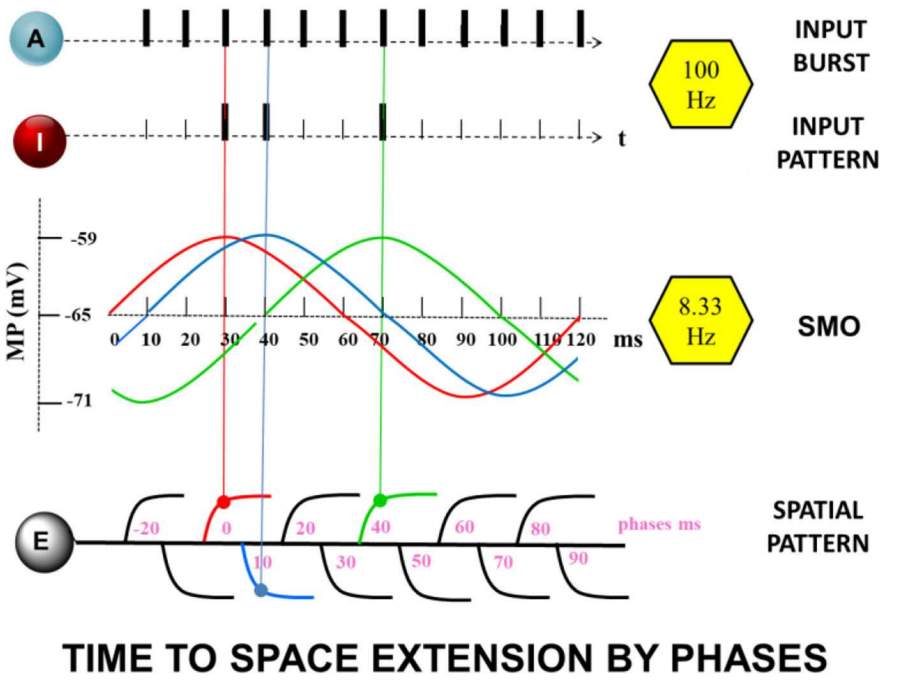

FIGURE 2 | The "Time to Space Extension Process" A continuous spike train of the attention neuron A together with randomly appearing, although precisely aligned spikes of the input neuron I, coincide with maximum values of subthreshold dendritic oscillations of the extension neuron E at different dendrites (same color as the oscillations) leading to enhanced oscillations or increased synaptic efficacy (STDP).

recall will be initiated via this pathway when the input neuron will be silent. Hence, the EPSP amplitudes from neuron A have to be sufficiently increased to compensate for the lack of the EPSPs from input neuron I.

The synaptic weights from neuron A to all branches of neuron E initially amount to $W(AE)_{ini} = 0.40$. In this model a single spike increases the synaptic weight to $W(AE)_{max} = 0.55$, (i.e., by about 37% of its initial value). This will increase further EPSP amplitudes which shall be sufficient to store the information of the timing of a spike input from neuron I in the corresponding dendrite of neuron E for later recall by neuron A. Such a potentiation seems biologically plausible. Remy and Spruston (2007) used a single burst of only five spikes at 100 Hz as stimulation and showed that hippocampal synapses potentiated robustly under this condition. LTP was triggered whenever there were dendritic spikes. Amplitudes of EPSPs could be robustly potentiated by 66% (from ∼7 to ∼12 mV).

Single somatic spikes do not significantly change the synaptic weights at dendrites. Since the proposed model is a single trial, involving a single spike burst, the propagation of somatic action potential will have a negligible effect on LTP.

Learning by SMO Amplification

In an alternative learning mechanism, the encoding occurs due to the enhancement of the amplitude of the SMO after dendritic spikes. This is consistent with the enhancement of amplitude of theta oscillations from cortical EEG recording in a working memory task in humans (Raghavachari et al., 2001) as well as with the findings that neuronal activities in the hippocampus change with individual theta phase in monkeys (Skaggs et al., 1996). Moreover, significant enhancement of oscillatory power observed during encoding has predicted subsequent recall. This effect has been found predominantly in the 4–8 Hz (theta) and 28–64 Hz (gamma) frequency bands (Sederberg et al., 2003). Ness et al. (2016) showed that local field potentials could indeed be utilized to characterize the properties and cellular distributions of active conductance.

In our model the SMO amplitudes are initially set to 6.0 mV. In those branches of neuron E in which a dendritic spike is generated and these values are enhanced by 50% to 9.0 mV which, again, is sufficient to store the information of spike times arriving from neuron I in spatially distributed branches of neuron E for later recall by neuron A.

Recall

Both learning mechanisms allow the encoding and recall of the input spike train. Two out of three neurons, A and E, play a role in the recall process. Recall is triggered by a continuous spike train from A to E, synchronized at the same gamma frequency (100 Hz) as used during the encoding. During this recall, the input neuron is silent. The impact of the input neuron during encoding is substituted by enhanced synaptic weights from neuron A. Furthermore, during recall, the dendritic branches fire at time-points corresponding to the input spike train that was initially encoded. Persistence of phase-shifted SMOs in the dendritic branches together

with enhanced synaptic weights would then result in the recall spikes at dendritic branches. Dendritic spikes generate somatic spikes in E matching the output during the encoding process. It should be noted that the recalled spike train evokes the somatic spike train, generated by the input spike train, not the input spike train itself. The input spike train is recalled only, if it reproduces an equivalent somatic spike train during encoding.

Simulation

A major goal of this study is to show that a single neuron would be able to store temporal properties of an input by the spatial pattern of dendritic branch activation. The current simulation study uses simple mechanisms, such as theta-gamma-oscillations combined with LTP, which is implemented in a small SNN.

Gamma oscillations are represented in the rhythm of spike inputs. Theta oscillations are explicitly modeled by phase-shifted sine functions f(OP):

$$f'(OP) = h{*}sin\,(0.002{*}\pi{*}fq{*}(t - j{*}ph)) \quad (1)$$

where $h = 6.0$ mV is the oscillation amplitude, fq = 8.33 is the oscillation frequency, t is the time in ms and ph = 10 ms is the steps size of phases, multiplied by $j = -2$ to 10.

Referring to the model of Legenstein and Maas (2011), each spike input to the dendrites of the extension neurons causes an EPSP modeled in the form of an alpha function:

$$\mathrm{f\,(EPSP)} \;=\; k{*}W{*}\delta t{*}g^{-\frac{\delta t}{\tau}} \quad (2)$$

where $k = 39$, $g = 2.0$, and $\tau = 1$ are constants and W is the synaptic weight (either from I to E or from A to E: the initial $W(IE)_{ini} = 0.15$ and initial $W(AE)_{ini} = 0.40$. δt is the time difference to the EPSP onset.

SMOs and EPSPs sum up, which generates a dendritic spike whenever a threshold $\varphi^{dend} = -48.7$ mV is reached. The branch spike potential (BsP) is again modeled in form of an alpha function:

$$\mathrm{f\,(BsP)} \;= k{*}\delta t{*}g^{-\frac{\delta t}{\tau}} \quad (3)$$

where $k = 40$, $g = 2.0$, and $\tau = 1$ are constants with δt as the time difference to the dendritic spike onset.

The total branch potentials are summed up by the SMO potentials, the local EPSPs and spike potentials.

EPSPs and spikes propagate to the soma, scaled by different weighting factors, assuming active conduction of dendritic spikes without attenuation ($u_B = 1.0$) and passive, electrotonic propagation of EPSPs with decay to $u_{pass} = 0.08$.

Alterations of somatic membrane potential MP introduced by an individual dendritic branch are calculated by:

$$\mathrm{MP} = \mathrm{EPSP}{*}u_{pass} + \mathrm{BsP}{*}u_B \quad (4)$$

Time delays and different distances of the different branches to the soma are not considered. Dendritic SMOs potentials are without any effects on the soma.

Somatic spikes emerge if the somatic potential exceeds the somatic spike threshold S and are represented as abrupt increase to +30 mV for 1 ms followed by a transient hyperpolarization with exponential reduction. Input spikes, dendritic and somatic spikes are represented by bars as shown in the summarizing **Figure 3**.

The simulations are calculated with time resolution of 1 ms.

The simulation program was written by the author (HL) in Python 3.7 (**Supplementary Materials** - **Datasheet 1** and **Table S2**).

Parameters for somatic, dendritic, synaptic and oscillation properties are shown in **Table S1**.

RESULTS

The model has been tested with input trains of random numbers of input spikes at randomized times, however, adjusted to the 100 Hz gamma rhythm of the spiking of the attention neuron. Test runs have been made with spike inputs and recall during a full cycle period of the dendritic SMOs (120 ms). In this period the attention neuron fires a continuous sequence of 12 spikes.

For encoding and learning, additionally a randomly chosen number of up to 11 spikes from the input neuron arrive at the dendrites of the extension neuron. These input spikes are synchronized with the spike times of the attention neuron but appear at random positions. In any case, due to the above described mechanisms, (i.e., superposition of EPSPs from the input neuron and the attention neuron in correct phase with the SMOs at the dendrites), (i.e., in their maximum, the input patterns also appears at the extension neuron).

Recall is initialized by the application of a spike burst from the attention neuron A while the input neuron I is silent. Due to the learning mechanism, (i.e., increased EPSPs by STDP or enhanced SMO amplitudes, the same pattern as during the input phase reappears at the extension neuron).

This was confirmed by numerous (100) trials with randomized input patterns each showing an exact reproduction of the initial input pattern in the extension neuron after learning and recall. An example is shown in **Figure 3**.

Details of functionally relevant alterations of dendritic and somatic mechanisms during encoding and learning and how these are enabling correct recall are shown in the following sections.

Encoding

Dendritic processes: The net membrane potential of the branch E2 is the sum of the intrinsic potentials of SMO and EPSPs from local synaptic activation by neuron I and neuron A, eventually superimposed by a dendritic spike. The time course of all these potentials is shown in **Figure 4A** with an example from the synaptic branch E2 in which a dendritic spike is generated. This happens only in the maximum of the SMO where the sum of EPSP is strong enough to reach the threshold $\varphi^{dend} = -48.7$ mV for spike generation.

In other branches, dendritic spikes will not be generated at this time, even when the EPSPs are the same. The reason is that potential of the phase shifted oscillations is too low. Spikes in other branches will be generated when their oscillations

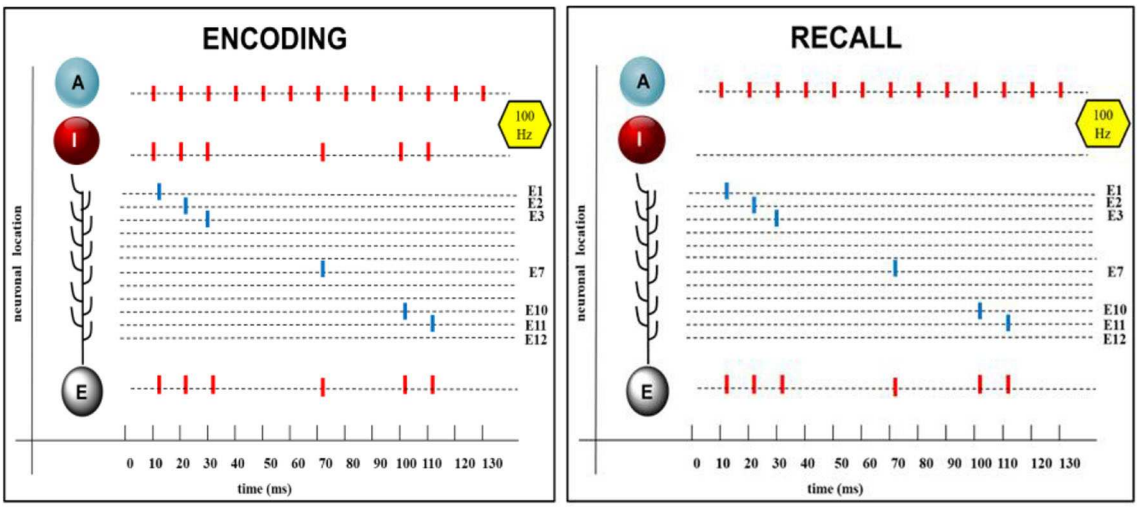

FIGURE 3 | Encoding and recall of a randomized spike pattern. From **Top** to **Bottom**: A: Continuous spike sequences of the attention neuron, which are the same during encoding and recall. I: randomized spike pattern of the input neuron, shown here with six spikes aligned to the spikes of the attention neuron. Single dendritic spikes at different synaptic branches (E1 to E12), appearing during encoding and also on recall, in both cases, summing up to produce a spike train in the soma of the extension neuron E, which is the same as in the input neuron. Short time delays (2 ms) are introduced between onset of the input spike and of the action potential in E.

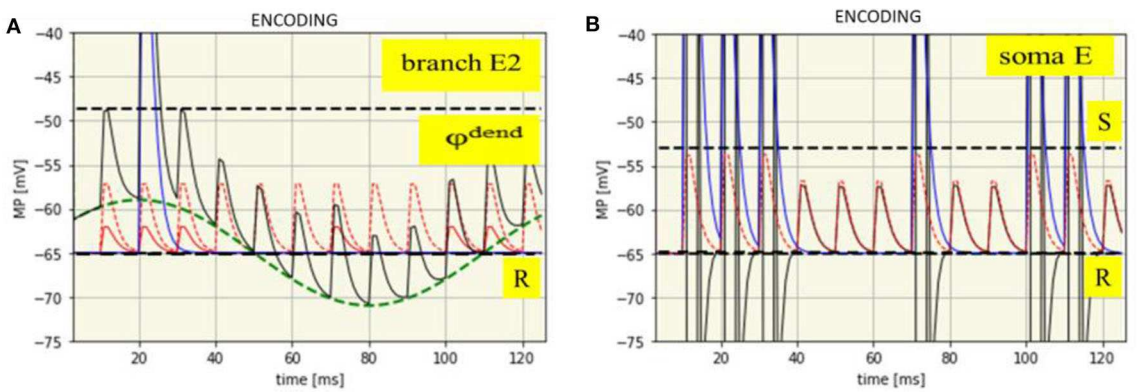

FIGURE 4 | Time course of membrane potentials during encoding of the spike pattern IP1: dendritic in branch E2 **(A)** and somatic **(B)**. R: Resting membrane potential (-65 mV, black dashed line), φ^{dend}: Threshold for dendritic spikes (-48,7 mV, black dashed line). Green dashed line: SMO (phase: -10 ms). Red dashed line: EPSPs from neuron A. Red solid line: EPSPs from neuron I. Blue solid line: Dendritic spike potential appearing once at t = 21 ms. Thin black line: Course of the net branch membrane potential.

are in their maxima, provided that there is again spike input from both neurons at the input level, (i.e., when in addition to the attention neuron the input neuron is also firing). In this example, already shown in **Figure 3**, this happens six times, always in the maximum of the oscillations. In different braches spikes are generated at different time-points due to their phase shifted oscillations.

Somatic processes: The dendritic spikes propagate to the soma together with EPSPs from all branches. However, as it needs the combination of both to exceed the threshold of spike generation the same spike pattern will appear at the extension neuron as received from the input neuron (**Figure 4B**).

Learning

Learning is implemented at the extension neuron either by enhanced EPSPs from the attention neuron or by increased oscillation amplitudes of the dendritic branches (**Figure 5**). These alterations, however, only appear at those dendrites in which a dendritic spike has been generated during encoding by synchronized input from A and I at the maxima of the dendritic SMO. In this way, the information about the temporal pattern of the previous spike input is stored, spatially distributed, in dendritic branches with accordingly altered properties to allow recall.

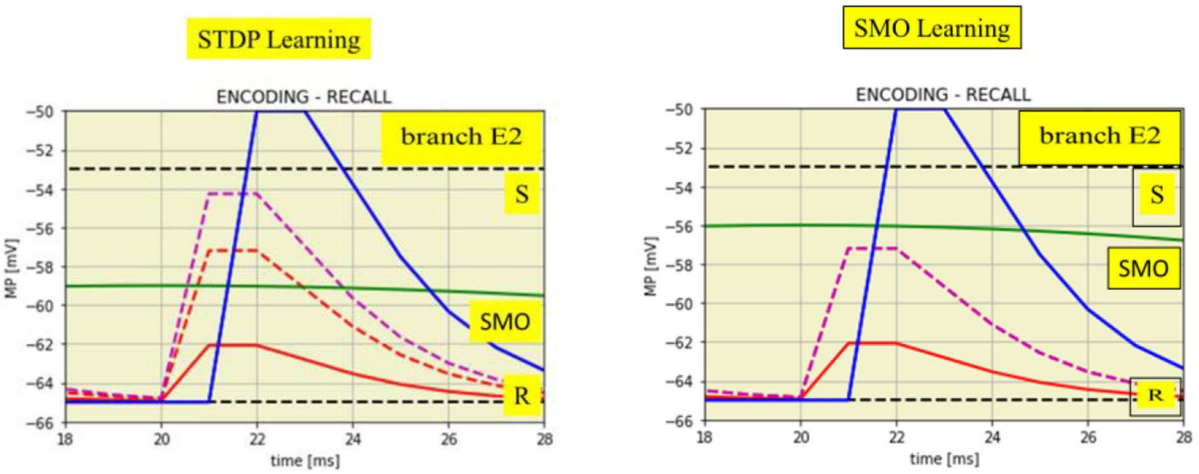

FIGURE 5 | Time expanded cutouts of **Figure 4A** (from $t = 18$ to $t = 28$ ms) also illustrating two different learning mechanisms. **Left**: STDP by enhanced EPSPs on spike input from the attention neuron. EPSPs from the input neuron (two red lines) change to the enhanced EPSP from attention neuron (magenta dashed line). **Right**: increased SMO amplitude, from 6.0 to 9.0 mV, shown by the green line in comparison to the green line in the left figure.

Recall

The previous spike pattern from the input neuron can be recalled by a burst of spikes from the attention neuron while the input neuron is silent. The attention neuron thereby has to fire at the same gamma rhythm as during encoding and the spike burst should have in minimum a duration of the length of the theta cycle over which the maxima of the SMOs are distributed. This guarantees that each dendrite will receive a spike input from the attention neuron when its SMO is close its maximum. In this case, a single spike from the attention neuron can elicit again a dendritic spike, however, only in those branches in which the learning mechanism, as described before, have either increased the oscillation amplitude or enhanced the synaptic efficacy due to enhanced EPSPs. The one like the other is sufficient to compensate for the lack of spikes from the input neuron. Dendritic branches without these learning effects will remain subthreshold even when a spike from the attention neuron input hits the oscillation maxima.

Figure 6A shows how a dendritic spike is generated only by an EPSP from the attention neuron, in this example due to enhanced synaptic efficacy. This example is again drawn for branch E2 in which learning mechanisms have been introduced as illustrated in **Figure 5**. The same happens at all other branches with enhanced EPSPs or increased oscillation amplitudes but not in those branches without learning effects. The spikes will be generated in the same sequence as the learning mechanisms have been introduced in the different braches by spikes from the input neuron.

Nothing has changed with spike generation at the soma due to the propagation of dendritic EPSPs and spikes. Propagation of dendritic EPSPs alone cannot generate somatic spikes. This needs the contribution of dendritic spikes. Hence, dendritic spikes, together with dendritic EPSP will propagate to the neuron's soma, there generating a sequence of spikes which exactly reflects the previous input pattern (**Figure 6B**), just recalled by a gamma spike burst from the attention neuron.

DISCUSSION

The simulation of the current model demonstrates how temporally precise spike trains can be stored by the spatial pattern of dendritic branches even in a single neuron. This mechanism can precisely store temporal information about onetime presented single spikes via the interaction of neural oscillations. The spike trains can be just recalled by a continuous spike burst. The key feature of this model is the spatial localization of spike-timing, which is established by phase-shifted theta oscillations of excitability in dendritic branches in combination with gamma-aligned input patterns. This mechanism links the temporal order of spikes of the input to different dendritic branches, allowing the transformation of temporal properties into the spatial pattern of dendritic activation. The spatial localization of the timing information enables the storage of temporal properties by learning mechanisms.

Other Approaches to Realize Linking Time to Space

Our prototypical model illustrates within a single neuron the general mechanism of transformation of time into spatial dimensions using cross frequency coupling of neuronal oscillations. The same approach can be adapted to ensembles of neurons. Analogous to the dendrites of a single neuron, an ensemble of neurons can work together as canonical microcircuits based on same principles. However, instead of the SMOs of dendrites, the somatic SMOs of ensembles vary in phases. The somatic spikes of single neurons correspond to the dendritic spikes in our prototypical model. The somatic outputs of the ensemble neurons have to be propagated to an additional

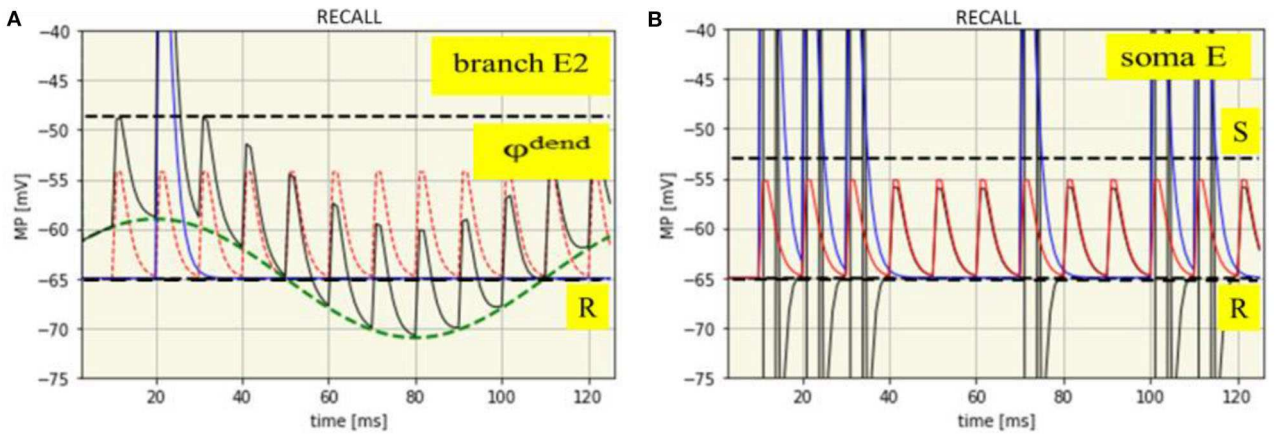

FIGURE 6 | Time course of membrane potentials during recall of the spike pattern IP1: dendritic in branch E2 **(A)** and somatic **(B)**. R: Resting membrane potential (-65 mV, black dashed line), φ^{dend}: Threshold for dendritic spikes (-48,7 mV, black dashed line). Green dashed line: SMO (phase: -10 ms). Red dashed line: EPSPs from neuron A. Blue solid line: Dendritic spike potential appearing once at t = 21 ms. Thin black line: Course of the net branch membrane potential.

output neuron, corresponding to the somatic output of the single neuron. A simulation trial of this alternative model-version has produced identical results. A bursting input by neuron A to the ensemble neurons reproduced an equal temporal pattern in the output neuron as was produced during the encoding process. Future simulation studies could reveal the effect of randomness in the spike trains or various oscillation parameters in neuronal ensembles on the spatiotemporal patterns.

Some further assumptions of our model can be changed without destroying its functionality. For example, the regular distribution of phases of SMO within spatial units can be replaced by random phases, if the number of dendritic branches was enhanced from 13 to 48. Preliminary simulations led to 96% right recalled spikes and 85% completely right recalled spike trains, using 1,000 random input patterns.

Moreover, oscillation-based conversion from temporal to spatiotemporal neural patterns can be also useful in the reverse direction. Since a temporally precise spike train can be represented by a more complex spatiotemporal neural pattern, a spatiotemporal pattern can be propagated to neuronal areas far apart from the origin by the corresponding simpler temporal code. The spatial part of the information can be rebuilt by synchronized oscillations between the area of origin and a target area. In summary, via oscillations spatial information can be propagated by temporal sequences, and these temporal sequences can be efficiently stored by transformation into neural spatial patterns.

Dendrites in Information Processing in Perception

The assumption of dendritic branches acting as independent subunits to process the memory of spike trains appears to be feasible (Golding et al., 2002; Behabady and Mel, 2013; Bono and Clopath, 2017). The independence of dendritic processing allows dendritic spikes to play an important part in information processing since they significantly increase the probability of somatic spikes (Oesch et al., 2005; Polsky et al., 2009). According to the current prototypical model, a coincidence detector, comprising of A neuron along with I neuron (providing inputs from hierarchical sensory areas) and the phase of the SMO when it peaks, is responsible for dendritic spike, which then leads to the learning by STDP. This learning mechanism is responsible for encoding. Furthermore, coincidence detection will occur only at those dendritic branches, where the magnitude of phase shift with respect to the first or a reference dendritic branch is quantitatively equal to an integer multiple of periodicity of continuous input from neuron A. This integer multiple of periodicity of continuous input, called the integration period (Bahmer and Gupta, 2018), encodes the information about the individual spikes in the spike train during the learning stage. STDP, which increases synaptic weights on spatially distributed synapses is responsible for the recall. Continuous spike train, synchronized by the same frequency, which was used during encoding, is responsible for the recall of the initial input from neuron I in this model. Moreover, the learning during the encoding stage will increase the certainty in neural circuits given the knowledge about the source of input that is neuron I.

As discussed previously by Gupta and Bahmer (2019), perception is contributed by both increase in surprisal as well as increase in certainty given the knowledge about sensory object. In the present prototypical model, surprisal would be due to the presence of inhibitory synaptic input at individual dendritic branches. The inhibitory inputs would suppress some of the dendritic spikes, preventing a complete recall. Thus, some of the spikes during the encoding process could be subjected to suppression by inhibitory synapses via a random process given the certainty about sensory object. Since sensory inputs may control the activity of inhibitory neurons in the cortex, the addition of inhibitory synapses could endow the prototypical model the ability to process sensory information for more complex, cognitive functions. Moreover, it is therefore

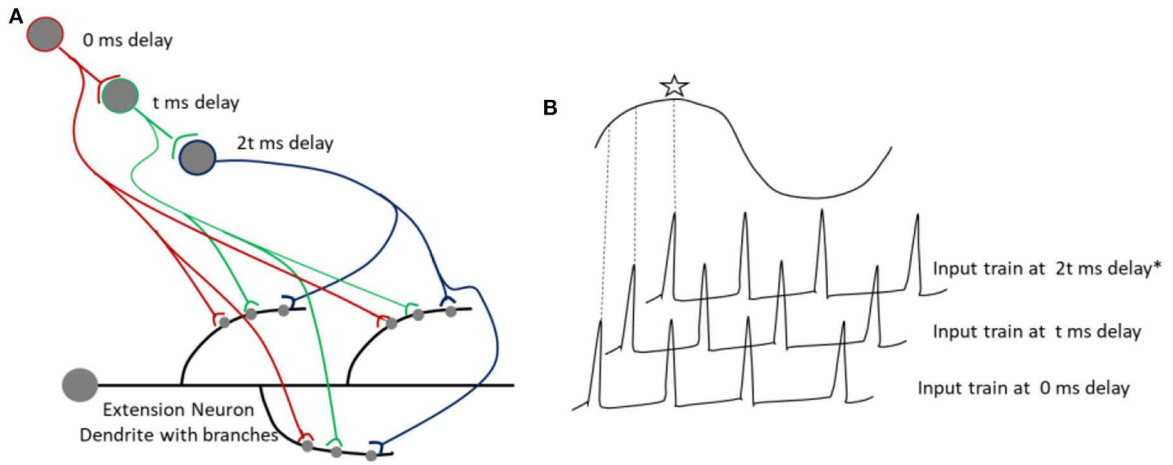

FIGURE 7 | (A) Shows the arrival inputs at dendritic branches, after being processed in parallel circuits with varying number of synapses. The parallel processing allows a single spike to arrive after various delays at the dendritic branches. A synaptic delay (*t*) can vary from 0.5 to few ms. **(B)** Shows a spike coinciding with the peak of a SMO, indicated with an asterix.

noteworthy that inhibitory synapses are shown to play key roles in cortical information processing.

During information processing in perception, spike trains can be produced by a single spike processed in parallel, which would arrive at multiple synapses between neurons I and E after various delays. Thus, the results of parallel processing can be encoded and recalled as a pattern of activation of dendritic branches. One such plausible mechanism is illustrated in **Figure 7**, wherein a single spike, processed in parallel, arrives at multiple synapses after various delays, which coincides with the peak of SMOs in a specific set of branches of dendrites, resulting in a spatial pattern. Processing of a single input in parallel circuits, as shown in **Figure 7B**, will result in multiple outputs, which will arrive at different dendritic branches after various delays, caused by synaptic transmission delay. Also note that a synaptic delay can vary between 0.5 and 5 ms. A spike is encoded as a spatial pattern, constructed by STDP induced at different synapses in a particular set of dendritic branches. During a recall, the same specific input as a result of parallel processing, coinciding with the SMO of dendritic branches, would be responsible for reproducing specific dendritic spikes. Thus, this proposed mechanism within the framework of current prototypical model can explain how a complex information about a simple stimulus (a single spike) can be temporarily stored in specific areas of the brain, where a specific parallel circuit configuration may be available.

Duration of Storable Spike Trains

Short spike trains lasting up to 120 ms as used by our simulation may include the reaction time and reaction intensity of a new input without an adaption process. A memory of this initial part of information is relevant for the further information processing. However, the length of spike trains that can be stored by the proposed model is not limited on the length of a single theta oscillation cycle. This restriction pertains only if one uses a single neuron as bursting input during encoding (in our model by neuron A). Nevertheless, the additional part of a longer lasting input spike train can be stored by a second bursting neuron via its connections to the same (or a new) extension neuron. A temporal driven sequence of bursting neurons [e.g., by time cells (Eichenbaum, 2014) can encode and recall any long input spike train]. Moreover, using lower SMO- frequencies (e.g., delta) the duration of storable spike trains increases. If, for instance, the theta SMOs of our model are replaced by low delta (1 Hz) SMOs and the gamma aligned spike trains are replaced by low beta aligned (12 Hz) trains of about 10 spikes within one second can be exactly stored. Such delta-beta frequency coupling underlying sleep spindles is often associated with memory consolidation.

Working Memory

The mechanisms of the model enable a single trial encoding of temporally precise spike trains by single neurons and their fast and simple retrieval via a non-specific input of spike bursts. These are requirements holding for a working memory, too. However, the working memory has been believed to be established by sustained neuronal spiking, triggered by external events. Yet, there is no necessity for sustained spiking activity of an input, if the contents of working memory can be reactivated immediately, which is possible with the proposed prototypical model. Indeed Stokes (2015) suggests, that a delay in mnemonic activity in the prefrontal cortex is not always critical for maintaining the continuity of working memory. It can be re-established when attention is directed to the task-relevant content. Furthermore, Lundqvist et al. (2016) and Fiebig and Lansner (2017) point out that without sustained spiking, energy would be conserved during inactive states and information is not lost when activity is disrupted, and attractors can hold multiple items in working memory. They extended an attractor network model for memory encoding and recall by oscillations. Experimentally, the authors observed gamma bursts for activation and reactivation of inputs. Similarly,

our model proposes gamma bursts of the attention neuron accompanying the input for encoding and gamma bursts again during the recall. Gamma bursting is a general form of the activation patterns of neurons in the central nervous system (Cooper, 2002).

Because in our model the recall is generated by a bursting input from the same neuron (A), which was active throughout the encoding process, a further interesting option presents itself: The continuity of bursting yields original output of the input spike train, even if the input train itself is finished. Thus, the output of the terminated input is replicated as long as the bursting lasts. Even a complex input pattern within a group of neurons can repeat oneself as long as the burst continues. In addition, using SMO-learning instead of LTP, a properly tuned bursting input from any neuronal source can recall the stored input patterns and realize the working memory.

CONCLUSION

Authors suggest that the realization of neuronal memory for temporal events is restricted to distributing temporal properties of events over spatially different units. The ubiquitous brain oscillations combined with the synchronization of neural activities can store temporal information as spatial patterns. Frequency, phase and amplitude—the three main characteristics of oscillations—can work together for this purpose. Significantly, the simulations presented in this study show that neural oscillations can allow time-dimension to be linked with the spatial dimensions in the brain circuits, which is important for the cognitive functions in interacting with the four-dimensional physical world. Moreover, simple spike bursts frequency aligned to the input events can serve as the trigger for retrieval as well as correlate for attention. The utility of the proposed memory mechanism for various brain functions, such as working memory is also evident. Preliminarily, our model is for now a theoretical hypothesis supported by computational simulations using physiologically plausible range of parameters. As a prototype model, it will be further enhanced by additional biological findings.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

HL and DG contributed conception and design of the study. HL simulated the model and wrote the first draft of the manuscript. DG wrote sections of the manuscript. Both authors contributed to manuscript revision, read and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fncom.2020.00060/full#supplementary-material

REFERENCES

Albers, C. H., Westkott, M., and Pawelzik, K. (2015). Learning of precise spike times with homeostatic membrane potential dependent synaptic plasticity. *PLoS ONE* 11:e0148948. doi: 10.1371/journal.pone.0148948

Axmacher, N. (2016). A useful code for sequence. *Nat. Neurosci*. 19, 1276–1277. doi: 10.1038/nn.4391

Bahmer, A., and Gupta, D. S. (2018). Role of oscillations in auditory temporal processing: a general model of temporal processing of sensory information in the brain? *Front. Neurosci.* 12:793. doi: 10.3389/fnins.2018.00793

Bahmer, A., and Langner, G. (2006). Oscillating neurons in the cochlear nucleus: I. experimental basis of a simulation paradigm. *Biol. Cybern.* 95, 371–379. doi: 10.1007/s00422-006-0092-6

Behabady, B. F., and Mel, B. W. (2013). Mechanisms underlying subunit independence in pyramidal neuron dendrites. *Proc. Natl. Acad. Sci. U. S. A. 111,* 498–503. doi: 10.1073/pnas.1217645111

Bono, J., and Clopath, C. (2017). Modeling somatic and dendritic spike mediated plasticity at the single neuron and network level. *Nat. Commun.* 8:706. doi: 10.1038/s41467-017-00740-z

Bosman, C. A., Lansink, C. S., and Pennartz, C. M. (2014). Functions of gamma-band synchronization in cognition: from single circuits to functional diversity across cortical and subcortical systems. *Eur. J. Neurosci*. 39, 1982–1999. doi: 10.1111/ejn.12606

Buzsáki, G. (2002), Theta oscillations in the hippocampus. *Neuron* 33, 325–340. doi: 10.1016/S0896-6273(02)00586-X

Cooper, D. (2002). The significance of action potential bursting in the brain reward circuit. *Neurochem. Int*. 41, 333–340. doi: 10.1016/S0197-0186(02)00068-2

Das, A., and Narayanan, R. (2017). Theta-frequency selectivity in the somatic spike-triggered average of rat hippocampal pyramidal neurons is dependent on HCN channels. *J. Neurophysiol.* 118, 2251–2266. doi: 10.1152/jn.00356.2017

Das, A., Rathour, R. K., and Narayanan, R. (2017). String on a violin: location-dependence of frequency tuning in active dendrites. *Front. Cell. Neurosci. 11*:72. doi: 10.3389/fncel.2017.00072

Düzel, E., Penny, W. D., and Burgess, W. (2010). Brain oscillations and memory. *Curr. Opin. Neurobiol.* 20, 143–149, doi: 10.1016/j.conb.2010.01.004

Eichenbaum, H. (2014). Time cells in the hippocampus: a new dimension for mapping memories. *Nat. Rev. Neurosci.* 15, 732–744. doi: 10.1038/nrn3827

Fiebig, F., and Lansner, A. (2017). A spiking working memory model based on hebbian short-term potentiation, *J. Neurosci.* 37, 83–96. doi: 10.1523/JNEUROSCI.1989-16.2016

Florian, R. V. (2012). The chronotron: a neuron that learns to fire temporally precise spike patterns. *PLoS ONE* 7:e40233. doi: 10.1371/journal.pone.0040233

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

Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu Rev Neurosci*, 32, 209–224. doi: 10.1146/annurev.neuro.051508.135603

Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235. doi: 10.1016/j.neuron.2015.09.034

Golding, N. L., Staff, N. P., and Spruston, N. (2002). Dendritic spikes as a mechanism for cooperative long-term potentiation. *Nature* 418, 326–331. doi: 10.1038/nature00854

Guo, L., Wang, Z., and Adjouadi, M. (2015). A novel biologically plausible supervised learning method for spiking neurons. *Int'l Conf. Artificial Intelligence (ICAI'15)* (At Las Vegas), 578–584.

Gupta, D. S., and Bahmer, A. (2019). Increase in mutual information during interaction with the environment contributes to perception. *Entropy* 21:365. doi: 10.3390/e21040365

Gupta, D. S., and Chen, L. (2016). Brain oscillations in perception, timing and action. *Curr. Opin. Behav. Sci.* 8, 161–166. doi: 10.1016/j.cobeha.2016.02.021

Gupta, D. S., and Merchant, H. (2017). Editorial: understanding the role of the time dimension in the brain information processing. *Front. Psychol.* 8:240. doi: 10.3389/fpsyg.2017.00240

Gütig, R., and Sompolinsky, H. (2006). The tempotron: a neuron that learns spike timing-based decisions. *Nat. Neurosci.* 9, 420–428. doi: 10.1038/nn1643

Hasselmo, M. E., and Stern, C. E. (2014). Theta rhythm and the encoding and retrieval of space and time. *Neuroimage* 85, 656–666. doi: 10.1016/j.neuroimage.2013.06.022

Häusser, A. C., Poeppel, D., Ezzyat, Y., and Davach, L. (2016). Episodic sequence memory is supported by a theta–gamma phase code. *Nat. Neurosci.* 19, 1374–1380. doi: 10.1038/nn.4374

Headley, D. B., and Paré, D. (2017). Common oscillatory mechanisms across multiple memory systems. *NPJ Sci. Learn.* 2:1. doi: 10.1038/s41539-016-0001-2

Langer, G. D. (2015). *The Neural Code of Pitch and Harmony*. Cambridge, UK: Cambridge University Press. doi: 10.1017/CBO9781139050852

Larson, E., Perrone, B. P., Sen, K., and Billimoria, C. P. (2010). A robust and biologically plausible spike pattern recognition network. *J Neurosci.* 30, 15566–15572. doi: 10.1523/JNEUROSCI.3672-10.2010

Lee, S.-H., Urbano, F. J., and Garcia-Rilla, E. (2018). The critical role of intrinsic membrane oscillation. *Neurosignals* 26, 66–76. doi: 10.1159/000493900

Legenstein, R., and Maas, W. (2011). Branch-specific plasticity enables self-organization of nonlinear computation in single neurons. *J. Neurosci.* 31, 10787–10802. doi: 10.1523/JNEUROSCI.5684-10.2011

Lisman, J. E., and Jensen, O. (2013). The theta-gamma neuronal code. *Neuron 77*, 1002–1016. doi: 10.1016/j.neuron.2013.03.007

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.* 10, 3008–3020. doi: 10.1162/jocn_a_00029

Lundqvist, M., Rose, J., Herman, P., Brincat,S. L., Buschman, T. J., Miller, E. K. (2016). Gamma and beta bursts underlie working memory. *Neuron* 90, 152–164. doi: 10.1016/j.neuron.2016.02.028

Maris, E., and Fries, P. (2016). D*iverse phase relations among neuronal rhythms and their potential function. Trends Neurosci.* 39, 86–99. doi: 10.1016/j.tins.2015.12.004

Masquelier, T., Hugues, E., Deco, G., and Thorpe, S. (2008). Oscillations, phase-of-firing coding, and spike timing-dependent plasticity: an efficient learning scheme. *J. Neurosci.* 29, 13484–13493. doi: 10.1523/JNEUROSCI.2207-09.2009

Memmesheimer, R. M., Rubin, R., Olveczky, B. P., and Sompolinsky, H. (2014). Learning precisely timed spikes. *Neuron* 82, 925–938. doi: 10.1016/j.neuron.2014.03.026

Mohemmed, A., Schliebs, S., Matsuda, S., and Kasabov, N. (2012). SPAN: spike pattern association neuron for learning spatiotemporal spike patterns. *Int. J. Neuronal. Syst.* 22:1250012. doi: 10.1142/S0129065712500128

Nadasdy, Z. (2009). Information encoding and reconstruction from the phase of action potentials. *Front. Syst. Neurosci.* 3:6. doi: 10.3389/neuro.06.006.2009

Nadasdy, Z. (2010). Binding by asynchrony: the neuronal phase code *S. Front. Syst. Neurosci.* 4:51. doi: 10.3389/fnins.2010.00051

Ness, T. B., Remme, M. W. H., and Einevoll, G. T. (2016). Active subthreshold dendritic conductances shape the local field potential. *J. Physiol.* 594, 3809–3825. doi: 10.1113/JP272022

Oesch, N., Euler, T., and Taylor, W. R. (2005). Direction-selective dendritic action potentials in rabbit retina. *Neuron* 47, 739–750. doi: 10.1016/j.neuron.2005.06.036

Polsky, A., Mel, B., and Schiller, J. (2009). Encoding and decoding bursts by NMDA spikes in basal dendrites of layer 5 pyramidal neurons. *J. Neurosci*. 29, 11891–11903. doi: 10.1523/JNEUROSCI.5250-08.2009

Ponulak, F., and Kasinski, A. (2010). Supervised learning in spiking neuronal networks with ReSuMe: sequence learning, classification, and spike shifting. *Neuro. Comput.* 22, 467–510. doi: 10.1162/neco.2009.11-08-901

Raghavachari, S., Kahana, M. J., Rizzuto, D. S., Caplan, J. B., Kirschen, M. P., Bourgeois, B., et al. (2001). Gating of human theta oscillations by a working memory task. *J. Neurosci.* 21, 3175–3183. doi: 10.1523/JNEUROSCI.21-09-03175.2001

Remme, M. W. H., Lengyel, M., and Gutkin, B. S. (2009). The role of ongoing dendritic oscillations in single-neuron dynamics. *PLoS Comput. Biol.* 5:e1000493. doi: 10.1371/journal.pcbi.1000493

Remy, S., Spruston, M. (2007). Dendritic spikes induce single-burst long-term potentiation. *Proc. Natl. Acad. Sci. U. S. A.* 104, 17192–17197. doi: 10.1073/pnas.0707919104

Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B., and Varela, F. J. (1999). Perception's shadow: long-distance synchronization of human brain activity. *Nature* 397, 430–433. doi: 10.1038/17120

Schreiner, C. E., and Langer, G. (1988). Periodicity coding in the inferior colliculus of the cat. *II. topographical organization. J. Neurophysiol.* 60:6. doi: 10.1152/jn.1988.60.6.1823

Sederberg, P., Kahana, M., Howard, M., Donner, E., and Madse, R. (2003). Theta and gamma oscillations during encoding predict subsequent recal. *J. Neurosci.* 23, 10809–10814. doi: 10.1523/JNEUROSCI.23-34-10809.2003

Sinha, M., and Narayanan, R. (2015). HCN channels enhance spike phase coherence and regulate the phase of spikes and LFPs in the theta-frequency rang. *Proc. Natl. Acad. Sci. U. S. A.* 112, E2207–E2216. doi: 10.1073/pnas.1419017112

Skaggs, W. E., McNoughton, B. L., Wilson, M. A., and Barnes, C. A. (1996). Theta phase precession in hippocampal neuronal populations and the compression of temporal sequences. *Hippocampus* 6, 149–172. doi: 10.1002/SICI1098-106319966:2149::AID-HIPO63.0.CO;2-K

Sporea, I., and Grüning, A. (2013). Supervised learning in multilayer spiking neuronal networks. *Neuro. Comput.* 25, 473–509. doi: 10.1162/NECO_a_00396

Stiefel, K. M., Fellous, J. M., Thomas, P. J., and Sejnowski, T. J. (2010). Intrinsic subthreshold oscillations extend the influence of inhibitory synaptic inputs on cortical pyramidal neurons. *Eur. J. Neurosci.* 31, 1019–1026. doi: 10.1111/j.1460-9568.2010.07146.x

Stokes, M. G. (2015). "Activity-silent" working memory in prefrontal cortex: a dynamic coding framework. *Trends Cog. Sci.* 19, 394–405 doi: 10.1016/j.tics.2015.05.004

Wiegrebe, L., and Meddis, R. (2004). The representation of periodic sounds in simulated sustained chopper units of the ventral cochlear nucleus. *J. Acoust. Soc. Am.* 116:1207–1218. doi: 10.1121/1.1643359

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.

Copyright © 2020 Löffler and Gupta. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Computational Neuroscience

ORIGINAL RESEARCH
published: 13 July 2020
doi: 10.3389/fncom.2020.00059

Top-Down Control of Inhibitory Granule Cells in the Main Olfactory Bulb Reshapes Neural Dynamics Giving Rise to a Diversity of Computations

Zhen Chen[1†] *and Krishnan Padmanabhan*[2*†]

[1] Department of Brain and Cognitive Sciences, University of Rochester, Rochester, NY, United States, [2] Department of Neuroscience, University of Rochester School of Medicine and Dentistry, Rochester, NY, United States

Growing evidence shows that top-down projections from excitatory neurons in piriform cortex selectively synapse onto local inhibitory granule cells in the main olfactory bulb, effectively gating their own inputs by controlling inhibition. An open question in olfaction is the role this feedback plays in shaping the dynamics of local circuits, and the resultant computational benefits it provides. Using rate models of neuronal firing in a network consisting of excitatory mitral and tufted cells, inhibitory granule cells and top-down piriform cortical neurons, we found that changes in the weight of feedback to inhibitory neurons generated diverse network dynamics and complex transitions between these dynamics. Changes in the weight of top-down feedback supported a number of computations, including both pattern separation and oscillatory synchrony. Additionally, the network could generate gamma oscillations though a mechanism we termed **T**op-down control of **I**nhibitory **N**euron **G**amma (TING). Collectively, these functions arose from a codimension-2 bifurcation in the dynamical system. Our results highlight a key role for this top-down feedback, gating inhibition to facilitate often diametrically different computations.

Keywords: dynamical system, olfactory bulb, oscillations, pattern separation, synchrony, bifurcation, feedback, top-down

OPEN ACCESS

Edited by:
Daya Shankar Gupta, Camden County College, United States

Reviewed by:
Markus Rothermel, RWTH Aachen University, Germany
Christiane Linster, Cornell University, United States

***Correspondence:**
Krishnan Padmanabhan
Krishnan_Padmanabhan@urmc.rochester.edu

†ORCID:
Zhen Chen
orcid.org/0000-0002-5590-3552
Krishnan Padmanabhan
orcid.org/0000-0002-3255-8346

Received: *31 March 2020*
Accepted: *22 May 2020*
Published: *13 July 2020*

Citation:
Chen Z and Padmanabhan K (2020) Top-Down Control of Inhibitory Granule Cells in the Main Olfactory Bulb Reshapes Neural Dynamics Giving Rise to a Diversity of Computations. Front. Comput. Neurosci. 14:59. doi: 10.3389/fncom.2020.00059

INTRODUCTION

Growing evidence suggests that top-down centrifugal feedback from higher cortical areas specifically target inhibitory interneurons in primary sensory regions. In the olfactory system, axons from excitatory neurons in the piriform cortex (PCx) synapse onto the inhibitory granule cells in olfactory bulb (OB), whereby they can modulate the function of the mitral/tufted cells (M/T), the principal relays of olfactory information from the bulb to the brain (Shipley and Adamek, 1984; Boyd et al., 2012; Markopoulos et al., 2012; Oswald and Urban, 2012b; Padmanabhan et al., 2016, 2019). This circuit motif results in piriform cortical neurons receiving input from only excitatory M/T cells but exerting influence on the local circuit dynamics in the OB via inhibitory populations. Consequently, the information relayed to piriform cortical neurons comes from M/T cells, but feedback intervenes in network dynamics through the local inhibitory granule cells. Although this

network motif constitutes major feature of the olfactory system, the computational role of this top-down control of inhibition remains largely unknown.

A number of studies have previously explored the dynamics of M/T cells and granule cells in OB as a two-population network of excitatory (E) and inhibitory (I) neurons (Cleland and Linster, 2005; Brea et al., 2009; Kay et al., 2009; Li and Cleland, 2017) in parallel with a broader literature on excitatory-inhibitory (E-I) networks (Wilson and Cowan, 1972; Ermentrout and Kopell, 1990; Tsodyks et al., 1997; Tiesinga and Sejnowski, 2009; Ledoux and Brunel, 2011; Franci et al., 2018). The studies in olfaction have revealed not only the mechanisms by which these dynamics emerge, but also how changes in the oscillatory power (Nusser et al., 2001) result in alterations in behavior, including in odor discrimination tasks (Abraham et al., 2010). The recent evidence that cortical feedback directly synapses onto inhibitory interneurons (Boyd et al., 2012; Markopoulos et al., 2012) suggests that the local dynamics of excitatory and inhibitory neurons in OB can be gated by centrifugal projections from olfactory cortex. How the local E-I network's activity in the bulb is changed by centrifugal input, what these changes mean more broadly for neural dynamics in the early olfactory system, and the role of these dynamics play in neural computation remains an open question.

To address this, we built a three-node network model consisting of an excitatory population of mitral/tufted cells (M), an inhibitory population of granule cells (G) and a top-down population of pyramidal/semilunar cells (P) in PCx, and studied how firing rate dynamics were influenced by top-down weights onto inhibition. Changing the weight of the top-down connections to local inhibitory neurons reshaped the dynamics of the local E-I circuit in a way that enhanced sensory discrimination as well as generated oscillatory synchrony including entraining gamma oscillations in the local circuit [**T**op-down control of **I**nhibitory **N**euron **G**amma, (TING)]. Finally, the mechanism underlying the dynamics, as well as the functional roles played by top-down control of inhibition occurred via a codimension-2 bifurcation in the dynamical system. By gating the weight of connections from piriform cortex to the inhibitory neurons in the bulb, a number of seemingly disparate computations could be supported by a single circuit, providing an additional framework for the diversity of inhibitory interneuron function in the olfactory bulb.

MATERIALS AND METHODS

Network Model

The network model was composed of three nodes, the local excitatory population, corresponding to mitral and tufted cells (M) and inhibitory population corresponding to granule cells (G) which were reciprocally coupled, and a top-down population corresponding to the principal neurons in piriform cortex (P) that received input from the local M population and projected back to the inhibitory G population (**Figure 1A**). In the model, $r_i(t)$, $i = 1, 2, 3$ represented the firing rates of the three neuron populations, respectively, whose dynamics were determined by Wilson-Cowan equations (Wilson and Cowan, 1972) as follows:

$$\begin{cases} \tau_1 \dot{r_1} = -r_1 + S(w_{11}r_1 + w_{12}r_2 + \mu) \\ \tau_2 \dot{r_2} = -r_2 + S(w_{21}r_1 + w_{22}r_2 + w_{23}r_3) \\ \tau_3 \dot{r_3} = -r_3 + tanh(w_{31}r_1 + w_{33}r_3) \end{cases} \quad (1)$$

where S is the sigmoid function:

$$S(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

which described the non-linear relationship between the mean synaptic input and average firing rate (normalized to a range between 0 and 1). The parameter $\tau_i, i = 1, 2, 3$ was the time constant for each population, characterizing how quickly the dynamics of each population evolved. The mitral/tufted cell population (M) received an external stimulus μ, that represented the only external input to the system. The connection weight from population j to population i was denoted by w_{ij}, $i, j = 1, 2, 3$, among which $w_{11}, w_{21}, w_{31}, w_{23}, w_{33} > 0$ and $w_{12}, w_{22} < 0$. The connection weight $w_{ij}, i, j = 1, 2, 3$ represented the average synaptic input received by the neuron population i from the population j. Throughout this paper, we set the parameters as follows: $w_{11} = 8.7$, $w_{12} = -10$, $w_{21} = 7.0$, $w_{22} = -13$, $w_{31} = 1.5$, $w_{33} = 0.5$. The parameters in the model were chosen based on previous studies of Wilson-Cowan rate model (Ermentrout and Kopell, 1990; Ledoux and Brunel, 2011; Veltz and Sejnowski, 2015), and their relative values were adjusted according to experimentally recorded excitatory and inhibitory postsynaptic inputs of M/T and granule cells in OB (Urban and Sakmann, 2002; Egger et al., 2005; Kapoor and Urban, 2006). For instance, the value of recurrent excitation (w_{11}) was determined by mapping recorded excitatory post-synaptic potentials (EPSP) in M/T cells and the weight of inhibition from granule cells (w_{12}) was determined using similar mappings of whole cell recordings of inhibitory post-synaptic potentials (IPSP) in M/T cells (Urban and Sakmann, 2002). The weights between populations were determined by integrating synaptic potentials with known connectivity densities using trans-synaptic viral tracers that allowed for estimates of the number of pre-synaptic cells for each population (Willhite et al., 2006; Miyamichi et al., 2011; Padmanabhan et al., 2016) to yield the relative weights for instance, $|w_{12}| > |w_{21}|$. Estimates of synaptic weights for the projections to piriform cortex and the feedback connections were estimated from Franks and Isaacson (2006), Suzuki and Bekkers (2011), and Boyd et al. (2012). Time constants, for example the time constant τ_2 for granule cells, were derived from data using both calcium imaging and whole cell recordings across the types of neurons that constituted the populations in our model (Franks and Isaacson, 2006; Kapoor and Urban, 2006; Suzuki and Bekkers, 2011). As there was no inhibitory synaptic input into the piriform pyramidal/semilunar cell population (P), the combination $w_{31}r_1 + w_{33}r_3$ was non-negative due to $w_{31} > 0$ and $w_{33} > 0$. Sigmoid function (2) for the third equation of system (1) would mean that the lowest r_3 value that an equilibrium could reach would be 0.5.

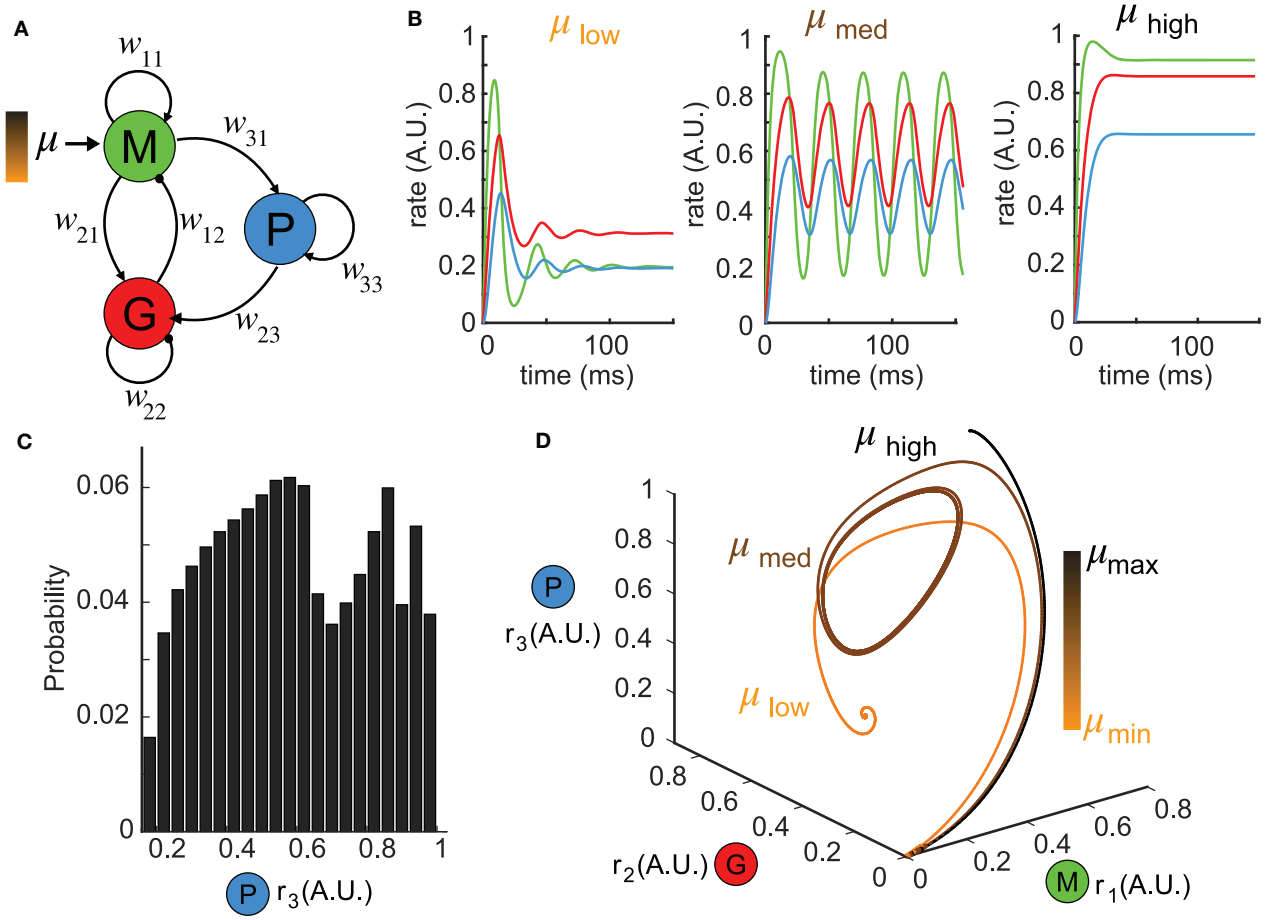

FIGURE 1 | Reduced network model exhibiting complex dynamics. **(A)** A schematic diagram illustrating the topology of the reduced network model. Each node denotes a neuron population and the connection weights are defined by w_{ij} (excitatory: arrows; inhibitory: circles). **(B)** Network responses [color-coded firing rates match population nodes in **(A)**] to different levels of stimulus strength (left: $\mu_{low} = 0.1$, middle: $\mu_{med} = 1.5$, right: $\mu_{high} = 3.0$) with the other parameters fixed. **(C)** A distribution of all possible steady states r_3 over a wide range of parameter choices: $\mu \in [0, 5]$, $w_{31} \in [0, 5]$, $w_{23} \in [0, 30]$ shows the diversity of responses the network can generate. **(D)** Trajectories of the firing rate responses plotted in **(B)** are visualized in the phase space spanned by (r_1, r_2, r_3). The color bar indicates stimulus strengths for three representative levels as in **(B)**.

In this framework, even without an input, at least half of the pyramidal/semilunar cells would keep firing. Thus, we used the hyperbolic tangent function $tanh\,(x)$ in order ensure that piriform cortical cell firing rates were low in the absence of odors (Stettler and Axel, 2009; Davison and Ehlers, 2011), and to exploit the entire range $[0, 1]$ of r_3 (**Figure 1C**). However, it should be noted that such choice of non-linearity did not affect our findings since all bifurcations supporting the computations of the network we identified were found in the system using the sigmoid function (2).

Definition of Period of a Limit Cycle

From the perspective of dynamical system, a limit cycle in the phase space corresponds to the oscillation of firing rates in the temporal space. Since the time constants in our model had units of millisecond, the frequency of oscillations was defined as $1{,}000/T$, with T denoting the period of limit cycles, which was defined as follows: if $r_i\,(t)$, $i = 1, 2, 3$ denote the firing rate of neural population in the model, then a limit cycle satisfies the periodicity:

$$r_i\,(t) = r_i\,(t + T)\,,\ i = 1, 2, 3 \tag{3}$$

for some $T > 0$ and all $t \in \mathbb{R}$. The minimal T for which the above equality holds is the period of the limit cycle.

Metric Definition

As representations of the network (ω-limit sets) could take on different forms, an equilibrium or a limit cycle, we defined two quantitative metrics: $d^E\,(\Omega_1, \Omega_2)$ and $d^S\left(r_3\,|_{\Omega_1}, r_3\,|_{\Omega_2}\right)$ which served to measure the distance between different types of ω-limit sets in responses to any given stimulus pair (μ_1, μ_2) where $\mu_2 = \mu_1 + \Delta$. $d^E\,(\Omega_1, \Omega_2)$ denoted the average Euclidean distance between Ω_1 and Ω_2 in the three-dimensional (3D) phase space of firing rates (r_1, r_2, r_3), and the spectrum distance

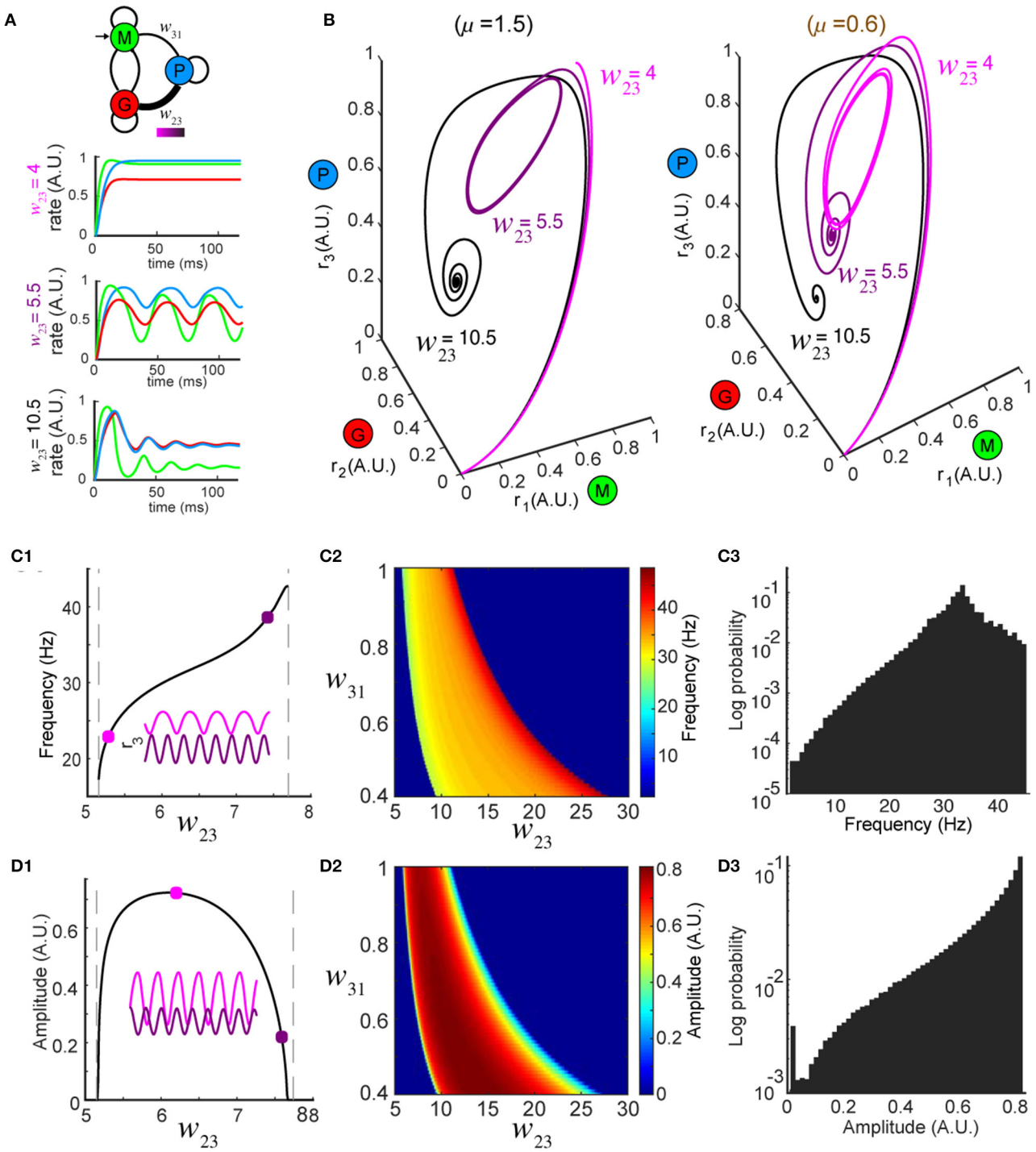

FIGURE 2 | Network dynamics are controlled by top-down modulation. **(A)** Top: changing top-down input w_{23} reshapes network activity to the same stimulus (color bar indicates values of w_{23}); bottom: firing rates at three representative values of w_{23} while the stimulus is held constant ($\mu = 1.5$). **(B)** Trajectories in the phase space for the same w_{23} as in **(A)**; left: oscillations occur around $w_{23} = 5.5$ for $\mu = 1.5$ (same as **A**); right: oscillations occur around $w_{23} = 10.5$ for a different stimulus $\mu = 0.6$, revealing that the dynamics are diverse across different combinations of stimuli and top-down input. **(C)** Modulation of oscillation frequency by top-down input (for $\mu = 1.5$). **(C1)** Dependence of frequency on top-down input w_{23}. Inset: time series of r_3 (t) for two example values of w_{23} (squares). **(C2)** Frequency modulation by top-down input occurred over a range of feedforward drive w_{31}. **(C3)** A distribution of oscillation frequencies that can be generated by the network for all possible combinations of $w_{23} \in [0, 30]$and $\mu \in [0, 5]$. **(D)** Similar to **(C)** but for oscillation amplitudes.

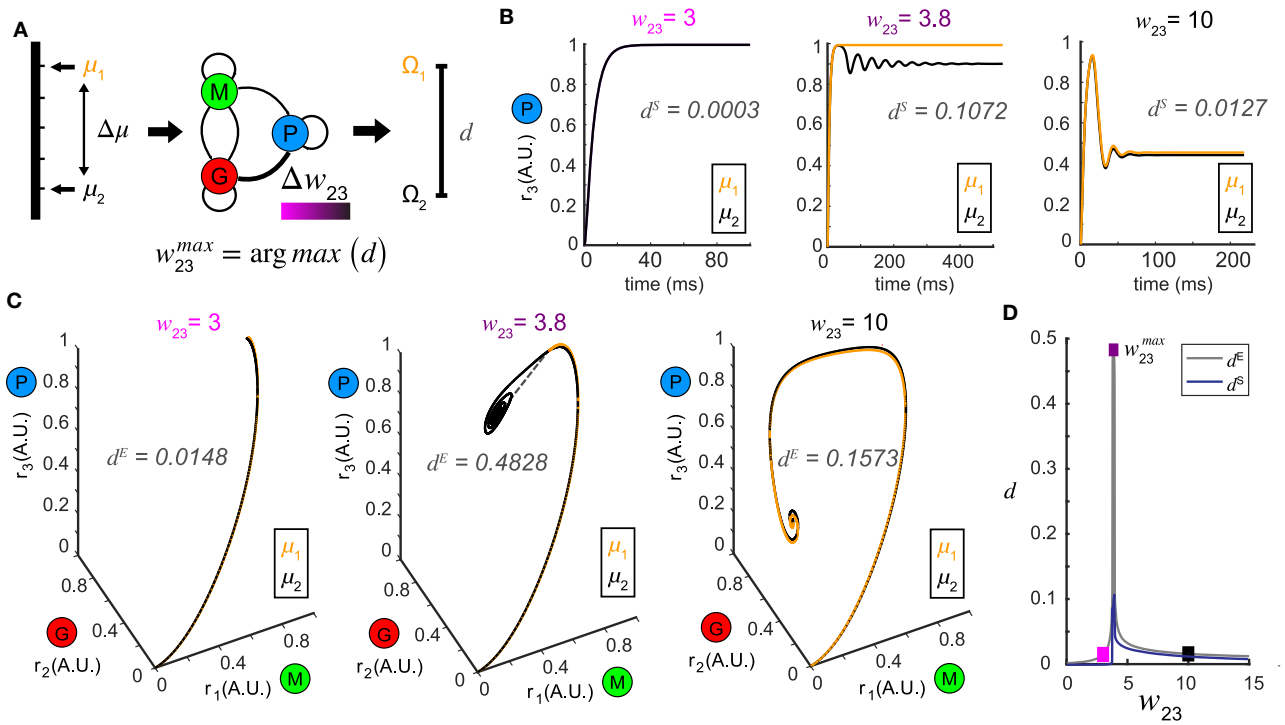

FIGURE 3 | Pattern separation via top-down control. **(A)** A schematic diagram illustrating the separation maximization between the response patterns to a pair of stimuli (μ_1 and μ_2) by changing top-down input w_{23} (color bar indicates values of w_{23}). **(B)** Time series of $r_3\,(t)$ in response to $\mu_1 = 1.0$ and $\mu_2 = 1.1$ at three representative values of w_{23} with the spectrum distance d^S indicated. The two responses $r_3\,(t)$ are close at some top-down inputs (left: $w_{23} = 3$; right: $w_{23} = 10$), but pushed apart at other top-down input (middle: $w_{23} = 3.8$). **(C)** Phase trajectories and network representations of two stimuli from which the Euclidean distance d^E is calculated. From left to right the top-down input w_{23} correspond to those in **(B)** for the same stimulus pair ($\mu_1 = 1.0$ and $\mu_2 = 1.1$). **(D)** Non-monotonic dependence of both d^E and d^S on top-down input w_{23}, with maximum achieved at $w_{23} = 3.8$. The squares are color coded as in **(B,C)**.

$d^S\left(r_3\,|_{\Omega_1}, r_3\,|_{\Omega_2}\right)$ was a sum of the squared differences between both direct components (DC) and alternating components (AC) in the amplitude-frequency domain of the Fourier transforms to the signals $r_3\,(t)$ associated with the two stimuli.

The Euclidean distance was defined as follows: supposing that Ω_1 and Ω_2 are two ω-limit sets composed of N_1 and N_2 discrete points in three-dimensional phase space $(r_1,\ r_2, r_3)$, respectively, denoted as $\{\boldsymbol{\alpha}_1,\ \boldsymbol{\alpha}_2,\ \ldots, \boldsymbol{\alpha}_{N_1}\}$ and $\{\boldsymbol{\beta}_1,\ \boldsymbol{\beta}_2,\ \ldots, \boldsymbol{\beta}_{N_2}\}$ where $\boldsymbol{\alpha}_i$ and $\boldsymbol{\beta}_j$ are three-dimensional vectors, then

$$d^E = \left\langle d\left(\boldsymbol{\alpha}_i,\ \boldsymbol{\beta}_j\right)\right\rangle,\ i = 1, \ldots,\ N_1,\ j = 1, \ldots,\ N_2 \qquad (4)$$

Where $\langle\cdot\rangle$ denotes the average and $d\left(\boldsymbol{x}, \boldsymbol{y}\right) = \sqrt{\left(\boldsymbol{x}-\boldsymbol{y}\right)^T\left(\boldsymbol{x}-\boldsymbol{y}\right)}$ which is the standard distance between any two points in the three-dimensional Euclidean space. Note that in the case of equilibria, the number of discrete points in the ω-limit set was one: $N = 1$, and in the case of a limit cycle, we set $N = T/dt$, where T was the period and dt was time bin for numerical integration.

The Euclidean distance d^E worked well in measuring the distance between two equilibria in the phase space, but when one ω-limit set was a limit cycle, the averaging operation in the definition made it only a coarse and lagged estimate of the separation for equilibrium vs. limit cycle and limit cycle vs. limit cycle. In particular, the activity of the P population should be decodable with respect to the stimulus information, something that was problematic for when using d^E. Therefore, we defined the spectrum distance d^S to address the question of distances between representations that were sensitive to those representations being oscillations. To calculate the spectrum distance d^S between two ω-limit sets Ω_1 and Ω_2, only the sequence of their r_3 component was decomposed by Fourier transform which converted the firing rate signal in the temporal domain into a representation in the frequency domain.

The single-sided amplitude spectrum for the Fourier transform of the firing rate signal $r_3\,(t)$, was used to obtain peaks around frequency values. For an equilibrium corresponding to constant firing rate $r_3 = A$, there existed only one peak around zero frequency with its amplitude proportional to A, since the Fourier transform of a constant function is a delta-function. We referred this component as the direct component (DC) of the signal. For the case of a limit cycle, in addition to one peak around zero frequency, there existed another peak around frequency $1{,}000/T$ where T denoted the period of the limit cycle. We referred this additional peak of a limit cycle as the alternating component (AC). Thus, the spectrum distance d^S

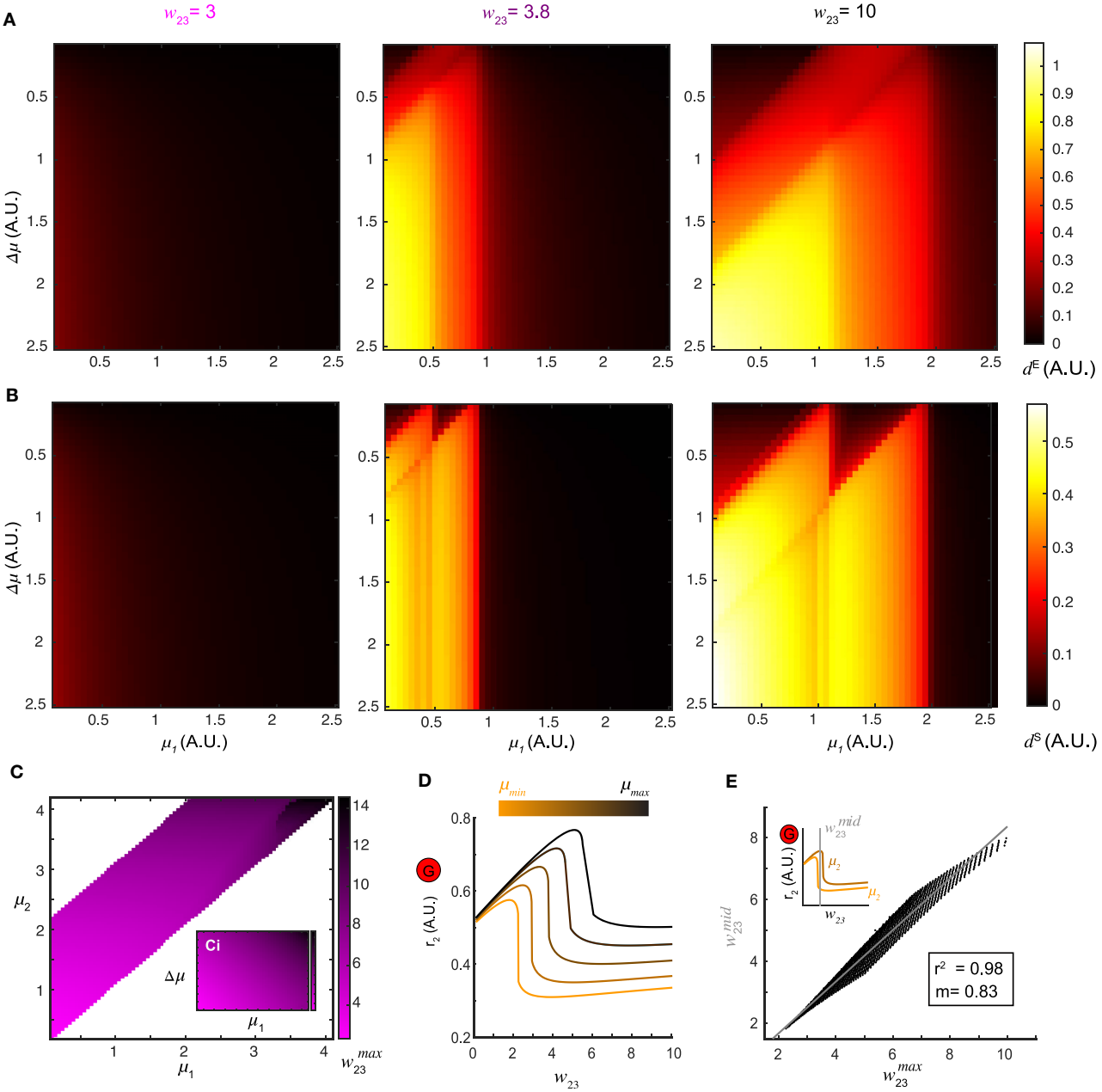

FIGURE 4 | Optimal top-down input for pattern separation. **(A)** Landscapes of the Euclidean distance d^E over all stimulus pairs at three representative values of top-down input. **(B)** Landscapes of the spectrum distance d^S unfolds as in **(A)**. **(C)** The matrix w_{23}^{max} which maximizes d between network representations in response to all combinations of stimuli (μ_1 and μ_2) presented to the network. Inset: the same matrix of w_{23}^{max} organized by one stimulus μ_1 vs. stimulus difference $\Delta\mu$. **(D)** Dependence of the stationary firing rate of inhibitory population r_2 on feedback w_{23} at different levels of input strength (indicated by color bar). **(E)** The correlation between w_{23}^{max} obtained from **(C)** for all pairs of stimuli μ_1 and μ_2 and the w_{23}^{mid} corresponding to the mid-point of the two inhibitory firing rate maxima associated with the same pair of stimuli (upper left inset) reveals that top-down input optimizes pattern separation by gating the G-M inhibition as well as recurrent G inhibition. The gray line denotes the utility line. Lower right inset shows the correlation coefficient and the slope of the linear regression.

was a sum of the differences between both direct components (DC) of two limit sets and alternating components (AC) of two limit sets, which was formalized as follows: supposing that D_1 and D_2 were the amplitudes of the peaks at zero frequency for two ω-limit sets Ω_1 and Ω_2, and $\mathbf{a}_i = (f_i, A_i)$, $i = 1, 2$ denoted the corresponding alternating components of Ω_i, where f_i was the non-zero frequency and A_i was the amplitude of the peak around f_i, then we have

$$d^S = |D_1 - D_2| + d(\boldsymbol{a}_1, \boldsymbol{a}_2) \quad (5)$$

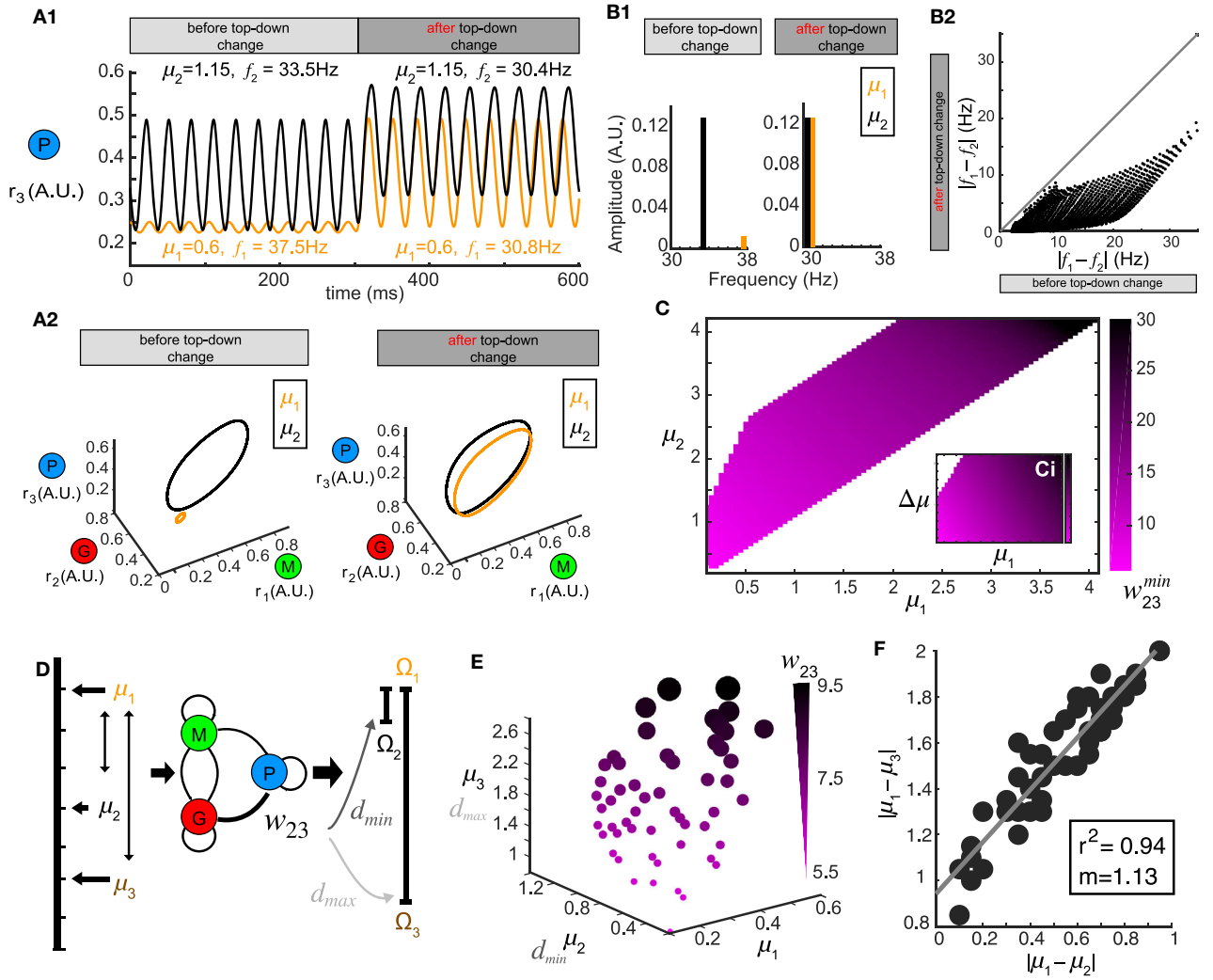

FIGURE 5 | Oscillation synchrony via top-down control. **(A)** Oscillatory responses to two example stimuli $\mu_1 = 1.15$ and $\mu_2 = 0.6$ become synchronized promptly after changing the top-down input w_{23}. **(A1)** Time series of r_3 (t) before and after changing the top-down input. **(A2)** Limit cycles corresponding to the oscillatory responses in **(A1)** are plotted in phase space (transitions not shown). **(B)** Changing w_{23} can make both the frequency and amplitude of two oscillations closer to each other. **(B1)** Frequency and amplitude components of the two oscillations shown in **(A)**. **(B2)** Changing top-down input reduces the frequency differences of responses to two distinct stimuli, effectively using frequency to synchronize the representations in the phase space. **(C)** The matrix w_{23}^{min} which minimizes the distance d between oscillatory responses to all combinations of stimuli μ_1 and μ_2. **(D)** Schematic diagram illustrating that the same value of w_{23} which minimizes the distance between oscillations responding to stimuli μ_1 and μ_2 can maximize the distance between responses to stimuli μ_1 and μ_3. **(E)** Scatter plot in μ_1- μ_2- μ_3 space where each sphere denotes a top-down input w_{23} as illustrated in **(D)** and is coded by color and size. **(F)** Correlation between the differences of those stimuli of which the response distances are simultaneously minimized and maximized. Inset: correlation coefficient and the slope of linear regression.

Note that we set $\boldsymbol{a}_i = (0, 0)$ if the ω-limit sets Ω_i was an equilibrium. Thus, when the two limit sets were both equilibria, d^S only contained the first term measuring the difference between the direct components. In this case, d^S was only a linear projection (up to a constant factor) of the Euclidean distance d^E.

When both Ω_1 and Ω_2 were equilibria, the spectrum distance d^S was a projection of the Euclidean distance d^E onto the r_3 axis. The spectrum distance was however more sensitive to bifurcations when one of the two ω-limit sets Ω_1 and Ω_2 transitioned into a limit cycle as the frequency of the limit cycle started from non-zero values at the onset of a bifurcation (**Figure 2C1**), causing discontinuous jumps in the spectrum distance. However, for the purpose of pattern separation, the two metrics did not give qualitatively different results when assessing the distances due to changes the feedback weight w_{23} (**Figure 4**). Additionally, we found that the non-monotonic dependence of distance in both d^E and d^S on the feedback weight ($w_{23}^{\max}$) persisted, and that an optimal value for any given pair of stimuli could be found (**Figure 3D**).

Bifurcation Analysis

We denoted the dynamical system of Equation (1) as a parameterized form

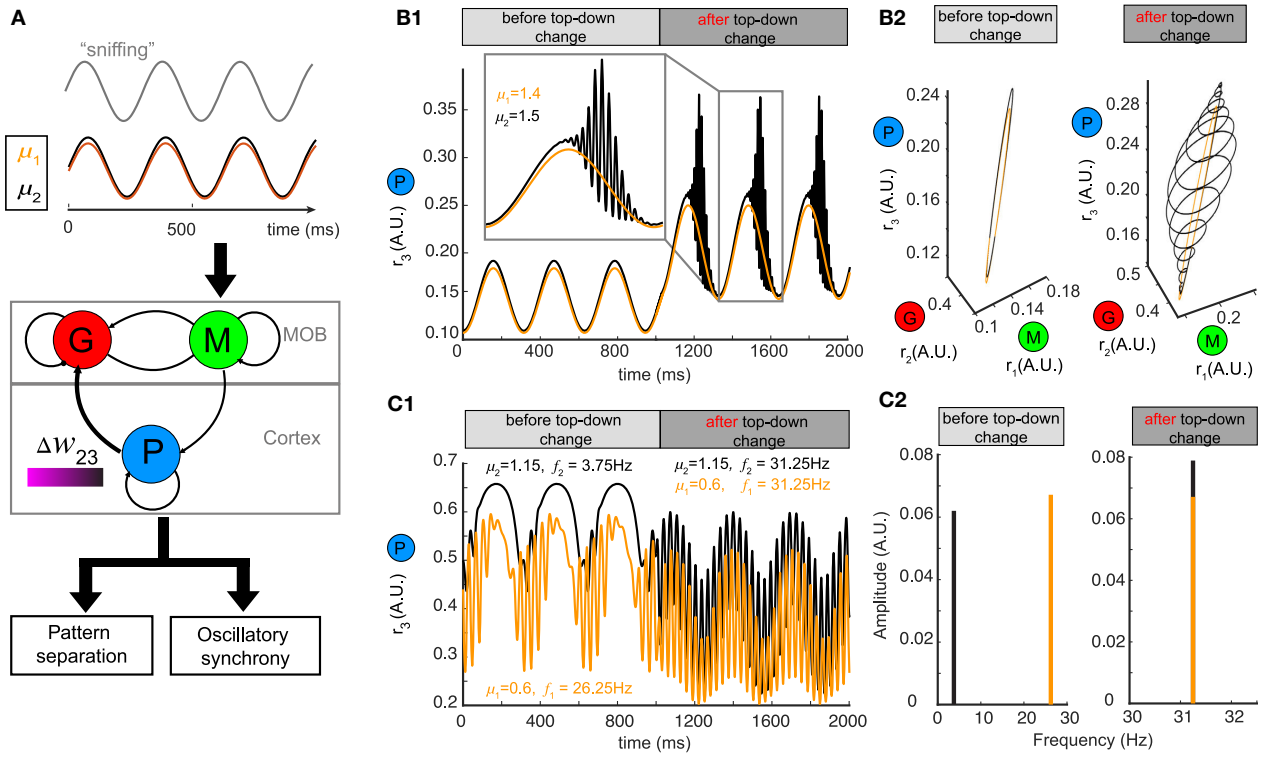

FIGURE 6 | Pattern separation and oscillation synchrony for sniff-modulated oscillatory stimulus. **(A)** Schematic diagram illustrating that varying the top-down weight of the network model (middle) can accomplish both pattern separation and oscillation synchrony (bottom) for a pair of oscillatory stimuli μ_1 and μ_2 modulated by sniffs (top). **(B)** Pattern separation for two oscillatory stimuli with closely related amplitudes: $\mu_1 = 1.4$ and $\mu_2 = 1.5$. **(B1)** Firing rate of the piriform population $r_3\,(t)$ before and after changing the top-down weight: before, $w_{23} = 26$; after, $w_{23} = 21$. **(B2)** The limit cycles in the phase space corresponding to the firing rate in **(B1)**. **(C)** Oscillation synchrony for two oscillatory stimuli with distinct amplitudes: $\mu_1 = 1.15$ and $\mu_2 = 0.6$. **(C1)** Changing the top-down weight from $w_{23} = 6.1$ to $w_{23} = 9.2$ synchronizes the intrinsic oscillations of the firing rate of the piriform population $r_3\,(t)$. **(C2)** Frequency and amplitude components of the two intrinsic oscillations shown in **(C1)**.

$$\dot{\boldsymbol{r}} = \boldsymbol{f}(\boldsymbol{r}, \Theta) \tag{6}$$

where $\boldsymbol{r} \in \mathbb{R}^3$ was the vector of firing rates and $\Theta \in \mathbb{R}^p$ was the vector of parameters. The vector field $\boldsymbol{f} = (f_1, f_2, f_3)^T$ was a smooth function on some open set of $\mathbb{R}^3 \times \mathbb{R}^p$. The dimensionality of Θ could be up to eight dimensions maximally to include all connection weights w_{ij} and the external stimulus μ. However, since we were only interested in the top-down control, Θ was restricted to be two dimensions ($p = 2$) including the top-down weight w_{23} and the external stimulus μ. All the other parameters were fixed as constants determined based on previous experimental work (Whittington et al., 2000). As the system Equation (6) had an equilibrium at $(\boldsymbol{r}, \Theta) = (\boldsymbol{r}_0, \Theta_0)$, i.e.,

$$\boldsymbol{f}(\boldsymbol{r}_0, \Theta_0) = \boldsymbol{0} \tag{7}$$

the stability of this equilibrium could be determined from the linearized vector field of Equation (6) given by

$$\dot{\xi} = D_{\boldsymbol{r}}\boldsymbol{f}(\boldsymbol{r}_0, \Theta_0)\,\xi, \xi \in \mathbb{R}^3 \tag{8}$$

where $D_{\boldsymbol{r}}\boldsymbol{f} \triangleq \frac{\partial \boldsymbol{f}}{\partial \boldsymbol{r}}$ was the Jacobian matrix of the vector field **f**.

If none of the eigenvalues of $D_{\boldsymbol{r}}\boldsymbol{f}(\boldsymbol{r}_0, \Theta_0)$ lied on the imaginary axis (i.e., the equilibrium was *hyperbolic*), the local stability of $(\boldsymbol{r}_0, \Theta_0)$ in the non-linear system (6) could be determined by the linear system (8). The equilibrium was stable if all eigenvalues of $D_{\boldsymbol{r}}\boldsymbol{f}(\boldsymbol{r}_0, \Theta_0)$ had negative real parts. In the case of a hyperbolic equilibrium, varying slightly the parameter Θ would not change the stability as taking Equation (7) and the invertibility of $D_{\boldsymbol{r}}\boldsymbol{f}(\boldsymbol{r}_0, \Theta_0)$, there existed a unique smooth function $\boldsymbol{h}: \mathbb{R}^p \to \mathbb{R}^3$ such that

$$\boldsymbol{h}(\Theta_0) = \boldsymbol{r}_0 \text{ and } \boldsymbol{f}(\boldsymbol{h}(\Theta), \Theta) = \boldsymbol{0} \tag{9}$$

for Θ sufficiently close to Θ_0. By continuity of the eigenvalues with respect to parameters, $D_{\boldsymbol{r}}\boldsymbol{f}(\boldsymbol{h}(\Theta), \Theta)$ had no eigenvalue on the imaginary axis for Θ sufficiently close to Θ_0. Therefore, the hyperbolicity of the equilibrium persisted and its stability type remained unchanged for in close vicinity of Θ_0. By contrast, when some of the eigenvalues of $D_{\boldsymbol{r}}\boldsymbol{f}(\boldsymbol{r}_0, \Theta_0)$ lied on the imaginary axis, for example, a zero eigenvalue or a pair of purely imaginary eigenvalues, new topologically different dynamical

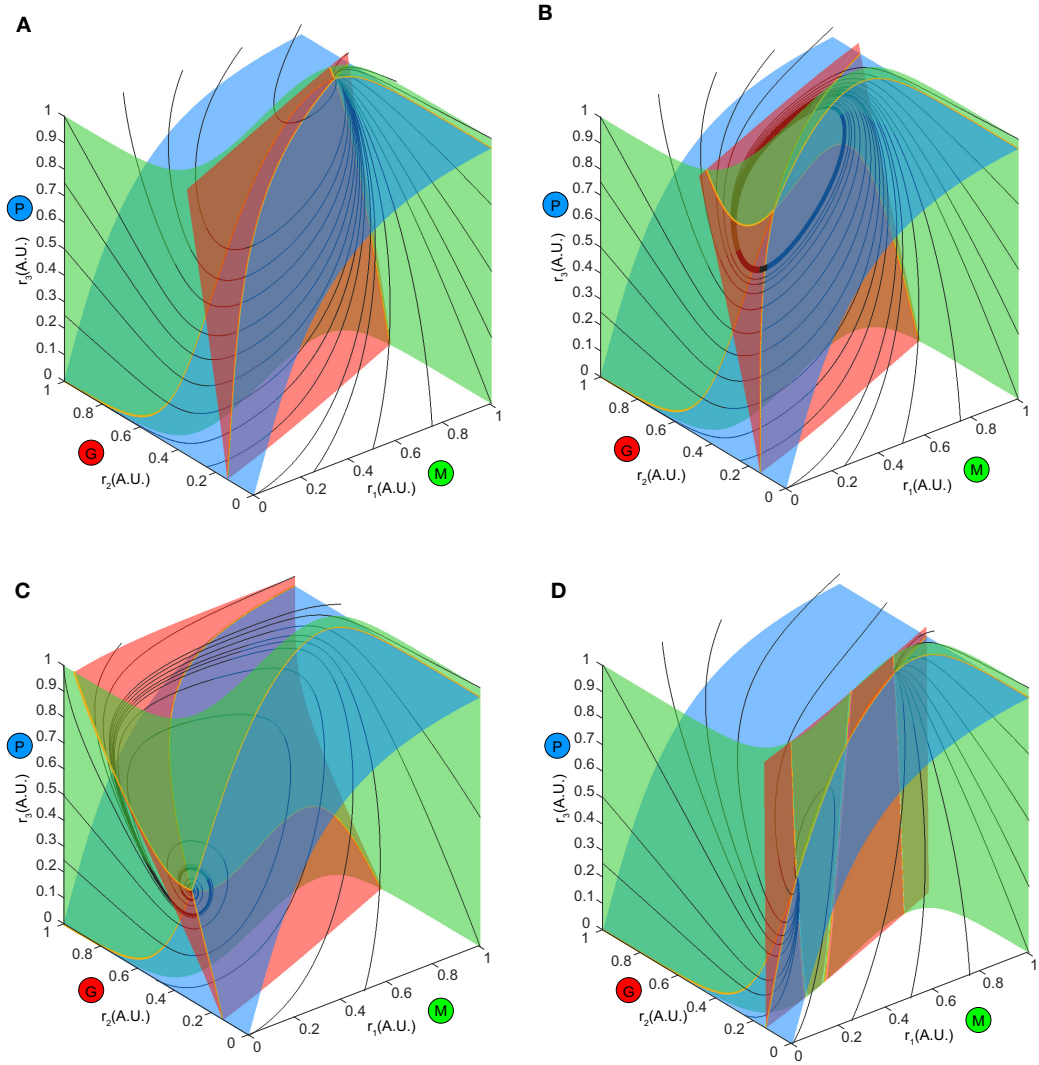

FIGURE 7 | Global phase structure changes with the top-down input. **(A–C)** Global phase structures showing the nullclines (yellow thick curves), nullplanes (transparent surfaces with the same color code as M, G, P population) and several representative trajectories (black thin curves) for the same stimulus $\mu = 1.5$ and three different values of w_{23}. A, $w_{23} = 4$; B, $w_{23} = 6$; C, $w_{23} = 15$. Varying the top-down input tilts the nullplanes, thus changing the position of the equilibrium as well as its stability. **(D)** An example phase structure where three equilibria were present simultaneously (two stable and one unstable), corresponding one of the blue regions of the manifold in **Figure 8A**.

behaviors occurred by a small change in Θ. Equilibria could be created or annihilated, and periodic dynamics could emerge.

The parameterized system (6) thus underwent a bifurcation at $(\boldsymbol{r}_0, \Theta_0)$ if the Jacobian matrix $D_{\boldsymbol{r}}\boldsymbol{f}\,(\boldsymbol{r}_0, \Theta_0)$ has an eigenvalue of zero real part. In our model, a saddle-node bifurcation (SN) occurred when $D_{\boldsymbol{r}}\boldsymbol{f}\,(\boldsymbol{r}_0, \Theta_0)$ had a single zero eigenvalue (in addition to some non-degenerate conditions), and a Hopf bifurcation (H) occurred when $D_{\boldsymbol{r}}\boldsymbol{f}\,(\boldsymbol{r}_0, \Theta_0)$ had a pair of purely imaginary eigenvalues. The bifurcation point was found numerically by XPPAUT or the Matlab toolbox MATCONT.

The number of parameters that must be varied simultaneously to evoke a bifurcation is defined as the *codimension* of this bifurcation (Guckenheimer and Holmes, 2013; Kuznetsov, 2013). Considering the infinite-dimensional space $\mathcal{H}$ of all vector fields defined on the n-dimensional Euclidean space $\mathbb{R}^n$, a vector field $\mathbf{f}_0$ undergoing a bifurcation, for example, a Hopf bifurcation, corresponds to a point in the space $\mathcal{H}$. All nearby vector fields with the same singularity as $\boldsymbol{f}_0$ (i.e., vector fields that are orbitally topologically equivalent to $\boldsymbol{f}_0$) form a submanifold $\mathcal{L}$ of co-dimension k, which is an equivalence class of the singular vector field $\boldsymbol{f}_0$. Therefore, within the space $\mathcal{H}$ it requires another submanifold $\tilde{\mathcal{L}}$ of at least k dimensions to intersect transversely with $\mathcal{L}$ at point $\boldsymbol{f}_0$, such that the singularity of $\boldsymbol{f}_0$ persists under small perturbations of the vector field. The submanifold $\tilde{\mathcal{L}}$ was obtained through a parametrized family of vector fields involving at least k parameters. The least number k is then defined as

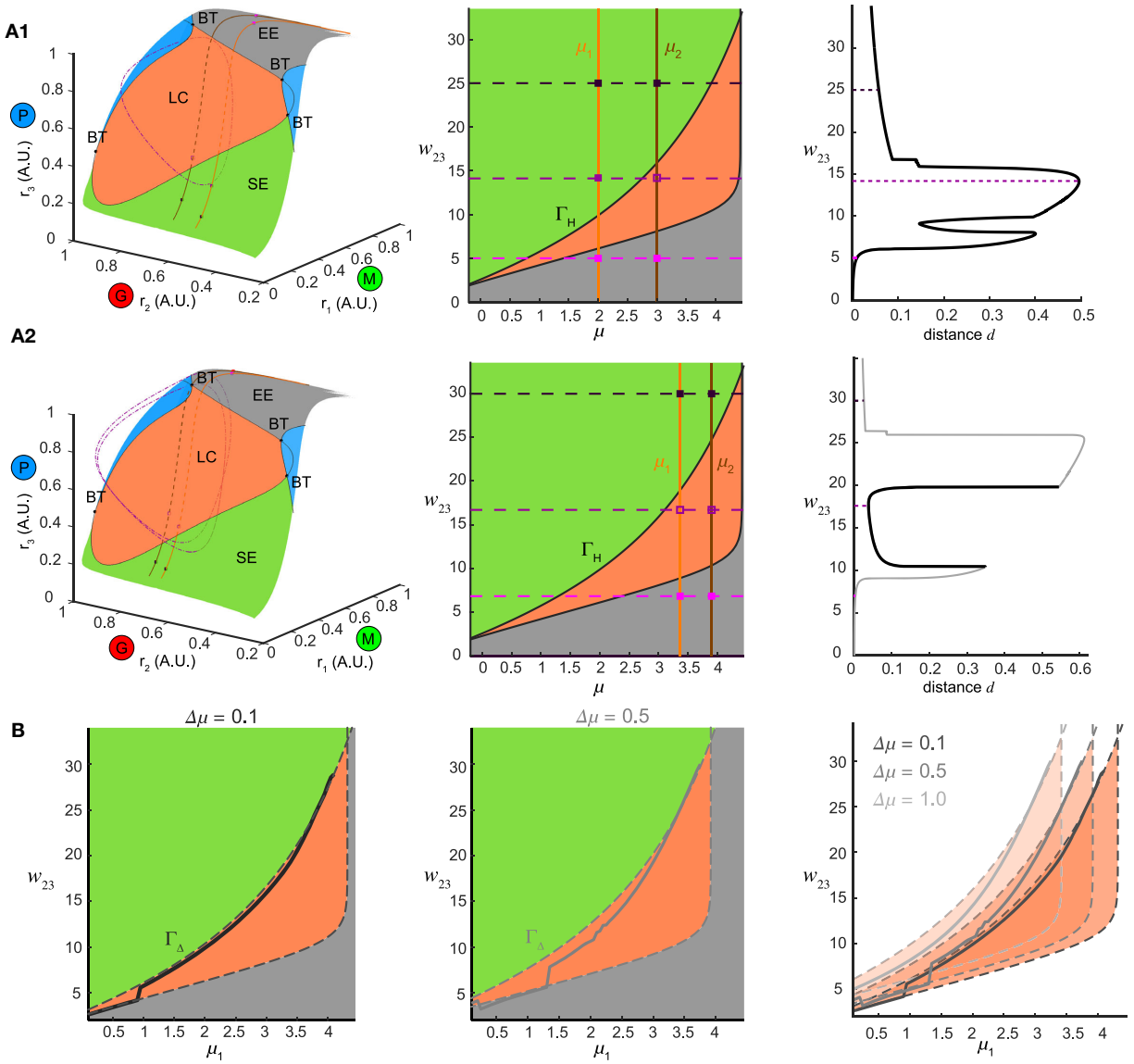

FIGURE 8 | Bifurcation mechanism of top-down control to support both pattern separation and oscillation synchrony. **(A)** Illustration of the bifurcation mechanism and the transition boundary of dynamics in phase space and parameter space. **(A1)** Pattern separation. Left: the equilibrium manifold in the firing rate phase space divided by the separatrix emitting from multiple codimension-2 bifurcation points: BT (Bogdanov–Takens bifurcation) into several regions. Region-LC: each equilibrium was unstable and had exactly one corresponding stable limit cycle (dot dashed cycle) arising from a Hopf bifurcation. Region-SE: each equilibrium was stable all trajectories spiraled into it. Region-EE: each equilibrium was stable all trajectories approached it exponentially. Blue regions: regions where multiple equilibria coexisted. For two example stimuli $\mu_1 = 2.0$ and $\mu_2 = 3.0$ given in the middle of **(A1)**, two paths of equilibria were induced on the equilibrium manifold and traversed across different regions as changing top-down input w_{23}. Middle: different regions on the equilibrium manifold corresponded to different regimes in the parameter space of μ and w_{23} in the same color scheme [parameters for blue regions in Left were largely beyond the range thus not shown]. The transition boundary Γ_H specified the pair (μ,w_{23}) at which the network underwent a Hopf bifurcation and corresponded to the separatrix enclosing the region-LC in Left. Two given stimuli were denoted by two vertical lines and three example values of w_{23} corresponded to three horizontal dashed lines, giving rise to a pair of junctions for each. These junctions were also plotted as squares in the left of **(A1)** denoting the corresponding ω-limit sets in the same color (solid square: stable equilibrium; empty square: unstable equilibrium). Right: the distance between the ω-limit sets to represent the two given stimuli. The maximal distance was achieved when the two junctions were on opposite sides of Γ_H. **(A2)** Same as **(A1)** but for oscillation synchrony occurring when the two junctions were both inside Γ_H. **(B)** Comparisons between the translated transition boundary Γ_Δ (dashed curve) depending on μ (left: $\Delta\mu = 0.1$, middle: $\Delta\mu = 0.5$) and the sliced section of the $w_{23}^{\max}$ at the same $\Gamma\mu$ (solid curve). Right: a series of translated Γ_Δ for three representative values of $\Gamma\mu$.

the codimension of f_0. The parametrized vector field $f(r, \Theta)$ in Equation (6) can be thought of as one realization of the submanifold $\tilde{\mathcal{L}}$ which passes through the vector field $f_0 \triangleq f(r_0, \Theta_0)$ undergoing a bifurcation with the two parameters corresponding to the top-down weight w_{23} and the external stimulus μ.

Simulating Sniffing With a Periodically Driven Non-autonomous System

Olfactory sensation is an active process, with sensory stimuli being sampled by sniffing on the time scale of 215Hz in animal models (Carey and Wachowiak, 2011; Wachowiak, 2011). To simulate sniffing, a vector field for the dynamical system that depended explicitly on time and was also periodic with fixed period $T = 2\pi/\omega > 0$, i.e.,

$$\dot{\boldsymbol{r}} = \boldsymbol{f}(\boldsymbol{r}, t) \ and \ \boldsymbol{f}(\boldsymbol{r}, t+T) = \boldsymbol{f}(\boldsymbol{r}, t) \tag{10}$$

could be rewritten in the form of an autonomous system by defining the function

$$\varphi(t) = \omega t, \ mod\, 2\pi \tag{11}$$

such that using Equation (11), Equation (10) became

$$\begin{array}{c} \dot{\boldsymbol{r}} = \boldsymbol{f}(\boldsymbol{r}, \varphi) \\ \dot{\varphi} = \omega \end{array} (\boldsymbol{r}, \varphi) \in \mathbb{R}^3 \times S^1 \tag{12}$$

where S^1 denoted a circle. To construct the Poincaré map, we defined a cross-section of Equation (12) by

$$\Sigma_0^{\overline{\varphi}} = \left\{ (\boldsymbol{r}, \varphi) \in \mathbb{R}^3 \times S^1 \,|\varphi = \overline{\varphi}_0 \in (0, \ 2\pi] \ \right\} \tag{13}$$

such that a fixed point of the Poincaré map $P_{\overline{\varphi}_0} : \Sigma^{\overline{\varphi}_0} \rightarrow \Sigma^{\overline{\varphi}_0}$ corresponded to a limit cycle of the extended system Equation (12), and a limit cycle of $P_{\overline{\varphi}_0}$ corresponded to a two-dimensional (2D) torus of Equation (12).

Topological changes in the ω-limit sets of the extended system Equation (12) could thus be understood via bifurcations of the discrete map $P_{\overline{\varphi}_0}$. Specifically, the bifurcation analysis we performed for autonomous system (6) also applied to the Poincaré map $P_{\overline{\varphi}_0}$. Hopf bifurcations undergone in autonomous system (6) which gave rise to limit cycles in 3D phase space corresponded to *Neimark-Sacker bifurcations* of $P_{\overline{\varphi}_0}$ which gave birth to a 2D torus in the extended space. The torus oscillation thus had two periodic components: one (the toroidal direction) driven extrinsically by the frequency of sniffs and the other (the poloidal direction) governed by the intrinsic network dynamics. Therefore, the Neimark-Sacker bifurcations provided an analogous bifurcation mechanism for non-autonomous system (10) as the Hopf bifurcations did for autonomous system (6).

RESULTS

Reduced Network Model Generates Complex Dynamics

To understand the functional role of top-down projections onto inhibitory neurons, we built a three-node network model (**Figure 1A**, see Methods) that recapitulated a circuit architecture identified both structurally (Padmanabhan et al., 2019) and functionally (Boyd et al., 2012; Markopoulos et al., 2012) across a number of brain areas. For different stimuli μ, the network exhibited a variety of dynamics (**Figure 1B**). For instance, when the stimulus was small, the firing rates $r_i, i = 1, 2, 3$ had a fast-transient increase followed by damping oscillations that converged to a stationary state (**Figure 1B**, left). A sufficiently large stimulus μ elevated the firing rates to near saturation, where they then remained at the upper bound of the non-linear sigmoid function throughout the duration of the stimulus (**Figure 1B**, right). For small or large stimuli, the network responses converged to a constant firing rate after transient dynamics. By contrast, for medium values of μ, more complex firing rate dynamics emerged, including oscillations (**Figure 1B**, middle). To visualize the collective behaviors of M, G, and P populations to these different stimuli, we turned to a three dimensional dynamical system representation of the model where the time evolution of the firing rates (i.e., *state variables*) was a *trajectory* (or an *orbit*) in the *phase space* (r_1, r_2, r_3) and the tangent vector defining the velocity of each point along a trajectory was given by the *vector field* $\boldsymbol{f} = \left(f_1, f_2, f_3\right)^T$ (see Methods) of Equation (1). The firing rates over time in **Figure 1B** thus corresponded to trajectories in **Figure 1D** starting from the origin $\mathcal{O}$ (where all three populations were silent). For small stimuli, the trajectory made an excursion before spiraling into an *equilibrium* indicated by the solid dot (**Figure 1D**, orange). Similarly, when the stimulus μ was large, the trajectory again settled into an equilibrium, but one that was translated within the phase space to the top-right corner (**Figure 1D**, black). Finally, for medium stimuli, the time-varying oscillation of firing rates manifested as a *periodic orbit* (or a *limit cycle*) in the 3D phase space (**Figure 1D**, brown). By convention, we defined the steady-state dynamics as the *ω-limit set* of the system.

Top-Down Weight Reshapes Network Dynamics and Modulates Neural Oscillations

Next, to explore how top-down down projections onto the inhibitory granule cell population (G) shaped the dynamics of the network, we studied the effects of changes in the connection weight w_{23} on firing rate dynamics. First, we varied the top-down weight w_{23} (**Figure 2A**, top) from the piriform population (P) to the inhibitory granule cell population (G) and studied the effect of these changes on the firing rate dynamics of the network. For a fixed stimulus ($\mu = 1.5$) the dynamics of the firing rates $r_i(t), i = 1, 2, 3$ were sensitive to different values of w_{23} (**Figure 2A**, bottom). When the top down weight was small ($w_{23} = 4$), firing rates approached the equilibrium exponentially (**Figure 2A**, bottom and **Figure 2B**, left, black traces). Conversely, when the top-down weight was large ($w_{23} = 10.5$), the firing rate of excitatory cells (r_1) increased initially, but was suppressed as inhibition reduced the activity, until the firing rates ultimately settled to an equilibrium (**Figure 2A**, top and **Figure 2B**, left, light magenta traces). When the magnitude of top-down weight was changed to an intermediate value ($w_{23} = 5.5$), the same stimulus generated oscillatory activity in the network, with the steady-state dynamics transitioning to a periodic orbit (a limit cycle). Changing the weight of top-down projections onto the local inhibitory population for a single

stimulus produced the same diversity of firing rate dynamics that occurred from changes in the stimulus. Furthermore, for a given top-down weight (w_{23}), the effects on the network dynamics stimulus was unique to that stimulus (**Figure 2B**, right vs. left).

In regimes where specific weights of top-down weight generated sustained oscillatory activity for a given stimulus μ, we characterized the frequency and amplitude of these oscillations (**Figure 2B**, left $w_{23} = 5.5$, right $w_{23} = 10.5$) as changes in both have been tied to circuit function and behavior (Buzsaki and Draguhn, 2004; Kay et al., 2009). For a given stimulus, oscillations emerged between two values of w_{23}, with the frequency of the oscillation varying monotonically (**Figure 2C1**). By contrast, while the amplitudes of the oscillations started from zero at the two critical values of w_{23}, they reached a maximum in between (**Figure 2D1**). The control of both the frequency and amplitude via changes in w_{23} occurred across an array of weights (w_{31}) associated with the feedforward drive from the mitral/tufted population (M) to the piriform population (P) (**Figures 2C2,D2**). Furthermore, the magnitude of synaptic weights from mitral/tufted cells to piriform cortical neurons established the dynamic range within which changes in top-down weights (w_{23}) influenced the frequencies (**Figure 2C3**, $0 - 45$Hz) and amplitudes (**Figure 2D3**, $0 - 1$ A.U.) of network oscillations, spanning frequencies in the alpha, beta and gamma bands.

Top-Down Weight Contributes to Pattern Separation

As changing the top-down weight onto inhibitory neurons could generate complex activity patterns we next asked what computations could be performed by this control. For example, both behavioral and neurophysiological measures show that as the representations of two stimuli by neuronal circuits become different, distinguishing between them becomes easier (Friedrich and Laurent, 2001; Leutgeb et al., 2007; Yassa and Stark, 2011) Control of inhibition, via top-down centrifugal projections, may be one way that such stimulus discrimination is implemented by the circuit.

To test this hypothesis, we presented our network with a pair of stimuli, denoted by μ_1 and μ_2 (corresponding to stimuli arranged along a one-dimensional axis) and studied how control of inhibition altered the representations of the two stimuli (**Figure 3A**). Conceptually, these two stimuli could be two different concentrations of an odor or two odors that share a similar physiochemical feature (two odors with different carbon chain lengths). For a set of stimuli $\mu_i, i = 1, 2$, we defined the steady-state *representation* of network activity as the ω-limit set $\Omega_i, i = 1, 2$. The distance between the two stimuli μ_1 and μ_2 in the stimulus space was defined as $\Delta\mu$, and the resultant distance in the firing rate phase space between the two ω-limit sets (Ω_1 and Ω_2) we defined as a metric d (see Methods). The smaller the $\Delta\mu$, the more similar the two stimuli were. We hypothesized that changes in the weight of feedback onto the inhibitory neuron population (w_{23}) could increase the value of d, making the representations of those stimuli more distinct (**Figure 3A**).

In a representative example where $\mu_1 = 1.0$ and $\mu_2 = 1.1$, when the top-down weight was low ($w_{23} = 3$), the representations of the two stimuli were close (**Figures 3B,C**, left). As w_{23} was increased, the representations of the two stimuli were pushed apart making them more separable ($w_{23} = 3.8$, **Figures 3B,C**). Interestingly, as w_{23} was increased further ($w_{23} = 10$), the representations of the two stimuli became close to one another again (**Figures 3B,C**, right). As representations could be either oscillations in the state space, or equilibria, we compared how the distances of these representations changed across different measures (see methods). Interestingly, although the absolute values given by the Euclidean distance d^E (**Figure 3C**) and the spectrum distance d^S (**Figure 3B**) were different, they both occurred at the same feedback weight (**Figure 3D**). We visualized the distance landscapes defined by d^E and d^S over all combinations of μ_1 and $\Delta\mu$ as a function of a change in the weight of the top-down weight (**Figures 4A,B**). Irrespective of which distance definition was exploited for measurement, we found an optimal value of $w_{23}^{\max}$ that maximized the distance between the two resultant representations for any given pair of stimuli.

A landscape of the optimal $w_{23}^{\max}$ across all pairs of stimuli (μ_1, μ_2) was shown in **Figure 4C**. Thus, changing the weight of top-down projections onto the inhibitory neuron population could be used to facilitate stimulus separation dynamically. To understand why, we examined the effect that varying the top-down weight had on the firing rate responses of both local excitatory mitral/tufted cell and inhibitory granule cell populations (r_1 and r_2). For a given stimulus, an increase in w_{23} led to a monotonic decrease in r_1, suggesting persistent suppression onto the local M population. By contrast, the response of the local G population r_2 was elevated first with increasing w_{23} until reaching the maximum $r_2^{\max}$, after which r_2 dropped significantly (**Figure 4D**). Across different stimuli μ, the shape of the firing rate r_2 as a function of w_{23} remained the same but shifted vertically. To determine if these differences in the firing rate of inhibitory neurons (r_2) were related to the values of top-down weights that maximally separated the distance between two stimulus representations, we plotted the $w_{23}^{\max}$ (abscissa) obtained from the landscape in **Figure 4C** vs. the midpoint w_{23}^{mid} (ordinate) between the $r_2^{\max}$ of the same stimulus pair (inset of **Figure 4E**). The response r_2 to one stimulus on the left of the midpoint dropped significantly, while the response to a similar stimulus on the right still had a high inhibitory firing rate. The optimal $w_{23}^{\max}$ was correlated to w_{23}^{mid} ($R^2 = 0.98$), the value at which inhibitory neuron activity from one stimulus was suppressed while activity from the other similar stimulus remained persistently high. Consequently, stimulus separation arose from the differential sensitivity of inhibitory neurons to the balance between top-down feedback and recurrent inhibition; an imbalance occurred between the top-down feedback and the recurrent inhibition for one stimulus while that balance was preserved for the second stimulus.

Top-Down Weight Contributes to Oscillation Synchrony

Stimulus-evoked oscillations also appeared in our model, and were modulated by the top-down weights (**Figure 2**) covering a wide range of frequency and amplitude. This suggested that oscillatory responses to different stimuli could be synchronized by tuning w_{23}. To explore this, we first examined the oscillations in the firing rate generated by two different stimuli μ_1 and μ_2. At a given value of top-down weight ($w_{23} = 12.0$), one stimulus ($\mu_1 = 1.15$) generated oscillations ($f_1 = 33.5$ Hz, **Figure 5A1**, before) in the piriform population's firing rate that were different in both frequency and amplitude from the oscillations ($f_2 = 37.$Hz, **Figure 5A1**, before) in response to a second stimulus ($\mu_2 = 0.6$). However, a change in the top-down weight ($w_{23} = 8$), resulted in firing rate oscillations becoming more similar for the same two stimuli (**Figure 5A1**, after, $f_1 = 30.4$ Hz, $f_2 = 30.8$ Hz). This increase in the firing rate synchrony was also apparent when visualized in the 3D phase space (**Figure 5A2**). To quantify the synchrony between the oscillations responses to μ_1 and μ_2, we calculated the spectrum distance d^S (see Methods) between the network representations for the two stimuli before and after changes in top-down weight (**Figure 5B1**). Changes in the feedback to inhibitory neurons w_{23} synchronized activity in the network stimuli (**Figure 5B2**), and while the effect was greatest when stimuli were similar, we found examples for stimuli that were initially as far apart as 20 Hz. As with stimulus discrimination, a systematic relationship emerged corresponding to the optimal top-down weight $w_{23}^{\min}$ across combinations of stimuli (μ_1 vs. μ_2) that was most effective at generating synchronous oscillations (**Figure 5C**).

Although we have thus far treated stimulus discrimination and synchrony separately, neural circuits perform both operations simultaneously, bringing the network representation of one stimulus closer to another, while simultaneously pushing the representation of that stimulus farther from a third. We therefore tested if a single change in the top-down weight w_{23} accomplish both of these operations; minimize the distance between the responses to one pair of stimuli (μ_1 vs. μ_2) while also maximize the response distance to another other pair of stimuli (μ_1 vs. μ_3, **Figure 5D**). To do this, we generated a 3D scatter plot of values of w_{23} that were optimal for synchrony between oscillations generated by stimulus μ_1 and μ_2 (**Figure 5C**) and also produced a maximum separation between the representations of stimulus μ_1 and μ_3 (**Figure 5E**). The values of top-down weight w_{23} for each point that fulfilled these diametrically distinct functions were coded by color and size (**Figure 5E**). We found that the top-down weight corresponding to both operations scaled with the stimuli, such that when μ_1, μ_2 and μ_3 were small, the top-down weight was also small, but as the three stimuli increased in magnitude, the top-down weight needed to synchronize one pair and separate the other pair also increased. Finally, we found a strong correlation between the values of stimulus differences: $|\mu_1 - \mu_2|$ and $|\mu_1 - \mu_3|$ (**Figure 5F**) at which an w_{23} weight was optimal for stimulus separation and oscillatory synchrony.

Generalization to Oscillatory Stimulus Driven by Sniffs

Although we used a constant stimulus μ to represent the average input to mitral/tufted cells, in mammals sniffing brings odors into the nasal epithelium in a periodic fashion (Wachowiak, 2011). Sniff cycles carry different amounts of information about odor identity and concentration (Miura et al., 2012) and a single sniff cycle is sufficient for animals to discriminate accurately between two odors (Uchida and Mainen, 2003; Wesson et al., 2008). To explore how changing top-down weights can reshape network responses to oscillatory stimuli, we modeled our stimulus μ as a sinusoidal function $\tilde{\mu}(t) = \mu \cos(\omega_s t + \varphi_0)$, where the different odors had different amplitudes μ, the sniffing frequency ω_s was set to ∼4 Hz and φ_0 characterized the initial phase of sniffing (Carey and Wachowiak, 2011; Shusterman et al., 2011) (**Figure 6A**).

For a pair of oscillatory stimuli $\tilde{\mu}(t)$ with two similar amplitudes μ_1 and μ_2, the firing rate responses were also similar for the piriform population (P) (**Figure 6B1**, before top-down change) and the entire network in the phase space (**Figure 6B2**, left). If we changed the top-down weight w_{23}, as we had done for a fixed stimulus, both the piriform population firing (**Figure 6B1**, after top-down change) and the network representations became more distinct (**Figure 6B2**, right). The conch-shaped limit cycle in **Figure 6B2** (right) arose from oscillations occurring at two different time scales (see methods): a slower oscillation on the time scale of sniff cycles and a faster oscillation governed by the intrinsic dynamics of the network. As a consequence, the same pattern separation achieved by changing top-down weight for constant stimuli could also be accomplished for oscillatory stimuli.

To explore if network representations of oscillatory stimuli could be made synchronous by changing top-down weights, we presented two stimuli with distinct amplitudes (μ_1, μ_2) to the network (**Figure 6A**). Following a change in the top-down weight w_{23}, the network representations became synchronous (**Figure 6C1**), with the oscillations of firing rates occurring at the same frequency (**Figure 6C2**, right). Importantly, these high frequency oscillations occurred at the gamma band, and rode on top of the slower oscillations corresponding to sniff cycles, further revealing the computational decoupling of sniffing and inhibitory dynamics across two different time scales. Taken together, the mechanisms giving rise to both pattern separation and oscillatory synchrony were general to constant and oscillatory inputs.

Bifurcation Mechanism for Top-Down Control of Inhibition

Finally, to understand mathematically how such operations emerged from changes in the top-down weight to inhibition, we studied the structure of the transitions in network firing rates dynamics (**Figure 2A**). These transitions were associated with qualitative or topological changes in the ω-limit sets of the system, indicative of the occurrence of bifurcations in the system.

To explore this further, we first examined the how the ω-limit sets of the system receiving constant stimuli changed with

different top-down weights. An equilibrium corresponding to constant firing rates in the network arose from the intersection of three nullclines (**Figure 7**, yellow thick lines), each resulting from pairwise intersections of three "nullplanes" that characterized the geometric surface on which the firing rate derivatives of one node equaled to zero (**Figure 7**, transparent surfaces). Global phase structures for two representative values of top-down weights (**Figure 7A**: $\omega_{23} = 4$, **Figure 7C**: $\omega_{23} = 15$) illustrated how these equilibria varied within the firing rate phase space. In these two examples, both equilibria were stable and attractive, with all nearby trajectories (**Figures 7A,C**, black thin lines) moving toward them. This was, however, not true for all values of w_{23}. At some critical values of w_{23}, the equilibrium lost stability, and a small-amplitude limit cycle branched from that unstable equilibrium, resulting in the oscillations observed in the dynamics (**Figure 7B**). This transition signified a *Hopf* bifurcation of the system (see Methods), which arose when the top-down weight w_{23} was within a specific regime. Therefore, across all combinations of external stimuli μ and top-down weight w_{23}, we obtained a smooth manifold in the phase space (**Figure 8A**, left), corresponding to a family of ω-limit sets on which the network dynamics settled from any set of initial conditions. Sustained oscillations corresponded to the red region-LC (LC: limit cycle) where each equilibrium (unstable) was paired with exactly one limit cycle born simultaneously via a Hopf bifurcation (the purple empty square vs. the dot-dashed curve). Constant firing rates corresponded to the gray region-EE (EE: exponential equilibrium) and green region-SE (SE: spiraling equilibrium), where the equilibria were stable, approached either exponentially (region-EE) or via damping-oscillations (region-SE). Finally, the two blue regions on the manifold were bounded by saddle-node bifurcations near Bogdanov–Takens (BT), an example global phase structure of which was shown in **Figure 7D**. The equilibrium manifold thus defined the entire family of network representations for all possible combinations of stimuli μ and top-down weigh tw_{23}.

Within the manifold of the stimulus μ and top-down weight w_{23}, we identified a transition boundary Γ_H (black solid curves, **Figure 8A**, middle) corresponding to the separatrix enclosing the region-LC. Γ_H specified the parameter pairs (μ, w_{23}) at which a Hopf bifurcation occurred, thereby dividing the parameter space into regimes with different dynamics (same color coded as **Figure 8A1**, left). For a given pair of stimuli (for example, $\mu_1 = 2.0$, $\mu_2 = 3.0$), changing w_{23} corresponded to shifting the horizontal dashed line vertically (three representatives were shown in **Figure 8A1**, middle), thereby shifting the junctions with the two stimuli (vertical solid lines, **Figure 8A1**, middle) across different regimes in the parameter space. In the firing rate phase space (**Figure 8A**, left), these changes in w_{23} for one stimulus moved the equilibrium through different regions of the manifold: EE-LC-SE, while for another stimulus, a parallel curve on the manifold could also be traced. When the two junctions in **Figure 8A1** (middle) were on different sides of the transition boundary, with one equilibrium in region-LC and the other in region-SE (**Figure 8A1**, left), the two network representations became topologically different from each other; the former a limit cycle, and the latter an equilibrium point. Thus for a combination of stimulus pairs, the optimal ${w_{23}}^{\max}$ for pattern separation was then achieved when the ω-limit sets were on different sides of transition boundary (**Figure 8A1**, right).

Furthermore, when the junction of feedback weight and stimulus pair were both inside the transition boundary (**Figure 8A2**, middle) two limit cycles emerged (one for each stimulus, for example, $\mu_1 = 3.45$, $\mu_2 = 3.95$, **Figure 8A2**) synchronize the network representations. Changes in top-down weight moved the junctions for pairs of stimuli within the parameter space, revealing a shared mechanism supported both stimulus separation and oscillation synchrony, depending on the relative positions of the junctions with respect to the transition boundary.

Finally we determined if the transition boundary identified via analysis of the dynamical system corresponded to the $w_{23}^{\max}$ matrix found in **Figure 4C**i. To do this, we considered a set of initial stimuli μ_1, and a set of distances to a second set of stimuli $\Delta\mu$, wherein each value was an array that defined a set of stimulus pairs $\{(\mu_1, \mu_1 + \Delta\mu) \mid \mu_1 \in [0, 4]\}$. For a given $\Delta\mu > 0$, distinguishing the pair $(\mu_1, \mu_1 + \Delta\mu)$ was the same as distinguishing $(\mu_1 + \mu, \Delta\mu_1)$ in terms of pattern separation. In this framework two different distances, for instance, $\Delta\mu = 0.1$ or $\mu = 0.5$ (**Figure 8B**, left and middle), the set of stimulus pairs had a unique transition boundary Γ_Δ (**Figure 8B**, right). The section of the $w_{23}^{\max}$ matrix in **Figure 4** Ci for the set of stimulus pairs $\left\{(\mu_1, \mu_1 + \Delta\mu) \mid \mu_1 \in [0, 4]\,,\ \Delta\mu \textit{ is given}\right\}$ was correlated with the transition boundary Γ_Δ of the same value μ (**Figure 8B**). For small $\Delta\mu = 0.1$, the slice of the $w_{23}^{\max}$ followed closely with the transition boundary Γ_Δ (**Figure 8B**, left). As the stimulus difference increased (**Figure 8B**, middle), $w_{23}^{\max}$ deviated from the boundary Γ_Δ. Larger $\Delta\mu$'s increased the bifurcation lag between two stimuli such that the stimulus that caused a bifurcation first had more parameter space to develop before the bifurcation of the other stimulus. Conversely, stimulus discrimination was harder as μ decreased because the range of top-down weights w_{23} that separate two stimuli shrank significantly around a close vicinity of the transition boundary. Thus, subtle adjustments of top-down weight around the transition boundary were required to separate similar stimuli from each other. The same analysis could also be performed for the non-autonomous system receiving oscillatory stimuli $\tilde{\mu}(t)$ by investigating bifurcations of fixed points of the constructed Poincaré map on a given cross section (see Methods) with the same computational mechanism arising via the discrete version of Hopf bifurcation, i.e., a Neimark-Sacker bifurcation (Kuznetsov, 2013). Taken together, these results provide a bridge linking the mechanisms that give rise to the dynamics of the neural circuit with the computations performed by the circuit.

DISCUSSION

Using a three-node model, which included top-down projections from piriform cortical cells onto inhibitory granule cells in the main olfactory bulb, we identified a network capable of complex dynamic behaviors, ranging from an attractor to stable oscillations across a range of frequencies and amplitudes. By

changing the weight of these top-down projections, the network could either facilitate pattern separation between two similar stimuli, or synchronize the oscillatory activity produced by two different stimuli. A bifurcation analysis of the dynamical system revealed that both mechanisms emerged from the transition boundary of Hopf bifurcations which branched from co-dimensional two bifurcation points (i.e., the Bogdanov-Takens bifurcation). Furthermore, these computations could be accomplished even when the stimuli were periodic, fluctuating at the frequency of sniffing (Neimark-Sacker bifurcation), suggesting that these findings are a general feature of this network. Our results provide both a mathematical framework for how top-down control of inhibition shapes the dynamics of a network, and a link between such dynamics and the computations that neural circuits can perform.

An important point to consider is how changes in top-down weights may be instantiated biologically? This point depends on the timescale of weight changes. On short time scales, changes in inhibitory drive to granule cells can facilitate olfactory discrimination (Abraham et al., 2010; Nunes and Kuner, 2015) and generate synchronous oscillatory activity among mitral cells in the bulb (Galan et al., 2006). Neuromodulators such as serotonin (Petzold et al., 2009; Kapoor et al., 2016) can act on fast sub second time scales to support both oscillatory synchrony and stimulus discrimination, providing one biological mechanism by which weights can be changed dynamically. By contrast, long-term changes in the bulb may be instantiated by classis synaptic plasticity mechanisms such as LTP (Cauthron and Stripling, 2014), or via the remodeling of synaptic connectivity (Arenkiel et al., 2011; Deshpande et al., 2013), for instance due to adult neurogenesis (Lledo et al., 2006). In these examples, the changes in feedback weight likely reflect slow alterations in network structure that result in stable changes in neural representations, possibly corresponding to learning.

While the biological mechanisms by which the top-down synaptic weights change onto inhibitory neurons may be diverse depending on timescale, we find that such alterations give rise to functionally equivalent changes supporting an array of computations. For instance, changes in the top-down weight would render two stimuli more distinct at the level of firing rates in the population, a process referred to as pattern separation (Cayco-Gajic and Silver, 2019) or decorrelation (Friedrich and Laurent, 2001). Our model predicts that pattern separation arises from the non-monotonic change in firing in granule cells (at the balance between op-down excitation and recurrent inhibition). The top-down weight onto inhibitory neurons sets a gate, allowing some stimuli to cross a threshold of recurrent inhibition, while others do not.

In parallel, changing top-down weights onto inhibitory neurons can increase the synchrony between two stimuli that were initially asynchronous. A number of experimental and theoretical studies have explored the privileged role that inhibitory interneurons play in generating gamma oscillations (Whittington et al., 1995; Hasenstaub et al., 2005; Cardin et al., 2009; Sohal et al., 2009; Tiesinga and Sejnowski, 2009). Among these, the two most common models are when gamma arises from reciprocal coupling between pyramidal cells and inhibitory interneurons (PING), and recurrent connections among inhibitory interneurons (ING) (Whittington et al., 2000; Tiesinga and Sejnowski, 2009). In both, oscillatory activity arises from the structure of local connectivity. In our work, we identified another motif by which gamma oscillations can arise—**T**op-down control of **I**nhibitory **N**euron **G**amma (TING). Local excitatory mitral and tufted cells broadcast activity patterns to a pyramidal/semilunar cell population in piriform cortex, that then synapses back onto inhibitory granule cells.

Studies on dynamics of local excitatory and inhibitory neurons in the olfactory system both experimentally and mathematically are extensive (Wilson and Cowan, 1972; Ermentrout and Kopell, 1990; Kay et al., 2009; Li and Cleland, 2017). To these models we add a description of how an external (in this case, top-down input from piriform cortex) source controlling the inhibitory neuron population can influence dynamics. In studying the dynamical system defined by this network, we found that the bifurcations largely result from the singularity (linearized Jacobian matrix is non-hyperbolic) inherently embedded in the system itself. Thus, although the exact parameter values (defined by the weights of connections) influence when the dynamics of the network undergoes a bifurcation, the types of bifurcations that arise are determined by the normal form of the system (Guckenheimer and Holmes, 2013; Kuznetsov, 2013); revealing that the behaviors observed in this three node population are a fundamental feature of the network architecture. For Wilson-Cowan equations we used, the Bogdanov–Takens bifurcation is the inherent codimension-2 singularity (Cowan et al., 2016), meaning that the diversity of dynamics exists for a broad range of parameter settings, and that the unfolding of these dynamics can be implemented by modulating the top-down connection weight. Our model address this in the context of olfaction (Oswald and Urban, 2012a), but it may be applicable to a number of other sensory systems that share a similar architecture. For instance, the axonal projections from the cingulate of frontal cortex to GABAergic inhibitory neurons in V1 of the mouse visual system are organized (Zhang et al., 2014), and may therefore serve an analogous function as piriform projections to granule cells. Consequently, we identified a generalized principle by which control of inhibition via top-down weights can support a number of computations essential for neural circuit function.

Finally, we found that the firing rate representations of mitral/tufted cells, granule cells, and piriform neurons resided within distinct domains on a manifold defined by the stimulus and the weight of feedback. These domains corresponded to transitions in the dynamics of the system. Changes in the top-down weights moved a transition boundary that delineating these domains across different stimuli. When two stimuli were on opposites sides of this transition boundary, their dynamics operated under two different regimes, and their representations were pushed further apart. By contrast, when the stimuli were both on the same side of the transition boundary, within regimes corresponding to similar dynamics, their activity became more synchronous; effectively binding those stimuli together. Changes in top-down weight were therefore changes in the location of the transition boundary that could either marshal the representations of two stimuli together or push them apart. In conclusion, we

identified a model that links the dynamics of neural systems with the computations they are hypothesized to perform and may be used as a generalized framework to study the diverse effects of feedback onto inhibitory populations.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

KP conceived and supervised the project. ZC performed all the experiments and analysis. ZC and KP made the figures and wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by funding from the National Institutes of Health (NIH) and the National Science Foundation (NSF). KP was funded by NIH R01 MH113924, NSF CAREER 1749772, the Cystinosis Research Foundation, and the Kilian J. and Caroline F. Schmitt Foundation. This manuscript has been released as a pre-print at https://www.biorxiv.org/content/10.1101/2020.02.25.964965v1 (Chen and Padmanabhan, 2020).

REFERENCES

Abraham, N. M., Egger, V., Shimshek, D. R., Renden, R., Fukunaga, I., Sprengel, R., et al. (2010). Synaptic inhibition in the olfactory bulb accelerates odor discrimination in mice. *Neuron* 65, 399–411. doi: 10.1016/j.neuron.2010.01.009

Arenkiel, B. R., Hasegawa, H., Yi, J. J., Larsen, R. S., Wallace, M. L., Philpot, B. D., et al. (2011). Activity-induced remodeling of olfactory bulb microcircuits revealed by monosynaptic tracing. *PLoS ONE* 6:e29423. doi: 10.1371/journal.pone.0029423

Boyd, A. M., Sturgill, J. F., Poo, C., and Isaacson, J. S. (2012). Cortical feedback control of olfactory bulb circuits. *Neuron* 76, 1161–1174. doi: 10.1016/j.neuron.2012.10.020

Brea, J. N., Kay, L. M., and Kopell, N. J. (2009). Biophysical model for gamma rhythms in the olfactory bulb via subthreshold oscillations. *Proc. Natl. Acad. Sci. U.S.A.* 106, 21954–21959. doi: 10.1073/pnas.0910964106

Buzsaki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929. doi: 10.1126/science.1099745

Cardin, J. A., Carlén, M., Meletis, K., Knoblich, U., Zhang, F., Deisseroth, K., et al. (2009). Driving fast-spiking cells induces gamma rhythm and controls sensory responses. *Nature* 459, 663–667. doi: 10.1038/nature08002

Carey, R. M., and Wachowiak, M. (2011). Effect of sniffing on the temporal structure of mitral/tufted cell output from the olfactory bulb. *J. Neurosci.* 31, 10615–10626. doi: 10.1523/JNEUROSCI.1805-11.2011

Cauthron, J. L., and Stripling, J. S. (2014). Long-term plasticity in the regulation of olfactory bulb activity by centrifugal fibers from piriform cortex. *J. Neurosci.* 34, 9677–9687. doi: 10.1523/JNEUROSCI.1314-14.2014

Cayco-Gajic, N. A., and Silver, R. A. (2019). Re-evaluating circuit mechanisms underlying pattern separation. *Neuron* 101, 584–602. doi: 10.1016/j.neuron.2019.01.044

Chen, Z., and Padmanabhan, K. (2020). Top-down control of inhibition reshapes neural dynamics giving rise to a diversity of computations. *bioRxiv [Preprint]*. doi: 10.1101/2020.02.25.964965

Cleland, T. A., and Linster, C. (2005). Computation in the olfactory system. *Chem. Sens.* 30, 801–813. doi: 10.1093/chemse/bji072

Cowan, J. D., Neuman, J., and van Drongelen, W. (2016). Wilson–cowan equations for neocortical dynamics. *J. Math. Neurosci.* 6, 1–24. doi: 10.1186/s13408-015-0034-5

Davison, I. G., and Ehlers, M. D. (2011). Neural circuit mechanisms for pattern detection and feature combination in olfactory cortex. *Neuron* 70, 82–94. doi: 10.1016/j.neuron.2011.02.047

Deshpande, A., Bergami, M., Ghanem, A., Conzelmann, K.-K., Lepier, A., Götz, M., et al. (2013). Retrograde monosynaptic tracing reveals the temporal evolution of inputs onto new neurons in the adult dentate gyrus and olfactory bulb. *Proc. Natl. Acad. Sci. U.S.A.* 110, E1152–E1161. doi: 10.1073/pnas.1218991110

Egger, V., Svoboda, K., and Mainen, Z. F. (2005). Dendrodendritic synaptic signals in olfactory bulb granule cells: local spine boost and global low-threshold spike. *J. Neurosci.* 25, 3521–3530. doi: 10.1523/JNEUROSCI.4746-04.2005

Ermentrout, G. B., and Kopell, N. (1990). Oscillator death in systems of coupled neural oscillators. *SIAM J. Appl. Math.* 50, 125–146. doi: 10.1137/0150009

Franci, A., Herrera-Valdez, M. A., Lara-Aparicio, M., and Padilla-Longoria, P. (2018). Synchronization, oscillator death, and frequency modulation in a class of biologically inspired coupled oscillators. *Front. Appl. Math. Stat.* 4:51. doi: 10.3389/fams.2018.00051

Franks, K. M., and Isaacson, J. S. (2006). Strong single-fiber sensory inputs to olfactory cortex: implications for olfactory coding. *Neuron* 49, 357–363. doi: 10.1016/j.neuron.2005.12.026

Friedrich, R. W., and Laurent, G. (2001). Dynamic optimization of odor representations by slow temporal patterning of mitral cell activity. *Science* 291, 889–894. doi: 10.1126/science.291.5505.889

Galan, R. F., Fourcaud-Trocme, N., Ermentrout, G. B., and Urban, N. N. (2006). Correlation-induced synchronization of oscillations in olfactory bulb neurons. *J. Neurosci.* 26, 3646–3655. doi: 10.1523/JNEUROSCI.4605-05.2006

Guckenheimer, J., and Holmes, P. (2013). *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. New York, NY: Springer Science & Business Media.

Hasenstaub, A., Shu, Y., Haider, B., Kraushaar, U., Duque, A., and McCormick, D. A. (2005). Inhibitory postsynaptic potentials carry synchronized frequency information in active cortical networks. *Neuron* 47, 423–435. doi: 10.1016/j.neuron.2005.06.016

Kapoor, V., Provost, A. C., Agarwal, P., and Murthy, V. N. (2016). Activation of raphe nuclei triggers rapid and distinct effects on parallel olfactory bulb output channels. *Nat. Neurosci.* 19, 271–282. doi: 10.1038/nn.4219

Kapoor, V., and Urban, N. N. (2006). Glomerulus-specific, long-latency activity in the olfactory bulb granule cell network. *J. Neurosci.* 26, 11709–11719. doi: 10.1523/JNEUROSCI.3371-06.2006

Kay, L. M., Beshel, J., Brea, J., Martin, C., Rojas-Libano, D., and Kopell, N. (2009). Olfactory oscillations: the what, how and what for. *Trends Neurosci.* 32, 207–214. doi: 10.1016/j.tins.2008.11.008

Kuznetsov, Y. A. (2013). *Elements of Applied Bifurcation Theory*. New York, NY: Springer Science & Business Media.

Ledoux, E., and Brunel, N. (2011). Dynamics of networks of excitatory and inhibitory neurons in response to time-dependent inputs. *Front. Comput. Neurosci.* 5:25. doi: 10.3389/fncom.2011.00025

Leutgeb, J. K., Leutgeb, S., Moser, M.-B., and Moser, E. I. (2007). Pattern separation in the dentate gyrus and CA3 of the hippocampus. *Science* 315, 961–966. doi: 10.1126/science.1135801

Li, G., and Cleland, T. A. (2017). A coupled-oscillator model of olfactory bulb gamma oscillations. *PLoS Comput. Biol.* 13:e1005760. doi: 10.1371/journal.pcbi.1005760

Lledo, P.-M., Alonso, M., and Grubb, M. S. (2006). Adult neurogenesis and functional plasticity in neuronal circuits. *Nat. Rev. Neurosci.* 7, 179–193. doi: 10.1038/nrn1867

Markopoulos, F., Rokni, D., Gire, D. H., and Murthy, V. N. (2012). Functional properties of cortical feedback projections to the olfactory bulb. *Neuron* 76, 1175–1188. doi: 10.1016/j.neuron.2012.10.028

Miura, K., Mainen, Z. F., and Uchida, N. (2012). Odor representations in olfactory cortex: distributed rate coding and decorrelated population activity. *Neuron* 74, 1087–1098. doi: 10.1016/j.neuron.2012.04.021

Miyamichi, K., Amat, F., Moussavi, F., Wang, C., Wickersham, I., Wall, N. R., et al. (2011). Cortical representations of olfactory input by trans-synaptic tracing. *Nature* 472, 191–196. doi: 10.1038/nature09714

Nunes, D., and Kuner, T. (2015). Disinhibition of olfactory bulb granule cells accelerates odour discrimination in mice. *Nat. Commun.* 6:8950. doi: 10.1038/ncomms9950

Nusser, Z., Kay, L. M., Laurent, G., Homanics, G. E., and Mody, I. (2001). Disruption of GABAA receptors on GABAergic interneurons leads to increased oscillatory power in the olfactory bulb network. *J. Neurophysiol.* 86, 2823–2833. doi: 10.1152/jn.2001.86.6.2823

Oswald, A. M., and Urban, N. N. (2012a). There and back again: the corticobulbar loop. *Neuron* 76, 1045–1047. doi: 10.1016/j.neuron.2012.12.006

Oswald, A. M. M., and Urban, N. N. (2012b). Interactions between behaviorally relevant rhythms and synaptic plasticity alter coding in the piriform cortex. *J. Neurosci.* 32, 6092–6104. doi: 10.1523/JNEUROSCI.6285-11.2012

Padmanabhan, K., Osakada, F., Tarabrina, A., Kizer, E., Callaway, E. M., Gage, F. H., et al. (2016). Diverse representations of olfactory information in centrifugal feedback projections. *J. Neurosci.* 36, 7535–7545. doi: 10.1523/JNEUROSCI.3358-15.2016

Padmanabhan, K., Osakada, F., Tarabrina, A., Kizer, E., Callaway, E. M., Gage, F. H., et al. (2019). Centrifugal inputs to the main olfactory bulb revealed through whole brain circuit-mapping. *Front. Neuroanat.* 12:115. doi: 10.3389/fnana.2018.00115

Petzold, G. C., Hagiwara, A., and Murthy, V. N. (2009). Serotonergic modulation of odor input to the mammalian olfactory bulb. *Nat. Neurosci.* 12, 784–791. doi: 10.1038/nn.2335

Shipley, M. T., and Adamek, G. D. (1984). The connections of the mouse olfactory bulb: a study using orthograde and retrograde transport of wheat germ agglutinin conjugated to horseradish peroxidase. *Brain Res. Bull.* 12, 669–688. doi: 10.1016/0361-9230(84)90148-5

Shusterman, R., Smear, M. C., Koulakov, A. A., and Rinberg, D. (2011). Precise olfactory responses tile the sniff cycle. *Nat. Neurosci.* 14:1039. doi: 10.1038/nn.2877

Sohal, V. S., Zhang, F., Yizhar, O., and Deisseroth, K. (2009). Parvalbumin neurons and gamma rhythms enhance cortical circuit performance. *Nature* 459, 698–702. doi: 10.1038/nature07991

Stettler, D. D., and Axel, R. (2009). Representations of odor in the piriform cortex. *Neuron* 63, 854–864. doi: 10.1016/j.neuron.2009.09.005

Suzuki, N., and Bekkers, J. M. (2011). Two layers of synaptic processing by principal neurons in piriform cortex. *J. Neurosci.* 31, 2156–2166. doi: 10.1523/JNEUROSCI.5430-10.2011

Tiesinga, P., and Sejnowski, T. J. (2009). Cortical enlightenment: are attentional gamma oscillations driven by ING or PING? *Neuron* 63, 727–732. doi: 10.1016/j.neuron.2009.09.009

Tsodyks, M. V., Skaggs, W. E., Sejnowski, T. J., and McNaughton, B. L. (1997). Paradoxical effects of external modulation of inhibitory interneurons. *J. Neurosci.* 17, 4382–4388. doi: 10.1523/JNEUROSCI.17-11-04382.1997

Uchida, N., and Mainen, Z. F. (2003). Speed and accuracy of olfactory discrimination in the rat. *Nat. Neurosci.* 6, 1224–1229. doi: 10.1038/nn1142

Urban, N. N., and Sakmann, B. (2002). Reciprocal intraglomerular excitation and intra- and interglomerular lateral inhibition between mouse olfactory bulb mitral cells. *J. Physiol.* 542, 355–367. doi: 10.1113/jphysiol.2001.013491

Veltz, R., and Sejnowski, T. J. (2015). Periodic forcing of inhibition-stabilized networks: nonlinear resonances and phase-amplitude coupling. *Neural. Comput.* 27, 2477–2509. doi: 10.1162/NECO_a_00786

Wachowiak, M. (2011). All in a sniff: olfaction as a model for active sensing. *Neuron* 71, 962–973. doi: 10.1016/j.neuron.2011.08.030

Wesson, D. W., Carey, R. M., Verhagen, J. V., and Wachowiak, M. (2008). Rapid encoding and perception of novel odors in the rat. *PLoS Biol.* 6:82. doi: 10.1371/journal.pbio.0060082

Whittington, M. A., Traub, R., Kopell, N., Ermentrout, B., and Buhl, E. (2000). Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *Int. J. Psychophysiol.* 38, 315–336. doi: 10.1016/S0167-8760(00)00173-2

Whittington, M. A., Traub, R. D., and Jefferys, J. G. (1995). Synchronized oscillations in interneuron networks driven by metabotropic glutamate receptor activation. *Nature* 373, 612–615. doi: 10.1038/373612a0

Willhite, D. C., Nguyen, K. T., Masurkar, A. V., Greer, C. A., Shepherd, G. M., and Chen, W. R. (2006). Viral tracing identifies distributed columnar organization in the olfactory bulb. *Proc. Natl Acad. Sci. U.S.A.* 103, 12592–12597. doi: 10.1073/pnas.0602032103

Wilson, H. R., and Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophys. J.* 12, 1–24. doi: 10.1016/S0006-3495(72)86068-5

Yassa, M. A., and Stark, C. E. (2011). Pattern separation in the hippocampus. *Trends Neurosci.* 34, 515–525. doi: 10.1016/j.tins.2011.06.006

Zhang, S., Xu, M., Kamigaki, T., Do, J. P. H., Chang, W.-C., Jenvay, S., et al. (2014). Long-range and local circuits for top-down modulation of visual cortex processing. *Science* 345, 660–665. doi: 10.1126/science.1254126

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.

Copyright © 2020 Chen and Padmanabhan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Psychology

ORIGINAL RESEARCH
published: 14 July 2020
doi: 10.3389/fpsyg.2020.01555

The Low-Frequency Fluctuation of Trial-by-Trial Frontal Theta Activity and Its Correlation With Reaction-Time Variability in Sustained Attention

*Yao-Yao Wang[1,2,3], Li Sun[4], Yi-Wei Liu[1,2,3], Jia-Hui Pan[1,2,3], Yu-Ming Zheng[1,2,3], Yu-Feng Wang[4], Yu-Feng Zang[1,2,3] and Hang Zhang[1,2,3]**

[1] Institute of Psychological Sciences, College of Education, Hangzhou Normal University, Hangzhou, China, [2] Center for Cognition and Brain Disorders, The Affiliated Hospital of Hangzhou Normal University, Hangzhou, China, [3] Zhejiang Key Laboratory for Research in Assessment of Cognitive Impairments, Hangzhou, China, [4] Institute of Mental Health, The Sixth Hospital, Peking University, Beijing, China

OPEN ACCESS

Edited by:
Daya Shankar Gupta, Camden County College, United States

Reviewed by:
Pietro Aricò, Sapienza University of Rome, Italy
Ariel Telpaz, General Motors, United States

***Correspondence:**
Hang Zhang
kevinhangbnu@foxmail.com

Specialty section:
This article was submitted to Cognitive Science, a section of the journal Frontiers in Psychology

Received: *22 February 2020*
Accepted: *10 June 2020*
Published: *14 July 2020*

Citation:
Wang Y-Y, Sun L, Liu Y-W, Pan J-H, Zheng Y-M, Wang Y-F, Zang Y-F and Zhang H (2020) The Low-Frequency Fluctuation of Trial-by-Trial Frontal Theta Activity and Its Correlation With Reaction-Time Variability in Sustained Attention. Front. Psychol. 11:1555. doi: 10.3389/fpsyg.2020.01555

Reaction-time variability is a critical index of sustained attention. However, researchers still lack effective measures to establish the association between neurophysiological activity and this behavioral variability. Here, the present study recorded reaction time (RT) and cortical electroencephalogram (EEG) in healthy subjects when they continuously performed an alternative responding task. The frontal theta activity and reaction-time variability were examined trial by trial using the measures of standard deviation (SD) in the time domain and amplitude of low-frequency fluctuation (ALFF) in the frequency domain. Our results showed that the SD of reaction-time variability did not have any correlation with the SD of trial-by-trial frontal theta activity, and the ALFF of reaction-time variability has a significant correlation with the ALFF of trial-by-trial frontal theta activity in 0.01–0.027 Hz. These results suggested the methodological significance of ALFF in establishing the association between neurophysiological activity and reaction-time variability. Furthermore, these findings also support the low-frequency fluctuation as a potential feature of sustained attention.

Keywords: sustained attention, EEG, frontal theta activity, reaction-time variability, trial-by-trial fluctuation, frequency-dependent fluctuation

INTRODUCTION

The capacity of sustained attention is of great importance. It refers to focusing on a certain task for a long period of time (Posner et al., 2014). Many occupations in our daily lives require a high level of sustained attention, e.g., driving vehicles, industrial control, and air traffic control (Aricò et al., 2016; Reinerman-Jones et al., 2016; Sebastiani et al., 2020). Declined sustained attention was documented in studies of several physiological states, e.g., alcoholism (Coles et al., 2002), sleep deprivation (Gunzelmann et al., 2009), and fatigue (Gunzelmann et al., 2011). Moreover, the deficit of sustained attention is usually identified as symptoms of various neuropsychiatric disorders,

e.g., attention-deficit hyperactivity disorder (ADHD) (Barkley, 1997), and autism spectrum disorders (ASD) (Christakou et al., 2013). The practical importance of sustained attention, therefore, attracted interests from the research community, and numerous studies have been devoted to the behavioral and neurophysiological explorations of sustained attention.

Behavioral studies on sustained attention always employ examine stimulus-response tasks, e.g., alternative responding task (Helps et al., 2010), Go/NoGo task (Kirmizi-Alsan et al., 2006), and Eriksen flanker task (Castellanos et al., 2005). Subjects performing these tasks were requested to continuously detect stimulus, and their behavioral data of response and reaction time (RT) were recorded simultaneously. Several measures of the behavioral data were used to assess the level of sustained attention, e.g., error rate and intraindividual variability of RT. The error rate did not exhibit good test–retest reliability, and it was suggested to be more related with the response strategy (Liu et al., 2017; Steinborn et al., 2018). Therefore, the measure of intraindividual variability of RT was employed more extensively. For the individual, RT variability could be assessed by calculating the standard deviation of the reaction time (RT-SD) (Castellanos et al., 2005; Flehmig et al., 2007), and this measure in some studies was also standardized with the mean value of reaction time (RT-Mean) (Epstein et al., 2011; Tamm et al., 2012). RT-SD has good test–retest reliability (Liu et al., 2017), and their functional significance has been demonstrated by many clinical investigations. Observation from these investigations confirmed the linking between increased RT variability and the symptoms of sustained attention deficit, e.g., ADHD (Di Martino et al., 2008; Epstein et al., 2011), bipolar disorder (Brotman et al., 2009), and traumatic brain injury (TBI) (Burton et al., 2002). This functional significance of RT variability not only attracted increasing number of clinical trials (Rosch et al., 2013; James et al., 2016; Salum et al., 2019) but also promoted the neurophysiological explorations.

Cortical electroencephalogram (EEG) explorations have contributed many insightful evidences for the neurophysiological underpinning of sustained attention. Studies examining EEG identified several rhythmic activities, e.g., theta (4–8 Hz), alpha (8–14 Hz), and beta (14–30 Hz) (Buzsaki, 2006; Clayton et al., 2015), and some experimental evidences suggested that these EEG rhythmic activities were in response to distinct cognitive processes of sustained attention. Alpha was mainly in response to the inhibition of task-irrelevant processes (Klimesch et al., 2007; Uusberg et al., 2013), and beta could affect the attentional engagement to stimulus (Oswal et al., 2012; Coelli et al., 2015; Li et al., 2017; Shapiro et al., 2017). Theta was suggested to play a role in attentional control (Cohen and Donner, 2013; Cavanagh and Frank, 2014) and action monitoring process (Cavanagh et al., 2012; Pandey et al., 2016). Furthermore, studies attempt to establish the association between EEG rhythmic activity and behavior, and theta has been proven to be an important target (Helfrich et al., 2018; Fiebelkorn and Kastner, 2019). Increased theta activity, specifically the theta activity of the frontal brain regions, was identified in the conditions requiring higher level of sustained attention (Jensen and Tesche, 2002; Mitchell et al., 2008; Sauseng et al., 2010). More importantly, several studies reported that frontal theta activity successfully predicted the behavior during attention-demanding tasks and was related to prolonged task performance (Clayton et al., 2015; Helfrich et al., 2018). It was observed that increased frontal theta activity was in response of greater error rates and prolonged RTs (Boksem et al., 2005; Cavanagh and Shackman, 2015; Cooper et al., 2019); nevertheless, these results were not consistent across studies (Loo et al., 2004; van Driel et al., 2012; Wascher et al., 2014). Similar to RT variability, continuous recording of EEG activity typically exhibited a pattern of waxing and waning across trials of tests (Ratcliff et al., 2009; VanRullen et al., 2011). Inspired by this, recent studies explored the trial-by-trial fluctuation of EEG activity (Truccolo et al., 2002; Fox et al., 2006; McLoughlin et al., 2014; Adamo et al., 2015). Studies used the measure of SD to assess the frontal theta activity across different trials and directly examined the correlation between the SD of trial-by-trial frontal theta activity and RT-SD; however, they did not identify any significant results (McLoughlin et al., 2014).

It was worthy to note that an increasing number of studies documented the periodical variability of RT in sustained attention (Castellanos et al., 2005; Di Martino et al., 2008; Helps et al., 2011). This observation was firstly reported by Castellanos et al. (2005). They employed frequency-dependent analyses to examine RT fluctuation of children with ADHD when they perform a continuous stimulus–response test. It was found that the RT variability of the children with ADHD had peak amplitude around 0.05 Hz (about 20 s a cycle), and this peak amplitude of RT could be eliminated by using the medications of methylphenidate. These results, for the first time, revealed the periodical variability of RT for sustained attention. Since this periodical fluctuation was observed in the relatively low-frequency band (<0.2 Hz, >5 s/cycle), some studies employed the term amplitude of low-frequency fluctuation to depict this measure of reaction time (RT-ALFF). RT-ALFF as a measure of the frequency domain exhibited good test–retest reliability (ICC = 0.69) (Liu et al., 2017), and it has been repeatedly employed as a validity measure in examining the subjects with sustained attention deficits (Johnson et al., 2007; Vaurio et al., 2009; Karalunas et al., 2013). Studies usually investigated RT-ALFF in several frequency bands, including <0.01, 0.01–0.027, 0.027–0.073, and 0.073–0.167 Hz (Di Martino et al., 2008; Adamo et al., 2014, 2015). These frequency bands are defined by Penttonen and Buzsáki (2003) based on specific properties and physiological function. It has been found that RT-ALFF in three specific frequency bands (0.01–0.027 Hz; 0.027–0.073 Hz; 0.073–0.20 Hz) was strongly related to the ratings of ADHD symptoms (Mairena et al., 2012). These behavioral findings suggested that low-frequency fluctuation may be a feature of sustained attention, which encouraged us to investigate the ALFF of trial-by-trial frontal theta activity. We expected that the ALFF of trial-by-trial frontal theta activity could exhibit association with the ALFF of RT variability.

To verify this view, we performed an exploratory investigation. Data of RT and EEG activity were recorded simultaneously when subjects performing a sustained attention test. Trial-by-trial fluctuation of frontal theta activity and RT variability were

examined with the measures of SD and ALFF (in the frequency bands of 0.01–0.027, 0.027–0.073, and 0.073–0.167 Hz). Then, EEG-behavior correlations were further assessed in each frequency band, respectively.

MATERIALS AND METHODS

Participants

Seventy healthy participants (37 females; 21 ± 2 years old) were recruited in this study. All of the participants were right-handed, and no individual reported any history of brain injury or mental disorders. The RT data and EEG recording were collected from all participants during a continuously performed test, i.e., alternative responding task. According to the inclusion criteria of previous studies (Geurts et al., 2008; Liu et al., 2017), 4 participants (3 females) showing extreme mean RT values (longer than 3 SDs beyond the mean RT for all the participants) were excluded from subsequent analyses. At last, data from 66 participants (34 females; 21 ± 2 years old) were included in this study. All participants gave written informed consent prior to their participation. The study was approved by the Center for Cognition and Brain Disorders (CCBD) Ethics Committee of Hangzhou Normal University.

Experimental Paradigm

The whole experiment involves two sessions of test, i.e., resting session and sustained attention task session, and the order of the two sessions was counterbalanced across all participants. Since the present study focused on trial-by-trial fluctuation, data from the resting session were not involved in the analysis. In the sustained attention task session, all participants performed the alternative responding task for 8 min (Helps et al., 2010). In the task, two kinds of stimuli, i.e., ">" and "<," were pseudo-randomly presented for 500 ms interleaved with a fixation cross, and the inter-trial-interval (ITI) was 3000 ms. Each participant was instructed to determine the direction (left or right) of the arrow and to press "F" or "J" on the computer keyboard with the index finger. Before this session, a practice involving six trials was employed to ensure each participant was familiar with the procedure of the task.

Electrophysiological Recording and Preprocessing

The EEG data were recorded using a 32-channel (Brain Products, Germany) extended 10–20 system montage. The original recording reference was positioned at FCz. A sampling rate of 500 Hz was used. The filter bandwidth at recording was 0.016–250 Hz. All impedances were kept below 5 kΩ. The EEG data from the task session was preprocessed using Vision Analyzer software (Brain Products, Germany). Specifically, the channel signals were firstly re-referenced to average reference. After applying a notch filter (50 Hz) and band pass filtering (0.1–70 Hz), eye movement artifacts were corrected using independent component analysis (ICA, Jung et al., 2000). Moreover, the stimulus-locked segments ranged from 0 to 3000 ms according to the ITI of the task. If the amplitude of the EEG data exceeded ± 100 mV at any electrode, a segment of 3 s around this artifact was excluded from further analyses; lastly, an average number of 158 (range from 137 to 160) artifact-free trials out of a total of 160 trials were available.

Data Analysis

Behavioral Data Analyses

Behavioral measures, including several conventional measures and RT-ALFF, were calculated. The conventional measures including RT-Mean, RT-SD, and error rate (including commission error and omission error) were firstly acquired from each participant. Then, RT-ALFF was assessed through the following steps: (1) Missing and anticipatory responses (RT < 100 ms) were interpolated by linear interpolation to reconstruct an integrated time series. (2) The RT time series (divided by RT-Mean) were transformed from the time domain to the frequency domain through fast Fourier transformation (FFT), and the amplitude at each frequency point was obtained. (3) RT-ALFF was calculated as the mean amplitude in a fixed frequency band. The examinable frequency band was 0.002–0.167 Hz according to Nyquist's sampling theorem (sampling rate is 0.33 Hz corresponding to the ITI of the task). RT-ALFF was calculated in three sub-frequency bands including 0.01–0.027, 0.027–0.073, and 0.073–0.167 Hz, which were widely explored in previous studies (Di Martino et al., 2008; Adamo et al., 2015).

Electrophysiological Data Analyses

Electroencephalogram recording of the whole testing was analyzed through FFT, and the theta activity (4–8 Hz) mainly located in the frontal area was acquired at the electrodes of F3, FZ, and F4, respectively (**Figures 1A,B**). In the same way, trial-by-trial theta activity of each electrode was acquired based on the EEG recording of each single trial (each trial lasting 3000 ms) (**Figure 1C**).

Trial-by-trial theta activity (F3, FZ, and F4) was examined through two measures, Theta-SD and Theta-ALFF. For each frontal electrode, Theta-SD was calculated as the SD value of theta activity across all available trials. Theta-ALFF for each frontal electrode was assessed using the same analysis procedure of RT-ALFF (see details in **Figure 2**). Concretely, the theta activity of missing trials was replaced using linear interpolation between the theta activity of adjacent trials. Then, the time series of theta activity were divided by the mean value and were further transformed from the time domain to the frequency domain through FFT. At last, the average amplitude for a fixed frequency band was calculated as Theta-ALFF. Here, Theta-ALFF was examined in three frequency bands, i.e., 0.01–0.027, 0.027–0.073, and 0.073–0.167 Hz, corresponding to the analysis procedure of RT-ALFF. All of the analyses were implemented through our own MATLAB code.

Behavior-EEG Correlation Analyses

The correlations between behavioral measures and EEG measures were examined using Kendall rank-correlation analysis, since the data is not normally distributed (Shapiro–Wilk test, p-values < 0.045). The extreme values (longer than 3 SDs

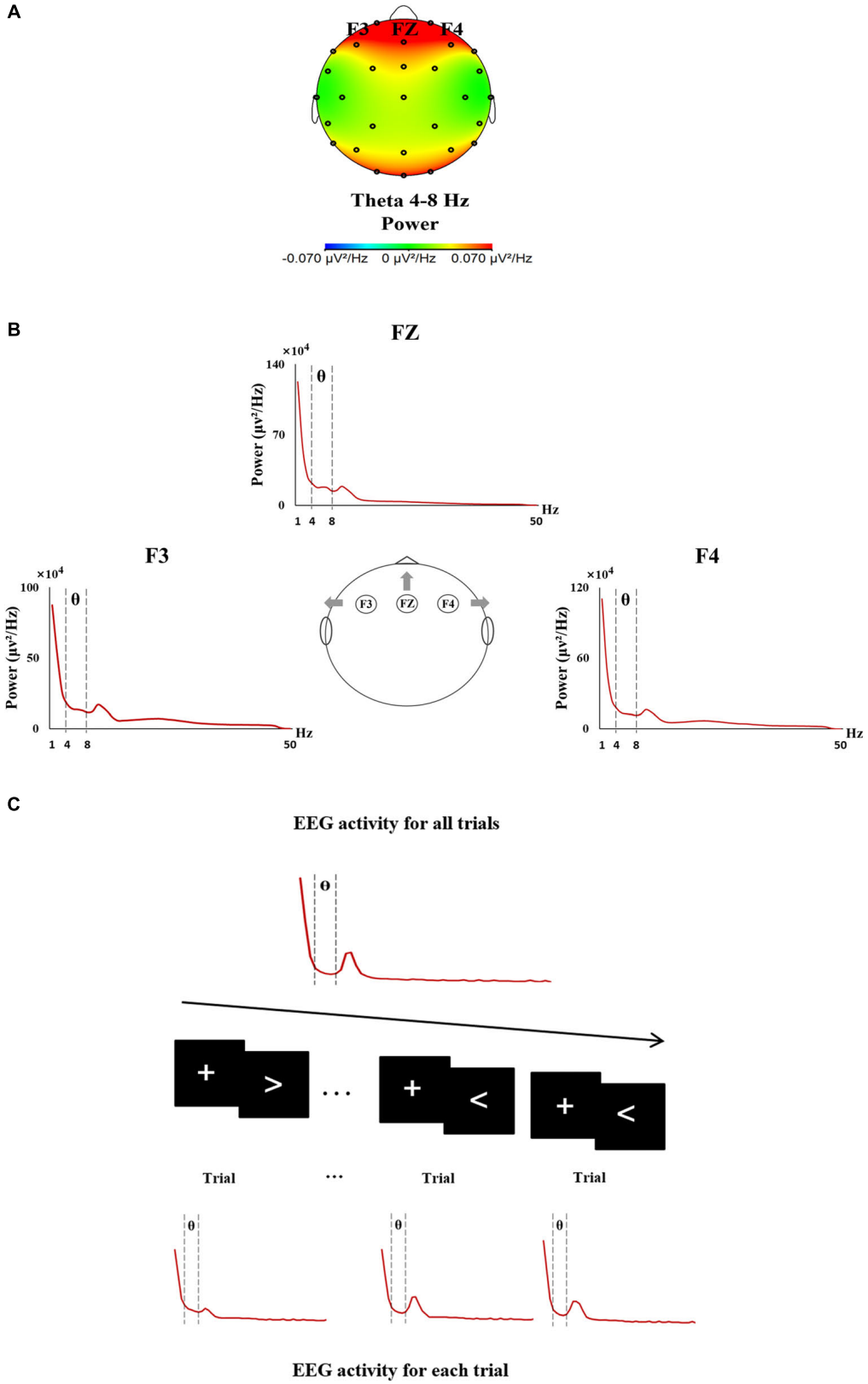

FIGURE 1 | The procedure of EEG data analyses. **(A)** EEG topographic map showing the theta activity mainly located in the frontal area during the sustained attention task. **(B)** Frontal theta EEG activity was analyzed from 4 to 8 Hz at F3, FZ, and F4. **(C)** The procedure of the trial-by-trial analysis of EEG activity.

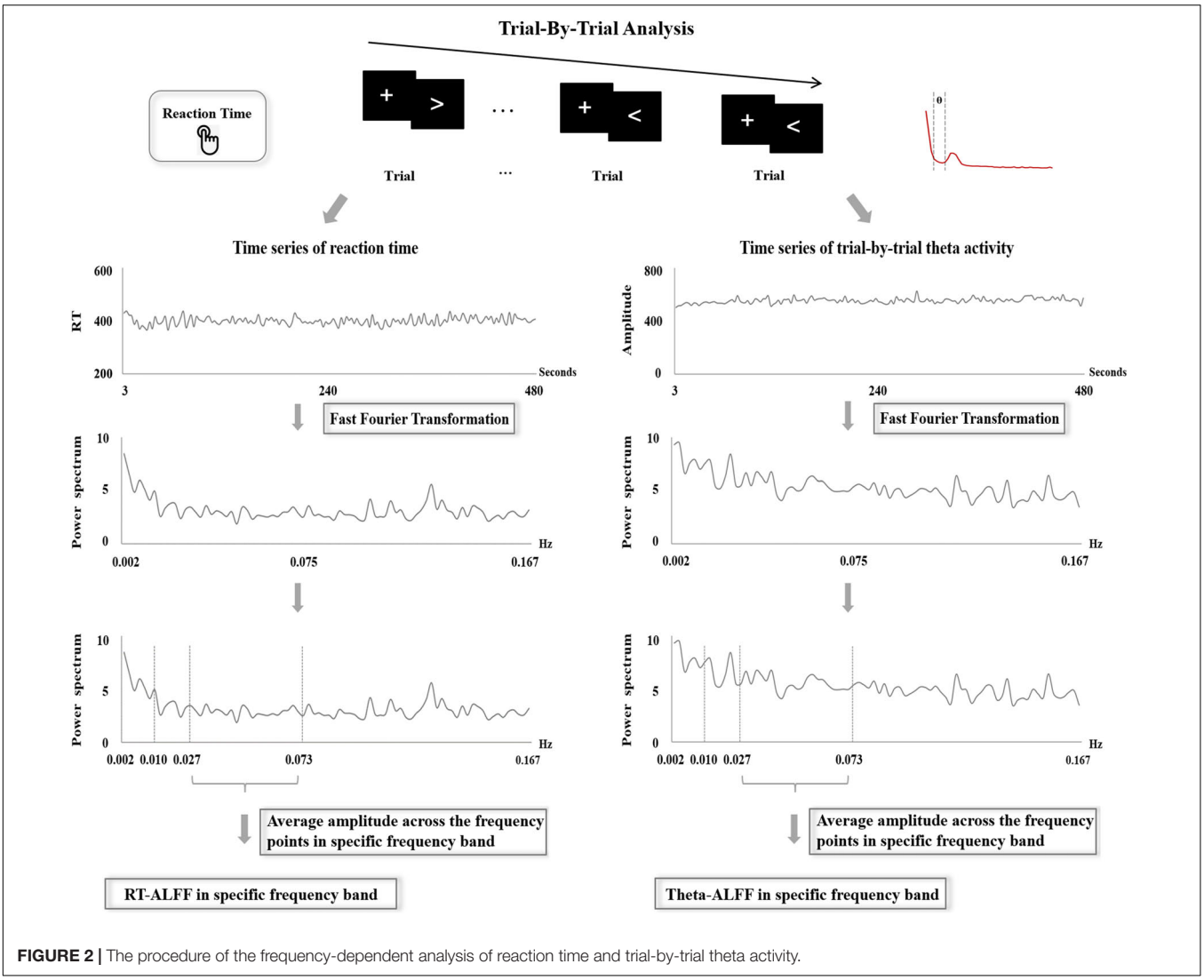

FIGURE 2 | The procedure of the frequency-dependent analysis of reaction time and trial-by-trial theta activity.

beyond the mean value of measure for all the participants) of RT data and EEG data were not involved in the analyses. The correlations between RT-SD and Theta-SD were first investigated, and then the correlations between RT-ALFF and Theta-ALFF were assessed in the frequency bands (0.01–0.027, 0.027–0.073, and 0.073–0.167 Hz). Moreover, we also examined the correlation between theta activity of whole testing and the above RT measures. All of these correlation analyses were performed using IBM SPSS Statistics (Version 20.0).

RESULTS

Behavioral Results

Descriptive statistics of conventional behavioral measures were depicted in **Table 1**, including RT-Mean, RT-SD, commission error rate, and omission error rate, and the measures of RT-ALFF in all frequency bands are shown in **Table 2**.

Correlations Between RT-SD and Theta-SD

The time series of RT are manifested in **Figure 3D**, and **Figures 3A–C** exhibited the time series of trial-by-trial theta activity at F3, FZ, and F4. The values of Theta-SD are shown in **Table 3**. The correlations between RT-SD and Theta-SD were assessed, and no significant correlation between RT-SD and Theta-SD was identified (each $r < 0.01$, $p > 0.87$) (**Table 3**). Moreover, no significant correlation between the theta activity of whole testing and RT-SD was identified (each $r < 0.05$, $p > 0.56$).

TABLE 1 | Conventional behavioral measures for the task performance.

Behavioral measures	Mean ± SD (*N* = 66)
RT-Mean (ms)	401.07 ± 39.88
RT-SD (ms)	63.87 ± 17.09
Commission error rate (%)	1.78 ± 1.75
Omission error rate (%)	0.42 ± 0.81

TABLE 2 | RT-ALFF in each frequency band for the task performance.

Frequency bands	RT-ALFF
	Mean ± SD (*N* = 66)
0.01–0.027 Hz	1.86 ± 0.61
0.027–0.073 Hz	1.69 ± 0.38
0.073–0.167 Hz	1.71 ± 0.37

TABLE 3 | Theta-SD and its correlations (Kendall rank correlation) with RT-SD.

Electrodes	Theta-SD	Correlations with RT-SD	
	Mean ± SD (*N* = 66)	*r*	*p*
F3	123.58 ± 56.95	0.01	0.87
FZ	140.04 ± 61.93	0.01	0.90
F4	120.05 ± 50.90	−0.003	0.97

Correlations Between RT-ALFF and Theta-ALFF

The frequency-dependent amplitude of reaction-time fluctuation is shown in **Figure 4D**, and **Figures 4A–C** exhibited the frequency-dependent amplitude of trial-by-trial theta activity at F3, FZ, and F4. Moreover, the values of Theta-ALFF in each frequency band are shown in **Table 4**.

As shown in **Figure 5**, a significant correlation was observed in 0.01–0.027 Hz, and RT-ALFF was negatively correlated with Theta-ALFF at F3 ($r = -0.26$, $p = 0.003$, $p < 0.05$ after Bonferroni correction for three bands across three electrodes), not at F4 and FZ (**Figure 5A**). Then, the frequency band of 0.01–0.027 Hz was divided into two sub-frequency bands, i.e., 0.01–0.019 and 0.020–0.027 Hz. Correlations between RT-ALFF and Theta-ALFF could be reserved in the two bands (0.01–0.019 Hz, $r = -0.22$, $p = 0.009$ and 0.020–0.027 Hz, $r = -0.22$, $p = 0.01$).

No other significant correlations were found between RT-ALFF and Theta-ALFF in the frequency bands of 0.027–0.073 and 0.073–0.167 Hz (each $r < 0.11$, $p > 0.18$) (**Figures 5B,C**). Moreover, no significant correlation between the theta activity of whole testing and RT-ALFF in any frequency bands was identified (each $r < 0.06$, $p > 0.46$). In addition, the trial-by-trial ALFF of delta (1–4 Hz), alpha (8–14 Hz), beta (14–30 Hz), and gamma (30–50 Hz) and its correlation with RT-ALFF in each frequency band was examined separately, and no significant correlation was observed (each $r < 0.13$, $p > 0.12$) (see details in **Supplementary Table S2**).

DISCUSSION

The present study explored the association between trial-by-trial fluctuation of frontal theta activity and RT variability. Measures

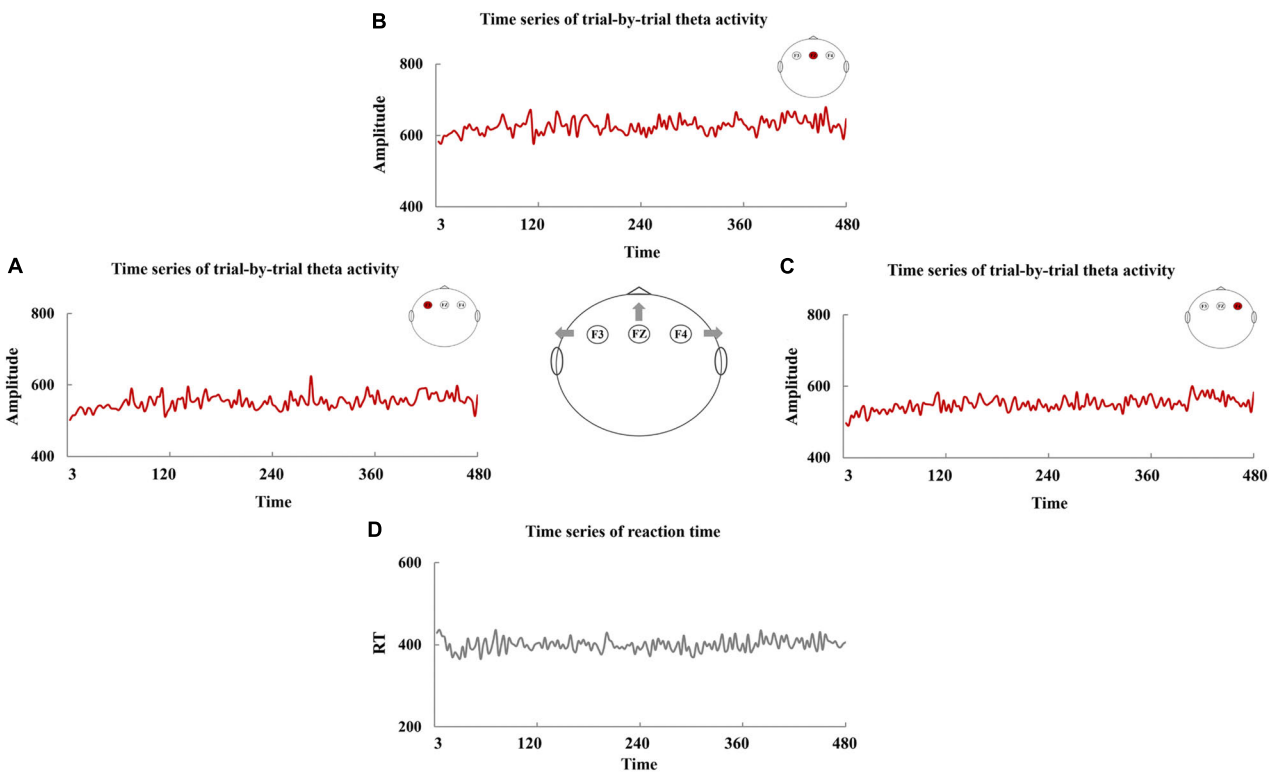

FIGURE 3 | The time series of trial-by-trial theta activity and reaction time (RT). The time series of trial-by-trial theta activity at F3 **(A)**, FZ **(B)**, and F4 **(C)** and the time series of RT **(D)**.

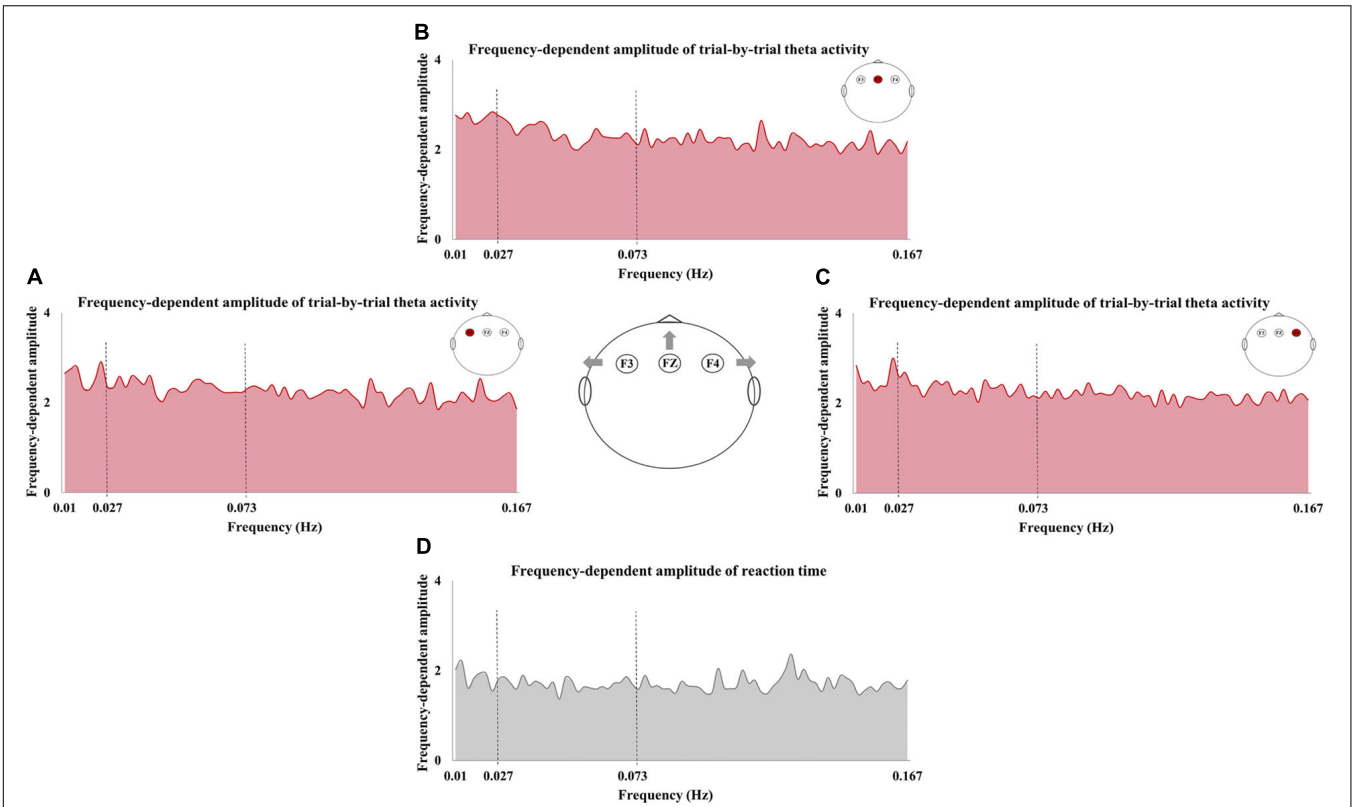

FIGURE 4 | The frequency-dependent amplitude of trial-by-trial theta activity and reaction time (RT). The amplitude of trial-by-trial theta activity at F3 **(A)**, FZ **(B)**, and F4 **(C)** and the amplitude of RT fluctuation **(D)** in the sampled frequencies (0.01–0.167 Hz).

of SD and ALFF were used to establish this association in the time domain and frequency domain, respectively. It was observed that RT variability did not show any relationship with the trial-by-trial frontal theta activity when we used SD as the measure. In contrast, a frequency-dependent correlation between RT variability and trial-by-trial frontal theta activity was revealed by the measure of ALFF. These results provided a methodological insight for future studies on the neural underpinning of sustained attention.

Reaction-time variability was considered as a critical index of the level of sustained attention (Betts et al., 2006; Flehmig et al., 2007). Albeit intensive investigation, measures for establishing the association between the RT variability and neurophysiological activity was lacking (Rabbi et al., 2009; Clayton et al., 2015; Fortenbaugh et al., 2017; O'Halloran et al., 2018). It is always failed to identify significant results when the researcher directly examined the correlation between EEG activity, e.g., beta, theta, and RT variability (Buyck and Wiersema, 2015). Similarly, the present study also did not find any significant association between theta activity at each electrode and RT variability. Experimental observations indicated that the EEG activity always fluctuates across trials in a similar way as RT variability (Ratcliff et al., 2009; VanRullen et al., 2011). Thus, recent EEG explorations highlighted the trial-by-trial fluctuation of EEG activity (Fox et al., 2006; McLoughlin et al., 2014; Adamo et al., 2015). These studies mostly focused on the frontal theta activity because experimental evidences indicated the functional role of frontal theta activity in attentional control (Cohen and Donner, 2013; Cavanagh and Frank, 2014; McLoughlin et al., 2014). Consistent with their finding, the present study did not identify any significant correlation results when using SD to assess the RT variability and trial-by-trial fluctuation of frontal theta activity (McLoughlin et al., 2014). Methodically, SD is a measure of time domain reflecting the overall fluctuation of the testing data. It is worthy to note, however, that RT variability for sustained attention may occur periodically (Castellanos et al., 2005; Di Martino et al., 2008; Helps et al., 2011). Overall fluctuation potentially masked the periodical variability. This is a possible explanation for the findings on SD, and it further encouraged us to examine the trial-by-trial frontal theta activity and RT variability in frequency domain.

TABLE 4 | Theta-ALFF in each frequency band.

Frequency bands	Theta-ALFF		
	Mean ± SD (F3, *N* = 66)	Mean ± SD (FZ, *N* = 66)	Mean ± SD (F4, *N* = 66)
0.01–0.027 Hz	2.57 ± 0.98	2.73 ± 1.09	2.54 ± 0.93
0.027–0.073 Hz	2.35 ± 0.64	2.35 ± 0.64	2.32 ± 0.66
0.073–0.167 Hz	2.17 ± 0.60	2.17 ± 0.47	2.15 ± 0.58

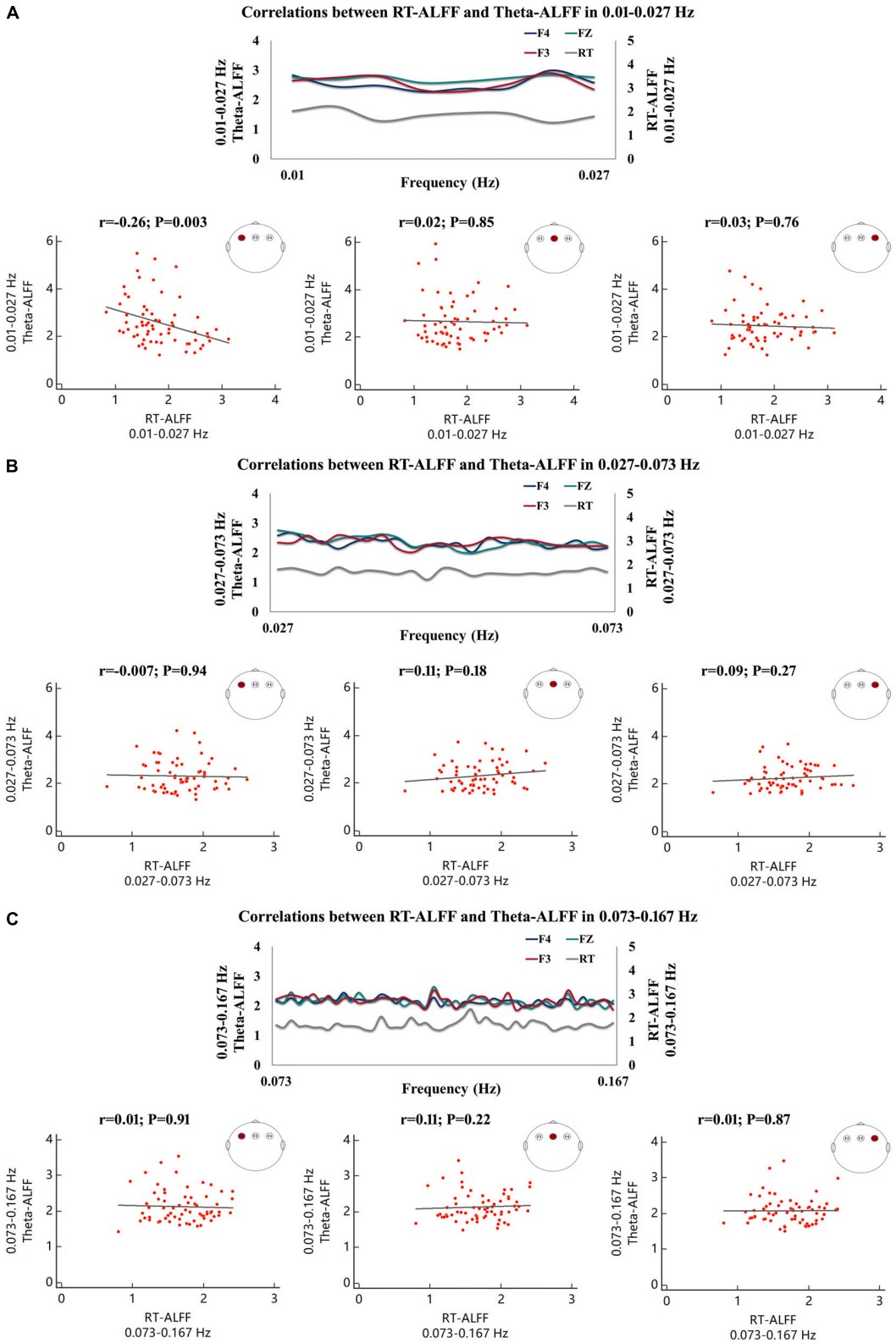

FIGURE 5 | Correlations between RT-ALFF and Theta-ALFF in each frequency band. The correlation (Kendall rank correlation) between RT-ALFF and Theta-ALFF (at the electrodes of F3, FZ, and F4) is shown in 0.01–0.027 Hz **(A)**, 0.027–0.073 Hz **(B)**, and 0.073–0.167 Hz **(C)**. The amplitude of reaction time fluctuation and trial-by-trial theta activity (at the electrodes of F3, FZ, and F4) were also depicted in the frequencies of 0.01–0.027 Hz **(A)**, 0.027–0.073 Hz **(B)**, and 0.073–0.167 Hz **(C)**.

Clinical investigations indicated that children with ADHD could show increased RT variability in the low-frequency band (<0.2 Hz) (Castellanos et al., 2005; Karalunas et al., 2013). This observation was confirmed by many studies (Johnson et al., 2007; Di Martino et al., 2008; Vaurio et al., 2009; Helps et al., 2011), and these evidences indicated that the low-frequency fluctuation was a potential feature of sustained attention. Thus, the association between trial-by-trial frontal theta activity and RT variability may be frequency-dependent. This idea was confirmed by our findings of the ALFF analyses. RT-ALFF exhibited a significant correlation with the Theta-ALFF in 0.01–0.027 Hz. Behaviorally, frequency-dependent RT variability has been linked with the attention lapse (Castellanos et al., 2005; Johnson et al., 2007; Adamo et al., 2014). It has been reported that RT-ALFF in 0.01–0.027 Hz could be used as a predictor to explain scale ratings of inattention of ADHD (Mairena et al., 2012). Here, our results may provide some preliminary evidences for understanding the neural underpinning of these behavioral observations. Nevertheless, the functional significance of the frequency-dependent fluctuation of frontal theta activity required to be further clarified. Notably, the frequency-dependent correlation in the present study had spatial specificity, and the correlation was only identified at the left frontal electrode (F3). The importance of frontal theta activity in sustained attention has been discussed (Missonnier et al., 2006; Gongora et al., 2015); however, the functional differences between the theta activity of the left and right frontal areas were less addressed in these EEG studies. It was reported that theta activity in the frontal brain areas may function in attention maintenance (Cavanagh and Frank, 2014; Wascher et al., 2014). Here, we found that the greater Theta-ALFF at the left electrode (F3) was negatively correlated with lower RT-ALFF. This finding reinforced the role of frontal theta activity in attention maintenance. Source location analysis further linked these findings with previous neuroimaging studies. The origin of the theta activity was implicated in the ventral medial frontal gyrus (vmPFC) (see details in the **Supplementary Material**), and this region serving as a role of state monitoring has been intensively reported by the functional magnetic resonance imaging (fMRI) studies on sustained attention (Ridderinkhof et al., 2004; Hasenkamp et al., 2012; Sepede et al., 2012; Clayton et al., 2015). Nevertheless, the EEG data, collected from 32 electrodes, could not provide robust localization results (Michel et al., 2004). Therefore, the linking between theta activity, vmPFC, and behavior variability should be examined with more experiments in the future. Moreover, significant correlation only appears on the left frontal electrode (F3). This spatial specificity suggested the functional differences between the theta activity of left and right areas, which should be taken into account by further explorations.

The present study sheds light on the low-frequency fluctuation of trial-by-trial EEG activity. Actually, the low-frequency fluctuation of brain activity has been documented in many previous studies. However, most of these explorations focused on fMRI but not on EEG activity (Sonuga-Barke and Castellanos, 2007; Zang et al., 2007). The findings of the present study supported the feature of low-frequency fluctuation of sustained attention, and thus, ALFF as a measure of frequency domain will subserve the establishment of the association between RT variability and EEG activity. Notably, the present study focused on EEG rhythmic activity, and we believe that it is also an interesting issue whether low-frequency fluctuation could be identified in other EEG components, e.g., event-related potential (ERP). Adamo et al. (2015) have attempted to explore this issue, and they reported that the trial-by-trial variability of P3 and RT variability coupled at 0.073–0.167 Hz when subjects performed the Go/NoGo tasks (Adamo et al., 2015). However, their findings may be confounded by many experimental factors. For example, the restriction of the signal-to-noise ratio makes it challengeable to extract robust ERP from data of a single trial. ERP is usually acquired from tasks with jitter in ITI, and the jitter results in more frequency complexity. So, this issue still needs to be examined with well-designed experiments in future studies. Moreover, the low-frequency fluctuation of trial-by-trial EEG activity, in the present study, is derived from the specific task, i.e., alternative responding task. This task has obvious advantages for frequency-dependent analysis. It has continuous response, and the time series does not need to be reconstructed statistically as is required by many non-responding trials in more complex tasks. Behavioral evidences suggested that the task complexity could induce the variation in the frequency band of fluctuation for sustained attention (Johnson et al., 2007; Di Martino et al., 2008; Helps et al., 2011; Karalunas et al., 2013; Salum et al., 2019). Therefore, further experimentation with different tasks is required to verify the band specificity of our findings.

Several limitations existed in the present study. First, the subjects were all healthy college students; further studies on the subjects with attention deficits were required to examine whether the trial-by-trial fluctuation of frontal theta activity has clinical significance. Second, the ITI of the task paradigm for the current study was 3000 ms, and thus, the frequencies that we could analyze only ranged from 0.002 to 0.167 Hz, according to the Nyquist sampling theorem. The explorations on higher frequencies (>0.167 Hz) are necessary, and tasks with fast behavior recording may be helpful for this issue. Third, our finding has spatial specificity in F3, and this electrode is spatially close to the original reference, FCz. This reference during data collection is a fixed setting for the EEG device. Therefore, it remains to be investigated whether this setting affects our findings in F3.

CONCLUSION

The present study is an exploratory investigation and, for the first time, reported that the correlation between RT variability and trial-by-trial frontal-theta activity was frequency-dependent. ALFF as a measure of the frequency domain exhibited methodological significance in establishing the association between RT variability and EEG activity. These findings supported the low-frequency fluctuation as a feature of sustained attention. Further explorations on this feature may facilitate the understanding of the neuro underpinning of sustained attention.

DATA AVAILABILITY STATEMENT

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

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by the Center for Cognition and Brain Disorders (CCBD) Ethics Committee of the Hangzhou Normal University. The patients/participants provided their written informed consent to participate in this study. Written informed consent was obtained from the individual(s) for the publication of any potentially identifiable images or data included in this article.

AUTHOR CONTRIBUTIONS

HZ conceived and designed the experiment. Y-YW, Y-WL, and J-HP collected the data. Y-YW, Y-WL, J-HP, and Y-MZ performed the data analysis. LS, Y-FW, and Y-FZ provided advice on the analysis and interpretation of the results. Y-YW and HZ wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was supported by the Key R&D Program of China (2018YFC1312600 and 2018YFC1312603), the National Natural Science Foundation of China (31471084, 81520108016, and 81771479), the National Basic Research Program of China (973 program 2014CB846104), and the Natural Science Foundation of Zhejiang Province (LY20H090006). HZ was supported by Scientific Research Staring Foundation from the Hangzhou Normal University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpsyg.2020.01555/full#supplementary-material

REFERENCES

Adamo, N., Baumeister, S., Hohmann, S., Wolf, I., Holz, N., Boecker, R., et al. (2015). Frequency-specific coupling between trial-to-trial fluctuations of neural responses and response-time variability. *J. Neural Trans.* 122, 1197–1202. doi: 10.1007/s00702-015-1382-8

Adamo, N., Di Martino, A., Esu, L., Petkova, E., Johnson, K., Kelly, S., et al. (2014). Increased response-time variability across different cognitive tasks in children with ADHD. *J. Attent. Disord.* 18, 434–446. doi: 10.1177/1087054712439419

Aricò, P., Borghini, G., Di Flumeri, G., Colosimo, A., Bonelli, S., Golfetti, A., et al. (2016). Adaptive automation triggered by EEG-based mental workload index: a passive brain-computer interface application in realistic air traffic control environment. *Front. Hum. Neurosci.* 10:539. doi: 10.3389/fnhum.2016.00539

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: constructing a unifying theory of ADHD. *Psychol. Bull.* 121:65. doi: 10.1037/0033-2909.121.1.65

Betts, J., Mckay, J., Maruff, P., and Anderson, V. (2006). The development of sustained attention in children: the effect of age and task load. *Child Neuropsychol.* 12, 205–221. doi: 10.1080/09297040500488522

Boksem, M. A., Meijman, T. F., and Lorist, M. M. (2005). Effects of mental fatigue on attention: an ERP study. *Cogn. Brain Res.* 25, 107–116. doi: 10.1016/j.cogbrainres.2005.04.011

Brotman, M. A., Rooney, M. H., Skup, M., Pine, D. S., and Leibenluft, E. (2009). Increased intrasubject variability in response time in youths with bipolar disorder and at-risk family members. *J. Am. Acad. Child Adolesc. Psychiatry* 48, 628–635. doi: 10.1097/CHI.0b013e3181a27527

Burton, C. L., Hultsch, D. F., Strauss, E., and Hunter, M. A. (2002). Intraindividual variability in physical and emotional functioning: comparison of adults with traumatic brain injuries and healthy adults. *Clin. Neuropsychol.* 16, 264–279. doi: 10.1076/clin.16.3.264.13854

Buyck, I., and Wiersema, J. R. (2015). Task-related electroencephalographic deviances in adults with attention deficit hyperactivity disorder. *Neuropsychology* 29:433. doi: 10.1037/neu0000148

Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford: Oxford University Press, doi: 10.1093/acprof:oso/9780195301069.001.0001

Castellanos, F. X., Sonuga-Barke, E. J., Scheres, A., Di Martino, A., Hyde, C., and Walters, J. R. (2005). Varieties of attention-deficit/hyperactivity disorder-related intra-individual variability. *Biol. Psychiatry* 57, 1416–1423. doi: 10.1016/j.biopsych.2004.12.005

Cavanagh, J. F., and Frank, M. J. (2014). Frontal theta as a mechanism for cognitive control. *Trends Cogn. Sci.* 18, 414–421. doi: 10.1016/j.tics.2014.04.012

Cavanagh, J. F., and Shackman, A. J. (2015). Frontal midline theta reflects anxiety and cognitive control: meta-analytic evidence. *J. Physiol. Paris* 109, 3–15. doi: 10.1016/j.jphysparis.2014.04.003

Cavanagh, J. F., Zambrano-Vazquez, L., and Allen, J. J. (2012). Theta lingua franca: a common mid-frontal substrate for action monitoring processes. *Psychophysiology* 49, 220–238. doi: 10.1111/j.1469-8986.2011.01293.x

Christakou, A., Murphy, C. M., Chantiluke, K., Cubillo, A. I., Smith, A. B., Giampietro, V., et al. (2013). Disorder-specific functional abnormalities during sustained attention in youth with attention deficit hyperactivity disorder (ADHD) and with autism. *Mol. Psychiatry* 18:236. doi: 10.1038/mp.2011.185

Clayton, M. S., Yeung, N., and Kadosh, R. C. (2015). The roles of cortical oscillations in sustained attention. *Trends Cogn. Sci.* 19, 188–195. doi: 10.1016/j.tics.2015.02.004

Coelli, S., Sclocco, R., Barbieri, R., Reni, G., Zucca, C., and Bianchi, A. M. (2015). "EEG-based index for engagement level monitoring during sustained attention," in *Proceedings of the 2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Piscataway, NJ: IEEE), 1512–1515. doi: 10.1109/embc.2015.7318658

Cohen, M. X., and Donner, T. H. (2013). Midfrontal conflict-related theta-band power reflects neural oscillations that predict behavior. *J. Neurophysiol.* 110, 2752–2763. doi: 10.1152/jn.00479.2013

Coles, C. D., Platzman, K. A., Lynch, M. E., and Freides, D. (2002). Auditory and visual sustained attention in adolescents prenatally exposed to alcohol. *Alcoholism* 26, 263–271. doi: 10.1111/j.1530-0277.2002.tb02533.x

Cooper, P. S., Karayanidis, F., McKewen, M., McLellan-Hall, S., Wong, A. S., Skippen, P., et al. (2019). Frontal theta predicts specific cognitive control-induced behavioural changes beyond general reaction time slowing. *Neuroimage* 189, 130–140. doi: 10.1016/j.neuroimage.2019.01.022

Di Martino, A., Ghaffari, M., Curchack, J., Reiss, P., Hyde, C., Vannucci, M., et al. (2008). Decomposing intra-subject variability in children with attention-deficit/hyperactivity disorder. *Biol. Psychiatry* 64, 607–614. doi: 10.1016/j.biopsych.2008.03.008

Epstein, J. N., Langberg, J. M., Rosen, P. J., Graham, A., Narad, M. E., Antonini, T. N., et al. (2011). Evidence for higher reaction-time variability for children with ADHD on a range of cognitive tasks including reward and event rate manipulations. *Neuropsychology* 25, 427. doi: 10.1037/a0022155

Fiebelkorn, I. C., and Kastner, S. (2019). A rhythmic theory of attention. *Trends Cogn. Sci.* 23, 87–101. doi: 10.1016/j.tics.2018.11.009

Flehmig, H. C., Steinborn, M., Langner, R., Scholz, A., and Westhoff, K. (2007). Assessing intraindividual variability in sustained attention: reliability, relation to speed and accuracy, and practice effects. *Psychol. Sci.* 49:132.

Fortenbaugh, F. C., DeGutis, J., and Esterman, M. (2017). Recent theoretical, neural, and clinical advances in sustained attention research. *Ann. N. Y. Acad. Sci.* 1396:70. doi: 10.1111/nyas.13318

Fox, M. D., Snyder, A. Z., Zacks, J. M., and Raichle, M. E. (2006). Coherent spontaneous activity accounts for trial-to-trial variability in human evoked brain responses. *Nat. Neurosci.* 9:23. doi: 10.1038/nn1616

Geurts, H. M., Grasman, R. P., Verté, S., Oosterlaan, J., Roeyers, H., van Kammen, S. M., et al. (2008). Intra-individual variability in ADHD, autism spectrum disorders and Tourette's syndrome. *Neuropsychologia* 46, 3030–3041. doi: 10.1016/j.neuropsychologia.2008.06.013

Gongora, M., Peressuti, C., Velasques, B., Bittencourt, J., Teixeira, S., Arias-Carrión, O., et al. (2015). Absolute Theta power in the frontal cortex during a visuomotor task: the effect of bromazepam on attention. *Clin. EEG Neurosci.* 46, 292–298. doi: 10.1177/1550059414535576

Gunzelmann, G., Gross, J. B., Gluck, K. A., and Dinges, D. F. (2009). Sleep deprivation and sustained attention performance: integrating mathematical and cognitive modeling. *Cogn. Sci.* 33, 880–910. doi: 10.1111/j.1551-6709.2009.01032.x

Gunzelmann, G., Moore, L. R., Gluck, K. A., Van Dongen, H., and Dinges, D. F. (2011). "Fatigue in sustained attention: generalizing mechanisms for time awake to time on task," in *Cognitive Fatigue: Multidisciplinary Perspectives on Current Research and Future Applications*, Ed. P. L. Ackerman (Washington, DC: American Psychological Association), doi: 10.1037/12343-004

Hasenkamp, W., Wilson-Mendenhall, C. D., Duncan, E., and Barsalou, L. W. (2012). Mind wandering and attention during focused meditation: a fine-grained temporal analysis of fluctuating cognitive states. *Neuroimage* 59, 750–760. doi: 10.1016/j.neuroimage.2011.07.008

Helfrich, R. F., Fiebelkorn, I. C., Szczepanski, S. M., Lin, J. J., Parvizi, J., Knight, R. T., et al. (2018). Neural mechanisms of sustained attention are rhythmic. *Neuron* 99, 854–865. doi: 10.1016/j.neuron.2018.07.032

Helps, S. K., Broyd, S. J., Bitsakou, P., and Sonuga-Barke, E. J. (2011). Identifying a distinctive familial frequency band in reaction time fluctuations in ADHD. *Neuropsychology* 25:711. doi: 10.1037/a0024479

Helps, S. K., Broyd, S. J., James, C. J., Karl, A., Chen, W., and Sonuga-Barke, E. J. (2010). Altered spontaneous low frequency brain activity in attention deficit/hyperactivity disorder. *Brain Res.* 1322, 134–143. doi: 10.1016/j.brainres.2010.01.057

James, S. N., Cheung, C. H., Rijsdijk, F., Asherson, P., and Kuntsi, J. (2016). Modifiable arousal in attention-deficit/hyperactivity disorder and its etiological association with fluctuating reaction times. *Biol. Psychiatry* 1, 539–547. doi: 10.1016/j.bpsc.2016.06.003

Jensen, O., and Tesche, C. D. (2002). Frontal theta activity in humans increases with memory load in a working memory task. *Eur. J. Neurosci.* 15, 1395–1399. doi: 10.1046/j.1460-9568.2002.01975.x

Johnson, K. A., Kelly, S. P., Bellgrove, M. A., Barry, E., Cox, M., Gill, M., et al. (2007). Response variability in attention deficit hyperactivity disorder: evidence for neuropsychological heterogeneity. *Neuropsychologia* 45, 630–638. doi: 10.1016/j.neuropsychologia.2006.03.034

Jung, T. P., Makeig, S., Westerfield, M., Townsend, J., Courchesne, E., and Sejnowski, T. J. (2000). Removal of eye activity artifacts from visual event-related potentials in normal and clinical subjects. *Clin. Neurophysiol.* 111, 1745–1758. doi: 10.1016/s1388-2457(00)00386-2

Karalunas, S. L., Huang-Pollock, C. L., and Nigg, J. T. (2013). Is reaction time variability in ADHD mainly at low frequencies? *J. Child Psychol. Psychiatry* 54, 536–544. doi: 10.1111/jcpp.12028

Kirmizi-Alsan, E., Bayraktaroglu, Z., Gurvit, H., Keskin, Y. H., Emre, M., and Demiralp, T. (2006). Comparative analysis of event-related potentials during Go/NoGo and CPT: decomposition of electrophysiological markers of response inhibition and sustained attention. *Brain Res.* 1104, 114–128. doi: 10.1016/j.brainres.2006.03.010

Klimesch, W., Sauseng, P., and Hanslmayr, S. (2007). EEG alpha oscillations: the inhibition–timing hypothesis. *Brain Res. Rev.* 53, 63–88. doi: 10.1016/j.brainresrev.2006.06.003

Li, Y., Kang, C., Wei, Z., Qu, X., Liu, T., Zhou, Y., et al. (2017). Beta oscillations in major depression–signalling a new cortical circuit for central executive function. *Sci. Rep.* 7, 1–15. doi: 10.1038/s41598-017-18306-w

Liu, Z. Z., Qu, H. J., Tian, Z. L., Han, M. J., Fan, Y., Ge, L. Z., et al. (2017). Reproducibility of frequency-dependent low frequency fluctuations in reaction time over time and across tasks. *PLoS One* 12:e0184476. doi: 10.1371/journal.pone.0184476

Loo, S. K., Hopfer, C., Teale, P. D., and Reite, M. L. (2004). EEG correlates of methylphenidate response in ADHD: association with cognitive and behavioral measures. *J. Clin. Neurophysiol.* 21, 457–464. doi: 10.1097/01.wnp.0000150890.14421.9a

Mairena, M. Á, Di Martino, A., Domínguez-Martín, C., Gómez-Guerrero, L., Gioia, G., Petkova, E., et al. (2012). Low frequency oscillations of response time explain parent ratings of inattention and hyperactivity/impulsivity. *Eur. Child Adolesc. Psychiatry* 21, 101–109. doi: 10.1007/s00787-011-0237-6

McLoughlin, G., Palmer, J. A., Rijsdijk, F., and Makeig, S. (2014). Genetic overlap between evoked frontocentral theta-band phase variability, reaction time variability, and attention-deficit/hyperactivity disorder symptoms in a twin study. *Biol. Psychiatry* 75, 238–247. doi: 10.1016/j.biopsych.2013.07.020

Michel, C. M., Murray, M. M., Lantz, G., Gonzalez, S., Spinelli, L., and de Peralta, R. G. (2004). EEG source imaging. *Clin. Neurophysiol.* 115, 2195–2222. doi: 10.1016/j.clinph.2004.06.001

Missonnier, P., Deiber, M. P., Gold, G., Millet, P., Pun, M. G. F., Fazio-Costa, L., et al. (2006). Frontal theta event-related synchronization: comparison of directed attention and working memory load effects. *J. Neural Trans.* 113, 1477–1486. doi: 10.1007/s00702-005-0443-9

Mitchell, D. J., McNaughton, N., Flanagan, D., and Kirk, I. J. (2008). Frontal-midline theta from the perspective of hippocampal "theta". *Progr. Neurobiol.* 86, 156–185. doi: 10.1016/j.pneurobio.2008.09.005

O'Halloran, L., Cao, Z., Ruddy, K., Jollans, L., Albaugh, M. D., Aleni, A., et al. (2018). Neural circuitry underlying sustained attention in healthy adolescents and in ADHD symptomatology. *Neuroimage* 169, 395–406. doi: 10.1016/j.neuroimage.2017.12.030

Oswal, A., Litvak, V., Sauleau, P., and Brown, P. (2012). Beta reactivity, prospective facilitation of executive processing, and its dependence on dopaminergic therapy in Parkinson's disease. *J. Neurosci.* 32, 9909–9916. doi: 10.1523/jneurosci.0275-12.2012

Pandey, A. K., Kamarajan, C., Manz, N., Chorlian, D. B., Stimus, A., and Porjesz, B. (2016). Delta, theta, and alpha event-related oscillations in alcoholics during Go/NoGo task: neurocognitive deficits in execution, inhibition, and attention processing. *Progr. Neuro Psychopharmacol. Biol. Psychiatry* 65, 158–171. doi: 10.1016/j.pnpbp.2015.10.002

Penttonen, M., and Buzsáki, G. (2003). Natural logarithmic relationship between brain oscillators. *Thalamus Relat. Syst.* 2, 145–152. doi: 10.1016/s1472-9288(03)00007-4

Posner, M. I., Rothbart, M. K., and Rueda, M. R. (2014). "Developing attention and self-regulation in childhood," in *The Oxford Handbook of Attention*, Ed. S. Kastner (Oxford: Oxford University Press), doi: 10.1093/oxfordhb/9780199675111.013.023

Rabbi, A. F., Ivanca, K., Putnam, A. V., Musa, A., Thaden, C. B., and Fazel-Rezai, R. (2009). "Human performance evaluation based on EEG signal analysis: a prospective review," in *Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society* (Piscataway, NJ: IEEE), 1879–1882. doi: 10.1109/iembs.2009.5333877

Ratcliff, R., Philiastides, M. G., and Sajda, P. (2009). Quality of evidence for perceptual decision making is indexed by trial-to-trial variability of the EEG. *Proc. Natl. Acad. Sci. U.S.A.* 106, 6539–6544. doi: 10.1073/pnas.0812589106

Reinerman-Jones, L., Matthews, G., and Mercado, J. E. (2016). Detection tasks in nuclear power plant operation: vigilance decrement and physiological workload monitoring. *Saf. Sci.* 88, 97–107. doi: 10.1016/j.ssci.2016.05.002

Ridderinkhof, K. R., Ullsperger, M., Crone, E. A., and Nieuwenhuis, S. (2004). The role of the medial frontal cortex in cognitive control. *Science* 306, 443–447. doi: 10.1126/science.1100301

Rosch, K. S., Dirlikov, B., and Mostofsky, S. H. (2013). Increased intrasubject variability in boys with ADHD across tests of motor and cognitive control. *J. Abnorm. Child Psychol.* 41, 485–495. doi: 10.1007/s10802-012-9690-z

Salum, G. A., Sato, J. R., Manfro, A. G., Pan, P. M., Gadelha, A., do Rosário, M. C., et al. (2019). Reaction time variability and attention-deficit/hyperactivity disorder: is increased reaction time variability specific to attention-deficit/hyperactivity disorder? Testing predictions from the

default-mode interference hypothesis. *ADHD* 11, 47–58. doi: 10.1007/s12402-018-0257-x

Sauseng, P., Griesmayr, B., Freunberger, R., and Klimesch, W. (2010). Control mechanisms in working memory: a possible function of EEG theta oscillations. *Neurosci. Biobehav. Rev.* 34, 1015–1022. doi: 10.1016/j.neubiorev.2009.12.006

Sebastiani, M., Di Flumeri, G., Aricò, P., Sciaraffa, N., Babiloni, F., and Borghini, G. (2020). Neurophysiological vigilance characterisation and assessment: laboratory and realistic validations involving professional air traffic controllers. *Brain Sci.* 10:48. doi: 10.3390/brainsci10010048

Sepede, G., De Berardis, D., Campanella, D., Perrucci, M. G., Ferretti, A., Serroni, N., et al. (2012). Impaired sustained attention in euthymic bipolar disorder patients and non-affected relatives: an fMRI study. *Bipolar Disord.* 14, 764–779. doi: 10.1111/bdi.12007

Shapiro, K. L., Hanslmayr, S., Enns, J. T., and Lleras, A. (2017). Alpha, beta: the rhythm of the attentional blink. *Psychonom. Bull. Rev.* 24, 1862–1869. doi: 10.3758/s13423-017-1257-0

Sonuga-Barke, E. J., and Castellanos, F. X. (2007). Spontaneous attentional fluctuations in impaired states and pathological conditions: a neurobiological hypothesis. *Neurosci. Biobehav. Rev.* 31, 977–986. doi: 10.1016/j.neubiorev.2007.02.005

Steinborn, M. B., Langner, R., Flehmig, H. C., and Huestegge, L. (2018). Methodology of performance scoring in the d2 sustained-attention test: cumulative-reliability functions and practical guidelines. *Psychol. Assess.* 30:339. doi: 10.1037/pas0000482

Tamm, L., Narad, M. E., Antonini, T. N., O'Brien, K. M., Hawk, L. W., and Epstein, J. N. (2012). Reaction time variability in ADHD: a review. *Neurotherapeutics* 9, 500–508. doi: 10.1007/s13311-012-0138-5

Truccolo, W. A., Ding, M., Knuth, K. H., Nakamura, R., and Bressler, S. L. (2002). Trial-to-trial variability of cortical evoked responses: implications for the analysis of functional connectivity. *Clin. Neurophysiol.* 113, 206–226. doi: 10.1016/s1388-2457(01)00739-8

Uusberg, A., Uibo, H., Kreegipuu, K., and Allik, J. (2013). EEG alpha and cortical inhibition in affective attention. *Int. J. Psychophysiol.* 89, 26–36. doi: 10.1016/j.ijpsycho.2013.04.020

van Driel, J., Ridderinkhof, K. R., and Cohen, M. X. (2012). Not all errors are alike: theta and alpha EEG dynamics relate to differences in error-processing dynamics. *J. Neurosci.* 32, 16795–16806. doi: 10.1523/jneurosci.0802-12.2012

VanRullen, R., Busch, N., Drewes, J., and Dubois, J. (2011). Ongoing EEG phase as a trial-by-trial predictor of perceptual and attentional variability. *Front. Psychol.* 2:60. doi: 10.3389/fpsyg.2011.00060

Vaurio, R. G., Simmonds, D. J., and Mostofsky, S. H. (2009). Increased intra-individual reaction time variability in attention-deficit/hyperactivity disorder across response inhibition tasks with different cognitive demands. *Neuropsychologia* 47, 2389–2396. doi: 10.1016/j.neuropsychologia.2009.01.022

Wascher, E., Rasch, B., Sänger, J., Hoffmann, S., Schneider, D., Rinkenauer, G., et al. (2014). Frontal theta activity reflects distinct aspects of mental fatigue. *Biol. Psychol.* 96, 57–65. doi: 10.1016/j.biopsycho.2013.11.010

Zang, Y. F., He, Y., Zhu, C. Z., Cao, Q. J., Sui, M. Q., Liang, M., et al. (2007). Altered baseline brain activity in children with ADHD revealed by resting-state functional MRI. *Brain Dev.* 29, 83–91. doi: 10.1016/j.braindev.2006.07.002

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.

Copyright © 2020 Wang, Sun, Liu, Pan, Zheng, Wang, Zang and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

frontiers in Computational Neuroscience

REVIEW
published: 15 September 2020
doi: 10.3389/fncom.2020.00082

Oscillatory Bursting as a Mechanism for Temporal Coupling and Information Coding

Idan Tal[1,2]*, *Samuel Neymotin*[2], *Stephan Bickel*[2,3,4], *Peter Lakatos*[2,5] and *Charles E. Schroeder*[1,2]*

[1] Department of Psychiatry, Columbia University Medical Center, New York, NY, United States, [2] Translational Neuroscience Division, Center for Biomedical Imaging and Neuromodulation, Nathan Kline Institute for Psychiatric Research, New York, NY, United States, [3] Feinstein Institutes for Medical Research, Northwell Health, New York, NY, United States, [4] Departments of Neurosurgery and Neurology, Northwell Health, New York, NY, United States, [5] Department of Psychiatry, New York University School of Medicine, New York, NY, United States

Even the simplest cognitive processes involve interactions between cortical regions. To study these processes, we usually rely on averaging across several repetitions of a task or across long segments of data to reach a statistically valid conclusion. Neuronal oscillations reflect synchronized excitability fluctuations in ensembles of neurons and can be observed in electrophysiological recordings in the presence or absence of an external stimulus. Oscillatory brain activity has been viewed as sustained increase in power at specific frequency bands. However, this perspective has been challenged in recent years by the notion that oscillations may occur as transient burst-like events that occur in individual trials and may only appear as sustained activity when multiple trials are averaged together. In this review, we examine the idea that oscillatory activity can manifest as a transient burst as well as a sustained increase in power. We discuss the technical challenges involved in the detection and characterization of transient events at the single trial level, the mechanisms that might generate them and the features that can be extracted from these events to study single-trial dynamics of neuronal ensemble activity.

Keywords: oscillations, transients, bursts, timing, single trial, methods

OPEN ACCESS

Edited by:
Joaquín J. Torres, University of Granada, Spain

Reviewed by:
Norman Forschack, Leipzig University, Germany
Carmen Castro Canavier, Louisiana State University, United States

***Correspondence:**
Idan Tal
idan.tal@nki.rfmh.org
Charles E. Schroeder
cs2388@cumc.columbia.edu

Received: *03 June 2020*
Accepted: *31 July 2020*
Published: *15 September 2020*

Citation:
Tal I, Neymotin S, Bickel S, Lakatos P and Schroeder CE (2020) Oscillatory Bursting as a Mechanism for Temporal Coupling and Information Coding. Front. Comput. Neurosci. 14:82. doi: 10.3389/fncom.2020.00082

INTRODUCTION

At a most basic level, neuronal oscillations reflect synchronous and rhythmic shifting of neuronal ensembles between high and low excitability states (Buzsaki, 2006; Schroeder and Lakatos, 2009). An obvious consequence is that most neurons in an ensemble are more likely to fire action potentials at a particular (high excitability) oscillatory phase. Neuronal oscillations have been proposed to underlie many critical brain operations including attentional selection of sensory input (Schroeder and Lakatos, 2009), parsing/chunking of complex input streams (Poeppel et al., 2008; Ding and Simon, 2014), generation of motor output (Baker et al., 1999; Parkkonen et al., 2015), memory encoding and retrieval (Jensen et al., 2007), ordering of information carried by spike trains through spike-phase coding (Kayser et al., 2009) and temporal coupling of distant ensembles to enhance information transfer (Varela et al., 2001; Fries, 2015; Singer, 2018). Key to their mechanistic role in these operations is the idea that neuronal oscillations in a particular frequency synchronize dynamically to couple a group of neurons into a cell assembly for a specific

brain task (Buzsáki, 2010), and then just as dynamically desynchronize so that neurons can regroup for the next brain task.

While oscillatory activity is often viewed as sustained, this perspective has been challenged in recent years by the notion that oscillations may occur as transient bursts of activity that may only appear as sustained activity when averaging across multiple trials (Lakatos et al., 2004; Jones, 2016; Sherman et al., 2016), in contrast to the idea that averaging in the time domain can diminish or obliterate oscillations when they are not phase-aligned across trials. Since some brain tasks last a few milliseconds (e.g., reacting to an alerting stimulus), while others require many seconds or even longer (solving a mathematical equation), both scenarios are likely to co-exist.

In this review, we will first contrast the ideas of oscillations as transient bursts vs. sustained events, and outline the technical challenges involved in the detection and characterization of transient events at the single trial level. After that we will discuss the circumstances and mechanisms that likely determine whether an oscillation will emerge as a transient or a more sustained brain event.

OSCILLATORY EVENTS

Methods to Identify Transient Oscillatory Bursts

One of the main challenges in single trial analysis is dealing with low signal-to-noise ratio (SNR). The idea behind averaging across trials is that the signals related to some event are enhanced compared to neural activity unrelated to the event, and other non-neural sources of noise, thus providing a representative signal for a "clean" neural response to the stimulus. Obviously, however, the brain has to operate on a single trial basis when performing a cognitive task. If indeed transient oscillatory bursts are involved in information processing, the first step would be to reliably detect such transients at the single trial level. However, as evident from the methods described below, reliable detection of power increases at the single trial level is not trivial. Due to the typically low SNR of single trial responses, frequency decomposition can yield a "bursty" time-frequency profile even in the simple case of constant-amplitude sustained oscillations (see **Figure 1**). This effect is particularly likely to occur when there is cross-frequency phase amplitude coupling (Lakatos et al., 2005; Schroeder and Lakatos, 2009) and an oscillation is obscured by noise in the non-ideal phase of the lower-frequency oscillation.

Due to the technical challenge presented by the low single-trial SNR, several methods for transient detection have been proposed. Most detection algorithms rely on filtering the data into frequency bands (with a wavelet convolution or a Hilbert transform) and detecting whether a power fluctuation exceeds an amplitude threshold on a trial-by-trial basis. This is known as the p-episode method and may also be combined with a duration threshold (Caplan et al., 2001). Variations of this method have been used to study single trial oscillations in recent years. Sherman et al. (2016) detected beta transients by finding the maxima in the single trial wavelet transformed data. The authors chose the highest beta event in each trial, sorted the events from low to high power and analyzed the top 50 highest power events. This procedure revealed a stereotypical time domain waveform that spans <150 ms of a beta burst (roughly three cycles). Lundqvist et al. (2016) used a similar approach to detect an increase in gamma power of two standard deviations above the mean spectral power in that band, but also added a duration constraint of an increase lasting at least three cycles. Neymotin et al. (2020a) demonstrates that below three cycles, any "length measure" becomes unreliable, in that it overestimates the number of cycles. Hughes et al. (2012) used an oscillation detection method to extract both sustained and transient rhythms from rat hippocampal recordings termed Better Oscillation Detection Method [BOSC; first described by Caplan et al. (2001)]. The BOSC method is applied to continuous signals to detect the incidence of oscillatory components that exceed amplitude and duration thresholds while ignoring the transient voltage fluctuations that may accompany artifacts or evoked potentials. The power threshold is set as the 95th percentile of the theoretical χ^2 distribution of wavelet power values and the duration threshold was set to three cycles (again, similar to previous methods). An extension to the BOSC method was recently suggested (eBOSC; Kosciessa et al., 2019) in which rhythmic and non-rhythmic episodes are automatically separated. An additional measure of "rhythmicity" termed lagged-coherence uses the present phase of a signal to predict future phases (Fransen et al., 2015). They show rhythmicity peaks detected in ongoing sensorimotor signals that are not visible using conventional power analysis, suggesting that rhythmicity measures are more suitable for identifying neuronal oscillations. Another approach to the detection and characterization of neuronal rhythms uses Hidden Markov Models (HMMs) to overcome some of the limitations of the amplitude-threshold approaches by avoiding a direct amplitude envelope threshold (Quinn et al., 2019). The HMM represents the signals as a system that moves through a set of discrete states, with each state having a probability of being "on" at each time point. Thus, the thresholding procedure is applied to the probabilities rather than the signals themselves. In addition, using temporal regularization, HMM can avoid state transitions due to small changes in the envelope close to the threshold [see Figure 2 in Quinn et al. (2019)]. One of the downsides of this method is that a fixed number of states must be defined in advance. In cases where the distribution of power values (or probabilities) is bi-modal, it is easy to define two states, but in many cases it is harder to define and interpret several states, specifically when studying wide-band phenomenon.

Single trial analysis and power-change detection can be computationally costly and might not be feasible for real-time closed-loop brain stimulation experiments or brain-machine interfaces. Karvat et al. (2019) suggested a method for the detection of transient oscillatory activity specifically designed for real-time data analysis and demonstrated its usefulness for analyzing volitional increase of beta-band burst-rate in the motor cortex of rats. The authors suggest defining a burst as a power peak in time and frequency, exceeding a threshold defined as a

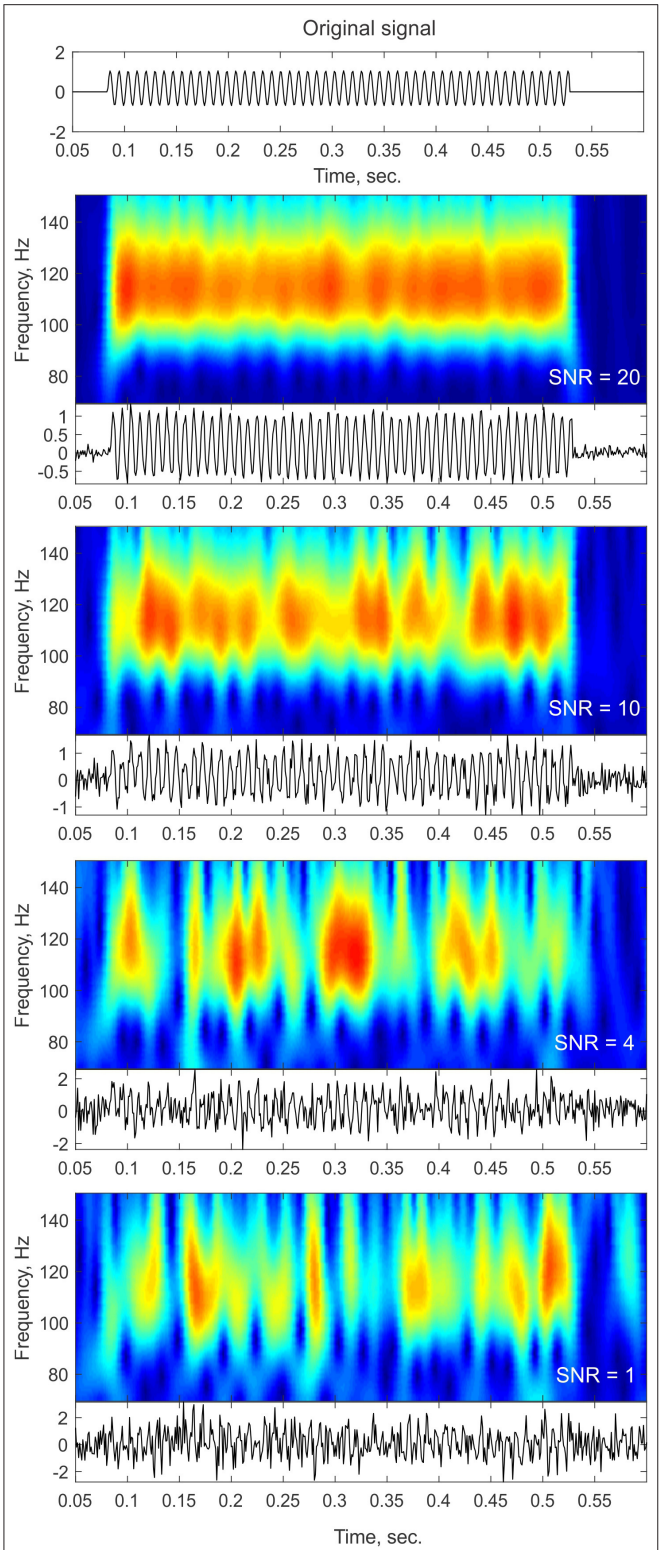

FIGURE 1 | Noise influence on time-frequency profiles. The top panel shows a trace of a pure sinusoid at 110 Hz with a constant amplitude lasting for a duration of ∼400 ms. The bottom panels show the time-frequency representation of the signal (top insets) and the time-domain trace (bottom insets) after adding different levels of white Gaussian noise. As the noise level

(Continued)

FIGURE 1 | increases (lower SNR), the time-frequency profile contains more gaps and becomes more "bursty" in appearance. In cases of low SNR, such as the ones that might be observed in single trial or ongoing data, sustained oscillations might be broken down into isolated peaks and thus might be considered as transient bursts that might even slightly vary in frequency as a result of the noise structure.

percentile to assure a statistically sound significance definition under non-normal distributions. The method is based on 32 real-time narrow bandpass FIR filters followed by peak and trough detection in the filtered signals that exceed the threshold set as the 98th percentile of power.

In addition to the spectrotemporal properties of oscillations, waveform shapes appear to matter as well. Robust differences in the waveform shapes of the oscillations mentioned above can be assumed to represent differences in the properties of their underlying generators (for review see Cole and Voytek, 2017). Due to the rich and possibly variant waveform across different cortical locations and cognitive tasks, detection techniques should combine power threshold, duration threshold and waveform specificity for each frequency band and recording location. An example of such an approach is to detect increases in the single trial time-domain or time-frequency data, then calculate the principal components of the time-domain waveforms and use the first principal component as a template for the detection of additional events in the time domain using a template matching scheme (see Abeles, 2014; Tal and Abeles, 2016, 2018). In brief, a segment of data is projected onto the template. The length of the projection is treated as the signal and the residual is treated as the noise. The threshold is then based on the signal-to-noise ratio. Such methods may lead to the creation of a "dictionary" of waveforms (similar to EEG atlases; Stern, 2005) that exhibit different oscillatory properties. Generating such a dictionary of oscillatory signatures may allow us to further test interactions between different neuronal populations under the assumption that a specific signature is generated by a specific cell population or a specific process (Siegel et al., 2012; Womelsdorf et al., 2014).

Some of the methods mentioned above might reduce concerns regarding low SNR. Occasionally, we may observe oscillatory bursts of sufficient amplitude that the SNR is less of a concern (see e.g., below). There is no trivial way to define a duration value that can serve as a boundary between transient and sustained oscillations. Thus, we can only say that with most commonly used analysis approaches, particularly due to the practice of averaging multiple trials prior to quantification, oscillatory activity might appear longer than they actually are. Clearly, both the nature of the task performed by the subject and the recording technique (e.g., invasiveness, electrode location) would influence the amplitude, duration and frequency of the recorded oscillations and thus also the SNR. Thus, there might be substantial variability in the characteristics of these oscillations across different studies, when they are studied at the single-trial level. Typically, invasive recording techniques provide higher SNR than non-invasive techniques.

Three independent examples of such invasive recordings in humans (sEEG) and non-human primates (laminar probe) are shown in **Figure 2** where oscillations are visible by eye in ongoing recordings. Some of these oscillations tend to be more sustained, specifically at lower frequency (e.g., 8–12 Hz; **Figure 2A**, middle panel) while higher frequencies reveal a "bursty" profile in this example (13–30 Hz; **Figure 2A**, bottom panel). We estimated the duration and number of cycles at a descriptive level for a few minutes of recording from these example (**Figure 2B**; see figure caption for more details). Even though these are merely selected examples and should not be considered as evidence for the specific durations of oscillatory activity, such recordings give us confidence that the basic phenomenon of a bursty oscillation exists. Recently, Neymotin et al. (2020a) quantified oscillation event features in resting-state invasive recordings from auditory cortex of humans and non-human primates. To our knowledge, this was the first attempt to characterize ongoing (single trial) oscillatory activity across different frequencies and species. They found that oscillations at all commonly studied frequency bands (i.e., delta—high gamma) exhibit multiple cycles (average of four cycles across all frequency bands; range: 1–44 cycles) with fluctuating frequency and amplitude. They also found that ∼90% of the time, oscillation events of at least one frequency band are occurring, suggesting that multicycle neuronal oscillations across a wide range of frequencies dominate auditory cortex dynamics. Interestingly, temporal predictability across bursts differed significantly from Poisson distribution assumption which indicates inter-burst quasi-rhythmicity.

However, none of the methods described above (including the ones that do not require a pre-defined threshold) can negate the possibility that a transient time-frequency profile results simply from low SNR (see **Figure 1**). Thus, the most convincing evidence suggesting that transient oscillatory activity exists and is meaningful come from studies relating features of single-trial oscillatory activity with behavioral or perceptual phenomenon and show that such signals add information on top of the traditional (e.g., averaged; sustained) view of neural oscillations.

Oscillatory Activity as Transient Bursts

Classic evidence for sustained oscillatory activity were based on averaging neural signals across many trials to form a representation of the neural activity with higher SNR. However, moment to moment perceptual representations in the real-world do not operate in such a way. To understand the neural basis of perception and action, and the involvement of oscillatory activity in these processes, one must explore trial-by-trial changes in oscillatory dynamics. In the following section, we will review evidence for transient or "bursty" oscillatory activity in both resting state (ongoing) and task-related activity, as well as their relationship with behavior.

Transient EEG events were first described by Berger (1930) during sleep and were later termed sleep spindles (Loomis et al., 1935). These 12–14 Hz bursts of oscillatory activity have been found in all mammals and the thalamocortical mechanisms generating them have been well-established (Dijk et al., 1993; Steriade et al., 1993). Yet, their function remains unclear (see De Gennaro and Ferrara, 2003 for a review). The first observations of oscillatory activity as transient bursts in awake subjects dates to 1966 when Jaffe and Weiss reported unilateral alpha bursts that are different from the alpha rhythm in several aspects (Jaffe and Weiss, 1966). These "alpha-range bursts" appear mainly in temporal EEG electrodes and correlate with clinical evidence of brain disease. They last for 3–4 s and increase during hyperventilation or drowsiness. The authors report that activity of this type has been occasionally observed in their lab and listed as an unusual finding of unknown significance. Alpha bursts of ∼3 s in duration were also found during periods of REM sleep (Cantero and Atienza, 2000). These bursts are different from sustained alpha in that they are not accompanied by an increase in EMG activity and thus might be indexing different functional role from REM background alpha. The authors hypothesize that such alpha bursts may work as a micro-arousal in human REM sleep to facilitate a connection between the dreaming brain and the external world. Transient oscillatory activity was also observed at lower frequency bands in humans. For example, Hebert and Lehmann (1977) found the emergence of theta bursts in healthy subjects practicing transcendental meditation. These bursts appeared every 2 min on average, had a duration of about 1.8 s and were preceded and followed by alpha rhythm. Since the subjects reported pleasant states during the theta bursts, the authors hypothesize that theta bursts may be the manifestation of a state adjustment mechanism that comes into play during prolonged low arousal states and related to relaxation. While these findings were reported for ongoing activity at different mental states, transient oscillatory activity at low frequencies was also observed in intracranial recordings in human epilepsy patients during a virtual reality environment navigation task (Bush et al., 2017). They reported that human theta oscillations appear in transient bursts that typically last several cycles around movement onset and throughout the movement, in contrast to the continuous rhythm in the rodent hippocampus (Watrous et al., 2013). Although it is not clear whether these sporadic oscillations could encode continuous self-motion information, it is possible that location estimates are updated intermittently during theta bursts, in accordance with the outcome of planned movements, rather than tracked continually throughout the movement by an ongoing theta oscillation.

In recent years, there has been a renewed interest in the bursty qualities of oscillatory activity, particularly in higher frequency bands. By analyzing the amplitude and frequency of gamma bursts above the auditory cortical regions of cats, Lakatos et al. (2004) found that while attention mostly affects amplitude, arousal affects the frequency of gamma oscillatory bursts. Sherman et al. (2016) found that spontaneous neocortical beta (15–29 Hz) from somatosensory and frontal cortex emerged as non-continuous beta events typically lasting <150 ms with a stereotypical waveform. These "beta events" occur with varying levels of alpha activity (that seemed more sustained) and their waveform seemed to be consistent across species (mice, monkeys, and humans). The authors determined that beta events do not necessarily depend on rhythmic inputs but on the relative timing and strength of synchronous proximal (i.e., proximal to the soma and basal dendrites) and distal (i.e., to apical dendrites in L2/3) drives. Beta bursts were also observed in local field potential

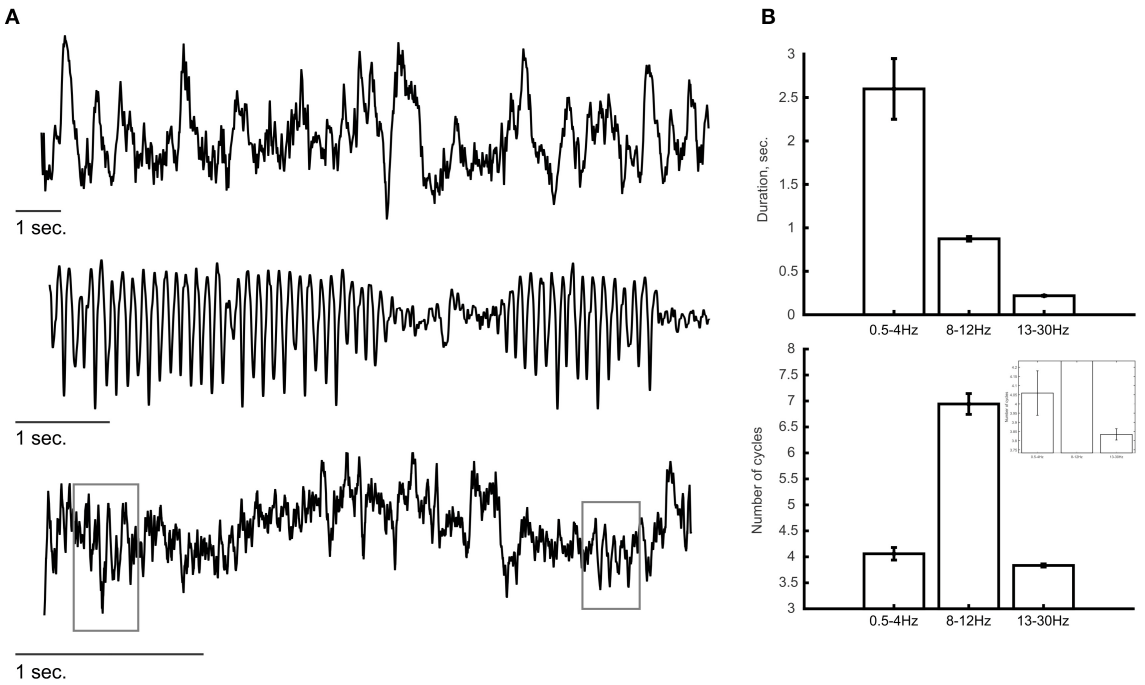

FIGURE 2 | Examples of ongoing oscillatory activity. **(A)** Top and middle panels show a trace of invasive recordings from sEEG electrodes showing low-frequency oscillations (0.5–4 Hz, top; 8–12 Hz, middle). (Bottom) Laminar probe recording from a non-human primate showing bursts of oscillatory activity at a frequency of 13–30 Hz. Gray rectangles indicate the detection of transient oscillatory bursts. **(B)** Descriptive analysis of the duration of oscillatory activity at each frequency band. The duration of oscillations (top) was estimated as the period of time in which the power at that frequency band exceeded the 95th percentile of the theoretical χ^2 distribution of wavelet power values. The number of cycles (bottom) were calculated by multiplying the duration (in seconds) by the peak frequency of the oscillation (in Hz). Lower frequency oscillations (i.e., 0.5–4 and 8–12 Hz) tend to show longer durations compared with the higher frequency oscillations (13–30 Hz) in which most oscillatory bursts consisted of ∼3.5 cycles. The inset shows a zoomed version to visualize the differences between 0.5–4 and 13–30 Hz. Error bars indicate standard error of the mean.

(LFP) signals recorded from the striatum and motor–premotor cortex of macaque monkeys performing a reaching task (Feingold et al., 2015). Using single trial analysis, they showed that beta bursts typically lasted 90–115 ms, and that extended periods of beta band synchronization reflected a modulation in the density of these short bursts. Burst probabilities were region and task-time specific such that in motor cortex they peak following the movement, while in striatum they peaked after reward and continued through the post-performance period. Lundqvist et al. (2016) used a trial-by-trial analysis and found that brief bursts of gamma-band activity (45–100 Hz) accompanied encoding and re-activation of sensory information in recording sites associated with spiking that reflected "to be remembered" items. Neuronal activity reflecting encoding or decoding correlated with changes in gamma burst rate. Additionally, they showed that gamma—alpha (8–10 Hz) coupling was not related to the periodicity of the gamma bursts but rather to the consistency in the duration of the gamma bursts, indicating that lower frequencies might modulate gamma-burst duration. Beta band oscillations (20–35 Hz) also appeared as transients in the Lundqvist et al. findings and were interpreted as reflecting a default (holding) state because it was interrupted by encoding and decoding. The authors concluded that working memory is not associated with sustained activity but rather discrete oscillatory dynamics and spiking. Beta range oscillations were also suggested to serve to clear memory states by resonantly driving transient bouts of spike synchrony which destabilize the network activity (Schmidt et al., 2018). Interestingly, the most effective oscillatory activity for allowing flexible switching between network states was burst-like with a sharp onset rather than a pure sinusoid. In addition, the authors demonstrate that such oscillatory bursts arise spontaneously in networks of excitatory and inhibitory neurons.

Transient oscillatory events may also provide an additional coding space for neuronal processes by utilizing the rate or timing of the transient events with regard to the stimulus or even with regard to other transient events. Shin et al. (2017) showed that differences in the rate of beta events predicted detection of stimuli at perceptual threshold and that non-detectable trials were more likely to have a beta event within ∼200 ms prior to the stimulus. Using MEG, Little et al. (2018) found that motor cortical beta in individual trials appears as high amplitude, transient infrequent bursts. Beta burst timing was a stronger predictor of single trial behavior than beta burst rate or single trial beta amplitude, with later bursts corresponding to delayed response times. The relative timing of transient events was studied in the context of sensorimotor synchronization using MEG to show that decoding of behavioral conditions using time-differences between transient events across brain regions is

significantly more accurate than other characteristics of the signal (Tal and Abeles, 2016, 2018; Felsenstein et al., 2019). Moreover, such transient events (treated as a point-process) might form more complex repeating sequences of activation with millisecond precision (Tal and Abeles, 2016, 2018; Felsenstein et al., 2019). These results suggest that relevant information might be encoded by subtle time differences or cascades of transient events across the brain. Although the study of oscillatory activity as transient events is still at an early stage, several features of these bursts (e.g., rate, duration, timing) have been linked back to behavior suggesting that the dynamics of such transient activity may be involved in cognitive processes.

The current review focuses on oscillatory events, however, there are also transient events that are not oscillatory in nature. A clear example of such event is an evoked potential generated in response to a stimulus. The nature of harmonic analyses techniques, such as wavelets and Fourier transform identifies these signals as oscillations. The BOSC method (Caplan et al., 2001; Hughes et al., 2012) attempts to avoid the detection of transient, non-oscillatory events by identifying oscillatory episodes at each frequency using both power and duration thresholds. Another approach to avoid bias due to externally-driven events removed the average evoked-responses waveforms from each cortical layer (Neymotin et al., 2020a), though if this removal uses simple subtraction of the trial-averaged response from each single trial, it runs the risk of creating artifacts (Knuth et al., 2006). Tal and Abeles (2016) used an algorithm that may detect both oscillatory and non-oscillatory transient events using a template matching scheme. They show that most of their detected events were not associated with clear periodic oscillations. They demonstrate that both oscillatory and non-oscillatory events showed similar increases in population activity around the times of these events (Tal and Abeles, 2018) and suggested them as markers of sudden increase in population activity that might indicate the recruitment of a new cell assembly within the cortical patch. It is not yet clear whether event-related-potentials trigger the same mechanisms in terms of the canonical circuits activated by internally generated oscillations. Sherman et al. (2016) argued that such brief sharp events are due to brief, strong excitation of the superficial cortical layers riding on the broader but weaker excitation of deep cortical layers. Laminar biophysical models of the thalamocortical system that accurately simulate recorded signals, such as local field potentials, can be used to predict the types of waveforms that are recorded *in vivo* after sensory stimulation, and offer mechanistic explanations for their features. For example, providing brief, strong thalamocortical activation to a hypothetical neocortical model would trigger production of a transient ERP-like event in the circuit, with a characteristic waveform (Neymotin et al., 2020b). Although running a wavelet filter on such a waveform will produce high power at a frequency inverse to the duration of the ERP, since it was produced by a punctate event, this type of waveform should not be considered an oscillation (Neymotin et al., 2020a). In contrast, specific synaptic connectivity and input patterns provided to a circuit model lead to production of sustained multi-cycle oscillations. Some of these circuit mechanisms and their implications in detecting oscillatory bursts from electrophysiology data *in vivo* are described in the next section.

Mechanisms of Transient Oscillatory Activity

The studies discussed above suggest that neocortical oscillations tend to be short-lived and bursty, however, some neurological disorders, such as Parkinson's disease, are clearly associated with prolonged rhythms (Tinkhauser et al., 2017a,b). In general, stronger activation of a particular circuit component that generates a specific oscillation (such as gamma), would produce a more sustained form of that oscillation. Weaker activation, either through reduction of the frequency and strength of AMPA/NMDA synaptic inputs to that component or from stronger suppressive inhibition, can result in gaps between oscillatory bursts (Lee and Jones, 2013; Neymotin et al., 2020b). **Figure 3** demonstrates the results of a neocortical column simulation of bursty Pyramidal-interneuron network gamma (PING) oscillations using the Human Neocortical Neurosolver (HNN) software (https://hnn.brown.edu). As shown in **Figure 3**, in PING, gamma is generated through a sequence of activations: (1) stochastic synaptic inputs drive spiking of pyramidal neurons, causing collateral activation of fast spiking (basket type) interneurons, (2) activation of the fast-spiking interneurons then causes feedback inhibition lasting a gamma cycle (∼20 ms for 50 Hz gamma), determined by the duration of the rise and fall of the $GABA_a$ synaptic conductance, and (3) after $GABA_a$ inhibition runs its course, pyramidal neurons are again able to spike and the cycle repeats. Synchronization of inhibitory interneurons, which is responsible for generation of gamma rhythms, is seen in the raster plot of **Figure 3A**, with the nearly vertical lines that recur at a gamma period (white and blue). In general, sustained gamma is produced when there is continued strong activation of pyramidal neurons, which causes continuing periodic activation of the interneurons, resulting in large amplitude/persistent gamma. Lowering the frequency or strength of excitatory synaptic inputs driving the pyramidal neurons, causes weaker, intermittent activation of the interneurons, and temporal gaps between interneuron-generated gamma bursts, as shown in **Figure 3A**. Note that in the raster plot, not all interneurons are activated at each gamma bout. Additionally, different subsets of the interneurons are activated. This firing pattern is a hallmark of weak PING, considered weak because the gamma amplitude occasionally waxes and wanes depending on the level of interneuron activation. In this example, pyramidal neurons (green, red dots) fire even less frequently, but are still synchronized by the interneurons (note the periodic gaps between the sparse pyramidal neuron firing). **Figure 3B** shows a single trial of the current dipole signal generated by HNN's biophysical cortical circuit model (top) and its associated time-frequency representation, using the Morlet wavelet spectrogram (bottom). As shown, the current dipole signal's gamma oscillation has a peak between 40 and 60 Hz, and has power waxing and waning. A close look at the spectrogram reveals that the gamma oscillation is only present at discrete times. However, when taking the average wavelet spectrogram from multiple trials of

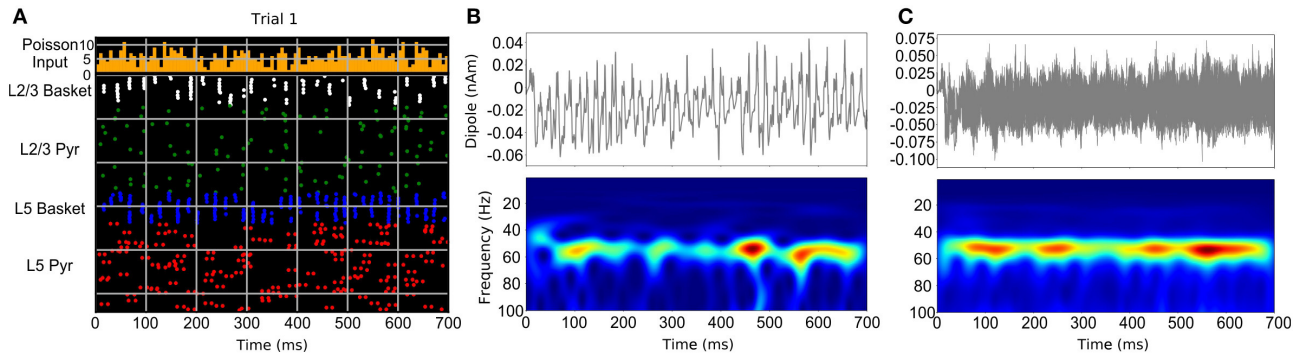

FIGURE 3 | Neocortical circuit model used to simulate bursty gamma oscillation events through a weak PING mechanism. **(A)** Raster spiking plot of neuronal firing times from single trial of weak PING simulation. Top panel shows histogram of low-frequency, noisy Poisson inputs used to drive pyramidal neurons and interneurons in the model. Bottom panel shows population color-coded firing times of individual neurons. **(B)** Single trial current dipole signal (top) and Morlet wavelet spectrogram from dipole signal (bottom) from weak PING model. **(C)** Simulation of one hundred trials of weak PING model produces one hundred current dipole signals (top). Although gamma oscillation events occur at different times in each trial, averaging the wavelet spectrogram across trials (bottom) produces an appearance of a sustained gamma oscillation [adapted with permission from Figure 10 of Neymotin et al. (2020b) under the license: https://creativecommons.org/licenses/by/4.0/].

this simulation (**Figure 3C**), the gamma oscillation appears to be more continuous. This is because the wavelet spectrogram always has positive values, and adding the gamma events which occur at different times across trials, produces the appearance of continuity. This is one mechanism for the bursty gamma observed in experimental data and highlights the importance of careful analysis of single trial data.

A related theme in cortical dynamics is the presence of multiple interacting oscillations caused by different time scales of inhibition provided by different classes of interneurons (for reviews see Whittington et al., 2000; Skinner, 2012; Kopell et al., 2014). For example, Neymotin et al. (2011) used a model with intermediate complexity to replicate normal and pathological hippocampal dynamics. In their model, oriens-lacunosum moleculare (OLM) interneurons produce theta through relatively long-lasting inhibition. OLM inhibition of fast-spiking basket and pyramidal neurons then modulated the faster gamma rhythm, which was produced through the standard PING mechanism. This interaction between OLM and basket interneurons caused gamma rhythms to increase and decrease based on the phase of the slower theta rhythm. This cross-frequency coupling mechanism could be used to model gamma bursts too, since a few strong cycles of gamma appear in between strong periods of OLM inhibition at the theta rhythm. As evidenced by a multitude of non-human (Lakatos et al., 2005; Buzsaki, 2006; Schroeder and Lakatos, 2009) and human (Canolty et al., 2006; Canolty and Knight, 2010) studies demonstrating phase-amplitude coupling, similar mechanisms should operate in the neocortex *in vivo*, which has an intricate circuitry with a multitude of interneuron types (Dienel and Lewis, 2019).

Another circuit model of neocortex aimed at determining the origin of beta oscillation events (Sherman et al., 2016). This biophysical model simulated current dipole signals produced by the circuit, allowing explicit comparison to source-localized current dipole signals from MEG/EEG studies. The neocortical model consisted of simplified models of pyramidal and inhibitory interneurons arranged in superficial and deep cortical layers and interconnected using AMPA and GABA synapses. Pyramidal neuron dendrites spanned the cortical layers. Synaptic inputs were provided to the pyramidal neurons to initiate network activity. These synaptic inputs were applied at proximal and distal locations on pyramidal neuron dendrites to model inputs from thalamic core (proximal) and thalamic matrix and corticocortical feedback (distal). Each of these types of synaptic input pushed current flow within the pyramidal neuron apical dendrites in opposite directions. The model was able to produce beta oscillation events through a series of ∼10 Hz stochastic synaptic inputs provided to proximal and distal pyramidal neuron dendrites. Beta oscillation events were produced when ∼100 ms duration proximal synaptic inputs (pushed current flow toward superficial layers) were truncated by a more synchronous 50 ms distal inputs (sharply pushed current flow toward deep layers), which produced a 50 ms current dipole waveform, matching the waveforms seen from source-localized human MEG experiments. Since the synaptic inputs were stochastic, the production of beta events was also stochastic, producing bursty oscillation events. Additionally, when the proximal and distal synaptic inputs arrived out of phase, instead of beta, alpha events were produced. This occurred because each set of synaptic inputs was provided to the model at the alpha period (100 ms interval). Invasive laminar electrophysiology recordings from non-human primate somatosensory cortex were used to confirm the model's accuracy.

CONCLUDING REMARKS

We reviewed several studies suggesting that brain rhythms tend to appear as short-lived bursts of oscillatory activity. The importance of these observations lies in our interpretation of the functional role of neural oscillations, the mechanisms generating them, the potential information they may carry, and the way we must analyze them. One of the major points raised here is that sustained oscillations can appear "bursty" in the

presence of noise and conversely, "bursty" oscillations can appear sustained when averaged across trials. Many cognitive studies report an increase in oscillatory activity that only lasts for a brief duration (e.g., a few tens of milliseconds). This might represent the main response to the stimulus but at the same time, might fail to explain the trial by trial variance in burst timing, frequency, and amplitude. In addition, the averaging approach might also miss other important responses that are either "smeared" in the averaged response or appear at different times across trials and are obscured by averaging. While the averaging approach has its advantages, specifically in increasing the SNR of the phase-locked response to a stimulus, single-trial analysis of oscillatory events should be performed to better understand the mechanisms of oscillatory activity by exploring the variability in different features of oscillatory activity across trials and, whenever possible, their relationship with behavioral and perceptual phenomenon. Estimating the characteristics of single-trial (or ongoing) oscillations is not trivial, and thus several methods for detecting transient events were suggested. Most of these methods depend on amplitude and duration thresholds or a probabilistic threshold. Due to the differences in the goals of each study, the design of the experiments, the recording techniques, and the frequency and time-windows studied, variability in the estimation of the characteristics of single-trial oscillations is to be expected and it is difficult to provide a single recommendation on the algorithm that should be used to study neuronal rhythms at the single trial. The simplest approaches (such as p-episodes) can be useful in cases where lower computational time is essential (such as in closed-loop experiments), while more computationally demanding approaches can achieve more fine-tuned results offline. When possible, we recommend applying more sophisticated algorithms, such as HMM or amplitude and duration thresholds combined with template-matching that carry less risk of false detection due to artifacts or noise. We identify five features of short-lived oscillations that may provide information-coding space for the brain: (1) Amplitude—indicates the size and synchronicity of the underlying neuronal population. (2) Temporal span (duration)—how long the synchrony within a population is maintained. (3) Frequency span—might index the participant neuron circuits and critically, the inherent conductances of their specific neuronal constituents. When studied at the single-trial level, these features may explain variability in behavioral performance across trials that cannot be observed in the averaged waveforms. (4) Inter-burst and stimulus-burst interaction–measures, such as the burst-rate, burst-timing, inter-burst interval, coefficient of variation, fano-factor and more complex spatio-temporal sequences comprised by transient bursts might be used to explore the single trial dynamics of oscillatory activity and non-oscillatory transient events. The general idea is to treat the timing of these transient events as a parallel point-process to study their temporal relationship with other events and with the stimulus. For example, determine the rhythmicity across events from a given oscillation frequency band (e.g., whether oscillatory events are rhythmic and predictable, or Poisson distributed). (5) Time-domain waveforms—might index different biophysical generators. We note that this feature is more abstract and challenging to measure but should be further studied to extract the meaningful features within the waveforms and test for repeating waveforms in the data. Time-domain waveforms may also reveal differences between oscillatory and non-oscillatory events that might differ in both their mechanisms and their role in information processing. Such features are necessary to study brain rhythms at the single trial level and take advantage of the temporal dynamics of neural oscillations to better understand their role in information transmission, processing, and coding.

AUTHOR CONTRIBUTIONS

IT and SN performed the simulations. SN designed the model shown in **Figure 3**. SB and CS provided the data for the examples shown in **Figure 2**. IT and CS wrote the manuscript with support from SN, SB, and PL. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by NIH grants MH109429, MH111439, and DC012947, by a grant from The James S. McDonnell Foundation, and by ARO W911NF-19-1-0402.

REFERENCES

Abeles, M. (2014). Revealing instances of coordination among multiple cortical areas. *Biol. Cybern.* 108, 665–675. doi: 10.1007/s00422-013-0574-2

Baker, S. N., Kilner, J. M., Pinches, E. M., and Lemon, R. N. (1999). The role of synchrony and oscillations in the motor output. *Exp. Brain Res.* 128, 109–117. doi: 10.1007/s002210050825

Berger, H. (1930). Ueber das Elektrenkephalogramm des Menschen. *J. Psychol. Neurol.* 40, 160–179.

Bush, D., Bisby, J. A., Bird, C. M., Gollwitzer, S., Rodionov, R., Diehl, B., et al. (2017). Human hippocampal theta power indicates movement onset and distance travelled. *Proc. Natl. Acad. Sci. U.S.A.* 114, 12297–12302. doi: 10.1073/pnas.17087 16114

Buzsaki, G. (2006). *Rhythms of the Brain*. Oxford: Oxford University Press.

Buzsáki, G. (2010). Neural syntax: cell assemblies, synapsembles, and readers. *Neuron* 68, 362–385. doi: 10.1016/j.neuron.2010.09.023

Canolty, R. T., Edwards, E., Dalal, S. S., Soltani, M., Nagarajan, S. S., Kirsch, H. E., et al. (2006). High gamma power is phase-locked to theta oscillations in human neocortex. *Science* 313, 1626–1628. doi: 10.1126/science.1128115

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

Cantero, J. L., and Atienza, M. (2000). Alpha burst activity during human REM sleep: descriptive study and functional hypotheses. *Clin. Neurophysiol.* 111, 909–915. doi: 10.1016/S1388-2457(99)00318-1

Caplan, J. B., Madsen, J. R., Raghavachari, S., and Kahana, M. J. (2001). Distinct patterns of brain oscillations underlie two basic parameters of human maze learning. *J. Neurophysiol.* 86, 368–380. doi: 10.1152/jn.2001.86.1.368

Cole, S. R., and Voytek, B. (2017). Brain oscillations and the importance of waveform shape. *Trends Cogn. Sci.* 21, 137–149. doi: 10.1016/j.tics.2016.12.008

De Gennaro, L., and Ferrara, M. (2003). Sleep spindles: an overview. *Sleep Med. Rev.* 7, 423–440. doi: 10.1053/smrv.2002.0252

Dienel, S. J., and Lewis, D. A. (2019). Alterations in cortical interneurons and cognitive function in schizophrenia. *Neurobiol. Dis.* 131:104208. doi: 10.1016/j.nbd.2018.06.020

Dijk, D.-J., Hayes, B., and Czeisler, C. A. (1993). Dynamics of electroencephalographic sleep spindles and slow wave activity in men: effect of sleep deprivation. *Brain Res.* 626, 190–199. doi: 10.1016/0006-8993(93)90579-C

Ding, N., and Simon, J. Z. (2014). Cortical entrainment to continuous speech: functional roles and interpretations. *Front. Hum. Neurosci.* 8:311. doi: 10.3389/fnhum.2014.00311

Feingold, J., Gibson, D. J., DePasquale, B., and Graybiel, A. M. (2015). Bursts of beta oscillation differentiate postperformance activity in the striatum and motor cortex of monkeys performing movement tasks. *Proc. Natl. Acad. Sci. U.S.A.* 112, 13687–13692. doi: 10.1073/pnas.1517629112

Felsenstein, O., Tal, I., Ben-Shachar, M., Abeles, M., and Chechik, G. (2019). Decoding multimodal behavior using time differences of MEG events. *arXiv*. arXiv1901.08093.

Fransen, A. M. M., van Ede, F., and Maris, E. (2015). Identifying neuronal oscillations using rhythmicity. *Neuroimage* 118, 256–267. doi: 10.1016/j.neuroimage.2015.06.003

Fries, P. (2015). Rhythms for cognition: communication through coherence. *Neuron* 88, 220–235. doi: 10.1016/j.neuron.2015.09.034

Hebert, R., and Lehmann, D. (1977). Theta bursts: an EEG pattern in normal subjects practising the transcendental meditation technique. *Electroencephalogr. Clin. Neurophysiol.* 42, 397–405. doi: 10.1016/0013-4694(77)90176-6

Hughes, A. M., Whitten, T. A., Caplan, J. B., and Dickson, C. T. (2012). BOSC: a better oscillation detection method, extracts both sustained and transient rhythms from rat hippocampal recordings. *Hippocampus* 22, 1417–1428. doi: 10.1002/hipo.20979

Jaffe, R., and Weiss, A. H. (1966). The significance of unilateral alpha-range bursts in the EEG. *Acta Neurol. Scand.* 42, 257–267. doi: 10.1111/j.1600-0404.1966.tb01181.x

Jensen, O., Kaiser, J., and Lachaux, J.-P. (2007). Human gamma-frequency oscillations associated with attention and memory. *Trends Neurosci.* 30, 317–324. doi: 10.1016/j.tins.2007.05.001

Jones, S. R. (2016). When brain rhythms aren't 'rhythmic': implication for their mechanisms and meaning. *Curr. Opin. Neurobiol.* 40, 72–80. doi: 10.1016/j.conb.2016.06.010

Karvat, G., Schneider, A., Alyahyaey, M., Diester, I., and Optogenetics, O. (2019). Real-time detection of neural oscillation bursts allows behaviourally relevant neurofeedback. *biorxiv* 1–27. doi: 10.1101/671602

Kayser, C., Montemurro, M. A., Logothetis, N. K., and Panzeri, S. (2009). Spike-phase coding boosts and stabilizes information carried by spatial and temporal spike patterns. *Neuron* 61, 597–608. doi: 10.1016/j.neuron.2009.01.008

Knuth, K. H., Shah, A. S., Truccolo, W. A., Ding, M., Bressler, S. L., and Schroeder, C. E. (2006). Differentially variable component analysis: identifying multiple evoked components using trial-to-trial variability. *J. Neurophysiol.* 95, 3257–3276. doi: 10.1152/jn.00663.2005

Kopell, N. J., Gritton, H. J., Whittington, M. A., and Kramer, M. A. (2014). Beyond the connectome: the dynome. *Neuron* 83, 1319–1328. doi: 10.1016/j.neuron.2014.08.016

Kosciessa, J. Q., Grandy, T. H., Garrett, D. D., and Werkle-Bergner, M. (2019). Single-trial characterization of neural rhythms: potential and challenges. *bioRxiv [Preprint]*. 356089. doi: 10.1101/356089

Lakatos, P., Shah, A. S., Knuth, K. H., Ulbert, I., Karmos, G., and Schroeder, C. E. (2005). An oscillatory hierarchy controlling neuronal excitability and stimulus processing in the auditory cortex. *J. Neurophysiol.* 94, 1904–1911. doi: 10.1152/jn.00263.2005

Lakatos, P., Szilágyi, N., Pincze, Z., Rajkai, C., Ulbert, I., and Karmos, G. (2004). Attention and arousal related modulation of spontaneous gamma-activity in the auditory cortex of the cat. *Cogn. Brain Res.* 19, 1–9. doi: 10.1016/j.cogbrainres.2003.10.023

Lee, S., and Jones, S. R. (2013). Distinguishing mechanisms of gamma frequency oscillations in human current source signals using a computational model of a laminar neocortical network. *Front. Hum. Neurosci.* 7:869. doi: 10.3389/fnhum.2013.00869

Little, S., Bonaiuto, J. J., Barnes, G. R., and Bestmann, S. (2018). Motor cortical beta transients delay movement initiation and track errors. *bioRxiv [Preprint]*. 384370. doi: 10.1101/384370

Loomis, A. L., Harvey, E. N., and Hobart, G. (1935). Potential rhythms of the cerebral cortex during sleep. *Science* 81, 597–598. doi: 10.1126/science.81.2111.597

Lundqvist, M., Rose, J., Herman, P., Brincat, S. L., Buschman, T. J., and Miller, E. K. (2016). Gamma and beta bursts underlie working memory. *Neuron* 90, 152–164. doi: 10.1016/j.neuron.2016.02.028

Neymotin, S. A., Barczak, A., O'Connell, M. N., McGinnis, T., Markowitz, N., Espinal, E., et al. (2020a). Taxonomy of neural oscillation events in primate auditory cortex. *bioRxiv [Preprint]*. doi: 10.1101/2020.04.16.045021

Neymotin, S. A., Daniels, D. S., Caldwell, B., McDougal, R. A., Carnevale, N. T., Jas, M., et al. (2020b). Human Neocortical Neurosolver (HNN), a new software tool for interpreting the cellular and network origin of human MEG/EEG data. *Elife*. 9:e51214. doi: 10.7554/eLife.51214

Neymotin, S. A., Lazarewicz, M. T., Sherif, M., Contreras, D., Finkel, L. H., and Lytton, W. W. (2011). Ketamine disrupts theta modulation of gamma in a computer model of hippocampus. *J. Neurosci.* 31, 11733–11743. doi: 10.1523/JNEUROSCI.0501-11.2011

Parkkonen, E., Laaksonen, K., Piitulainen, H., Parkkonen, L., and Forss, N. (2015). Modulation of the ∼20-Hz motor-cortex rhythm to passive movement and tactile stimulation. *Brain Behav.* 5:e00328. doi: 10.1002/brb3.328

Poeppel, D., Idsardi, W. J., and Van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. B Biol. Sci.* 363, 1071–1086. doi: 10.1098/rstb.2007.2160

Quinn, A. J., van Ede, F., Brookes, M. J., Heideman, S. G., Nowak, M., Seedat, Z. A., et al. (2019). Unpacking transient event dynamics in electrophysiological power spectra. *Brain Topogr.* 32, 1020–1034. doi: 10.1007/s10548-019-00745-5

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

Schroeder, C. E., and Lakatos, P. (2009). Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci.* 32, 9–18. doi: 10.1016/j.tins.2008.09.012

Sherman, M. A., Lee, S., Law, R., Haegens, S., Thorn, C. A., Hämäläinen, M. S., et al. (2016). Neural mechanisms of transient neocortical beta rhythms: converging evidence from humans, computational modeling, monkeys, and mice. *Proc. Natl. Acad. Sci. U.S.A.* 113, E4885–E4894. doi: 10.1073/pnas.1604135113

Shin, H., Law, R., Tsutsui, S., Moore, C. I., and Jones, S. R. (2017). The rate of transient beta frequency events predicts behavior across tasks and species. *Elife* 6:e29086. doi: 10.7554/eLife.29086.025

Siegel, M., Donner, T. H., and Engel, A. K. (2012). Spectral fingerprints of large-scale neuronal interactions. *Nat. Rev. Neurosci.* 13, 121–134. doi: 10.1038/nrn3137

Singer, W. (2018). Neuronal oscillations: unavoidable and useful? *Eur. J. Neurosci.* 48, 2389–2398. doi: 10.1111/ejn.13796

Skinner, F. K. (2012). Cellular-based modeling of oscillatory dynamics in brain networks. *Curr. Opin. Neurobiol.* 22, 660–669. doi: 10.1016/j.conb.2012.02.001

Steriade, M., Nuñez, A., and Amzica, F. (1993). Intracellular analysis of relations between the slow (<1 Hz) neocortical oscillation and other sleep rhythms of the electroencephalogram. *J. Neurosci.* 13, 3266–3283. doi: 10.1523/JNEUROSCI.13-08-03266.1993

Stern, J. M. (2005). *Atlas of EEG patterns*. Lippincott Williams & Wilkins.

Tal, I., and Abeles, M. (2016). Temporal accuracy of human cortico-cortical interactions. *J. Neurophysiol.* 115, 1810–1820. doi: 10.1152/jn.00956.2015

Tal, I., and Abeles, M. (2018). Imaging the spatiotemporal dynamics of cognitive processes at high temporal resolution. *Neural Comput.* 30, 610–630. doi: 10.1162/neco_a_01054

Tinkhauser, G., Pogosyan, A., Little, S., Beudel, M., Herz, D. M., Tan, H., et al. (2017a). The modulatory effect of adaptive deep brain stimulation on beta bursts in Parkinson's disease. *Brain* 140, 1053–1067. doi: 10.1093/brain/awx010

Tinkhauser, G., Pogosyan, A., Tan, H., Herz, D. M., Kühn, A. A., and Brown, P. (2017b). Beta burst dynamics in Parkinson's disease OFF and ON dopaminergic medication. *Brain* 140, 2968–2981. doi: 10.1093/brain/awx252

Varela, F., Lachaux, J.-P., Rodriguez, E., and Martinerie, J. (2001). The brainweb: phase synchronization and large-scale integration. *Nat. Rev. Neurosci.* 2, 229–239. doi: 10.1038/35067550

Watrous, A. J., Lee, D. J., Izadi, A., Gurkoff, G. G., Shahlaie, K., and Ekstrom, A. D. (2013). A comparative study of human and rat hippocampal low-frequency oscillations during spatial navigation. *Hippocampus* 23, 656–661. doi: 10.1002/hipo.22124

Whittington, M. A., Traub, R. D., Kopell, N., Ermentrout, B., and Buhl, E. H. (2000). Inhibition-based rhythms: experimental and mathematical observations on network dynamics. *Int. J. Psychophysiol.* 38, 315–336. doi: 10.1016/S0167-8760(00)00173-2

Womelsdorf, T., Valiante, T. A., Sahin, N. T., Miller, K. J., and Tiesinga, P. (2014). Dynamic circuit motifs underlying rhythmic gain control, gating and integration. *Nat. Neurosci.* 17:1031. doi: 10.1038/nn.3764

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.

Copyright © 2020 Tal, Neymotin, Bickel, Lakatos and Schroeder. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

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