HHS Public Access

Author manuscript

IEEE Trans Neural Syst Rehabil Eng. Author manuscript; available in PMC 2021 September 09.

Published in final edited form as:

IEEE Trans Neural Syst Rehabil Eng. 2021 ; 29: 1651–1660. doi:10.1109/TNSRE.2021.3105432.

Region-Level Functional and Effective Network Analysis of Human Brain During Cognitive Task Engagement

Sandeep Avvaru [Graduate Student Member, IEEE],
Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA

Noam Peled,
Athinoula A. Martinos Center for Biomedical Imaging, Department of Radiology, Massachusetts General Hospital, Charlestown, MA 02129 USA

Nicole R. Provenza,
Center for Biomedical Engineering, Brown University School of Engineering, Providence, RI 02912 USA

Alik S. Widge,
Department of Psychiatry and Behavioral Sciences, University of Minnesota, Minneapolis, MN 55455 USA

Keshab K. Parhi [Fellow, IEEE]
Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, MN 55455 USA

Abstract

Mental disorders are a major source of disability, with few effective treatments. It has recently been argued that these diseases might be effectively treated by focusing on decision-making, and specifically remediating decision-making deficits that act as “ingredients” in these disorders. Prior work showed that direct electrical brain stimulation can enhance human cognitive control, and consequently decision-making. This raises a challenge of detecting cognitive control lapses directly from electrical brain activity. Here, we demonstrate approaches to overcome that challenge. We propose a novel method, referred to as *maximal variance node merging* (MVNM), that merges nodes within a brain region to construct informative inter-region brain networks. We employ this method to estimate functional (correlational) and effective (causal) networks using local field potentials (LFP) during a cognitive behavioral task. The effective networks computed using convergent cross mapping differentiate task engagement from background neural activity with 85% median classification accuracy. We also derive *task engagement networks* (TENs): networks that constitute the most discriminative inter-region connections. Subsequent

Corresponding author: Keshab K Parhi. parhi@umn.edu.
Alik S. Widge and Keshab K. Parhi contributed equally to this work.

This article has supplementary downloadable material available at https://doi.org/10.1109/TNSRE.2021.3105432, provided by the authors.

graph analysis illustrates the crucial role of the dorsolateral prefrontal cortex (dlPFC) in task engagement, consistent with a widely accepted model for cognition. We also show that task engagement is linked to prefrontal cortex theta (4-8 Hz) oscillations. We, therefore, identify objective biomarkers associated with task engagement. These approaches may generalize to other cognitive functions, forming the basis of a network-based approach to detecting and rectifying decision deficits.

Keywords

Multi-source interference task; task engagement network; cognitive control; functional connectivity; effective connectivity; maximal variance node merging; local field potential

I. Introduction

NEUROPSYCHIATRIC disorders impose an enormous global disease burden that leads to premature mortality and degraded quality of life. Existing treatments for these disorders are less than 50% effective [1]. Many patients with mental illness do not get relief from gold-standard clinical therapies, resulting in a pressing need for new treatments. Recent studies indicate that these treatments might emerge from measuring and remediating cognitive deficits underpinning mental illness, e.g., through electric stimulation [2–4]. Dysfunctional decision-making and cognitive control are common features in a wide range of mental disorders such as depression, addiction, anxiety disorders, autism, and schizophrenia [5–7]. Cognitive control is a set of interrelated executive functions, including updating (i.e., monitoring working memory), inhibition (resisting prepotent responses), and shifting (switching between mental sets) [8]. It is, therefore, crucial to accommodating daily life requirements, ultimately affecting the quality of life [9]. The ability to remediate cognitive control might thus be a new approach to treating mental disorders.

Cognitive control can be measured in real time through standard laboratory tasks and direct brain stimulation can improve the performance during these tasks [10–12]. It is, however, unclear when this type of intervention should be automatically triggered to achieve a desired psychological effect. Increasing evidence indicates, however, that cognitive dysfunction in mental disorders can be described as aberrant patterns of interactions between neural elements in a large-scale brain network [3], [13–15]. It should therefore be possible to decode this dysfunction by examining changes in functional network communication.

The majority of the existing brain network studies related to cognitive control are limited to functional connectivity [9], [16]. Although some effective connectivity studies exist, they are based on signals acquired through functional magnetic resonance imaging (fMRI) [17], [18]. fMRI is an indirect measure of neural activity and recorded at a low temporal resolution, which is not ideal for capturing cognitive functions that involve finer temporal dynamics. A better approach would be direct decoding of cognitive control and control lapses from electrical brain recordings [11].

A recent attempt to decode cognitive control task engagement reported successful classification [19], [20], but only analyzed functional connectivity. Such correlational

analyses cannot identify the direction of information flow in brain networks, a critical variable for deciding how to stimulate. The prior paper's analysis also suffered from data leakage, which may have led to unreliable evaluation of the classifiers. More specifically, (1) the feature extraction utilized test data that should be independent of the training data, and (2) the training and test data were not temporally separated. Another challenge is that the underlying datasets come from human participants with highly variable placement of their brain electrodes in each brain region. Data interpretation requires methods for measuring and controlling for that variability across participants.

Deriving meaningful causal networks from task-related activity has remained a challenge. Granger causality has been used in numerous neuroscience applications [21]. However, Granger's approach could lead to ambiguity in biological and ecological networks due to non-separable dynamics [22]. This paper uses *directed information* and *convergent cross-mapping*, causal inference techniques that are not based on Granger's theory. The problem of electrode variability can be addressed by merging several channels (electrodes) associated with a region to *form region-level networks*. In most prior research, region-level signals were computed by averaging signals from multiple channels in a region [23]. These averaged signals do not necessarily correlate to the tasks, as classifiers based on these networks fail to achieve high accuracy (as described in Section IV). This paper overcomes this challenge by using the proposed novel *maximal variance node merging* (MVNM) approach. Causal cognitive control networks based on electrophysiological signals have not been constructed before. These networks not only confirm known network properties of cognitive control but also lead to new findings that can be explored further.

The contributions of this paper are five-fold.

1. We construct and compare three different brain networks: one functional and two effective networks). We derive causal networks based on a technique called convergent cross-mapping (CCM) [24] and show that the causal networks help identify regions of interest associated with task engagement. Thus, we utilize distributed brain connectivity analysis to not only detect task engagement, but also identify potential biomarkers for cognitive control.
2. We propose a novel technique referred to as maximal variance node merging (MVNM) to estimate region-level interactions. Unlike channel-level networks, region-level networks are more interpretable and more relevant for clinical translation.
3. We introduce and present *task engagement networks* (TENs) by combining the most explanatory network interactions from multiple subjects. The TENS can be further analyzed to identify significant regions that may be useful as stimulation sites.
4. We demonstrate that the causal inter-region networks can differentiate mental states associated with task performance from resting-state activity with 85.2% median accuracy. A previous analysis using the same data attained 78% accuracy [19].

5. We show that subband networks constructed from bandpass filtered signals also encode task specific activity. Especially, theta band (4–8 Hz) networks play a major role in detecting task engagement, consistent with prior findings that theta band oscillations are modulated during cognitive control.

II. Materials and Methods

A. Subjects

Fourteen subjects participated in the experiment when they were hospitalized for invasive epilepsy monitoring and subsequent seizure localization. Use of intracranial recordings of patients undergoing epilepsy monitoring to study cognitive phenomena is gaining popularity [25–28]. Each subject had a history of drug-resistant complex-partial seizures. We discarded the data from 4 out of the 14 subjects due to lack of statistically sufficient task/non-task recordings. The remaining ten subjects were included this study. All surgical decisions, including the location, type, and the number of electrodes, were made by clinicians independent of this study. The participants were informed that participation in the experiment would not affect their treatment. They were allowed to withdraw at any point during the task. According to the study sponsor guidelines, each participant gave fully informed consent. The original study was approved by the Institutional Review Board of Massachusetts General Hospital and the US Army Human Research Protection Office. The present study re-analyzed a publicly available, de-identified copy of the published dataset, and thus did not require further review [19].

B. The Multi-Source Interference Task

The MSIT is a well-established paradigm that involves multiple dimensions of cognition, including but not limited to attention, object recognition, and decision making, and is a useful tool to study the network basis of cognitive control specifically [29]. The MSIT has been shown to evoke connectivity changes related to cognitive impairment in major depressive disorder (MDD), obsessive compulsive disorder (OCD), and schizophrenia [30–32]. Its sensitivity to electrical stimulation makes it a good candidate to study cognitive control [10], [11]. In an MSIT trial (see Fig. 1), the participants were presented with a fixation cross for 2 s followed by a stimulus (until response) in the form of 3 digits – one of which is the ‘target.’ Two of the three numbers, known as ‘distractors,’ have the same value. The target is either 1, 2, or 3. Each trial is categorized as either ‘congruent’ or ‘incongruent’ depending on its difficulty level. In the congruent condition, the distractors are always ‘0’, and the target’s position matches its value. In the incongruent (also known as interference) condition, the distractors are picked from potential targets, and the position of the unique target is different from its keyboard position (Simon effect). In a successful trial, the participant reports, via a button press on a keypad, the target’s value regardless of its position. Examples of congruent and incongruent trials are presented in Fig. 1(a).

The experiment comprised up to 5 blocks of trials, each with approximately 32 or 64 trials in each block (Fig. 1(b)). Signals recorded between the blocks, before the first block, and after the last block were labeled as non-task data. The task contained a roughly equal number

of congruent and incongruent trials. Median success rates of 100%±2.47% and 97.1±5.52% were reported during congruent and incongruent conditions, respectively [19].

C. Signal Acquisition and Preprocessing

1) Data Acquisition: Local field potential (LFP) signals were recorded through depth electrodes surgically implanted for seizure monitoring in each participant. Between five and nine electrodes with diameters 0.8-1.0 mm were placed in each hemisphere. Each electrode consisted of 8-16 platinum/iridium contacts. The distribution of electrodes is illustrated in Fig. 2. The Multi-Modality Visualization Tool was used to create the visualization [33]. The signals were acquired at a 2 kHz sampling rate via neural signal processor recording systems from Blackrock Microsystems Inc., Salt Lake City, UT. All signals were referenced to a scalp EEG electrode. Electrodes with excessive line noise (60 Hz), close to seizure focus (based on clinical reports), and other artifacts found on visual inspection were removed. Each channel was down-sampled to 1000 Hz. The line noise and its harmonics were removed. Adjacent channels were then bipolar re-referenced to each other to alleviate the effect of volume conduction [34]. This data was previously reported in [19], and was analyzed using a different approach.

2) Electrode Localization: Spatial coordinates of the electrodes were determined manually through post-operative computerized tomography (CT). Pre-operative T1 weighted MRI were aligned with the anatomical CT images through a volumetric image co-registration method utilizing the FreeSurfer software package [35], [36]. An electrode labeling algorithm was employed to estimate the probability that a particular brain region contributes to the signal's source at each electrode [37]. The regions of interest were parcellated based on the Desikan-Killiany-Tourville brain atlas [38]. The number of bipolar re-referenced channels ranged between 64 and 195 based on the subjects' electrode montage, and these channels were mapped to 17-23 regions.

D. Defining Task and Non-Task Segments

The neural activity recorded during MSIT blocks is referred to as *task data*, and the data recorded during rest periods (before or after task blocks) is referred to as *non-task data*. To differentiate between the task and non-task states, the signals were divided into multiple task segments and non-task segments. On average, each MSIT trial was approximately 4 s long; each trial's actual duration varies depending on the subject's reaction time. The time when the fixation cross was presented during a trial was marked as the trial's start. The time duration between two consecutive fixation crosses determined the length of each trial. The minimum trial length for most subjects was approximately 3.8 s. Therefore, the 3.8-second time segments from every trial's start were labeled as 'task' data. The signals recorded during rest periods were windowed into 'non-task' segments with a window length equal to the minimum task duration.

Overlapping windows were used when the number of non-task segments was less than the number of task segments. If the amount of task data was less than the amount of non-task data, only a subset of the non-task data was used for classification. Thus, the two classes were balanced to make sure the classifiers were not biased. The multidimensional

time-series corresponding to task and non-task segments were then used to construct task and non-task networks.

E. Connectivity Measures

Network science provides a particularly appropriate framework to study brain mechanisms by treating neural elements (a population of neurons, a subregion) as nodes in a graph and neural interactions (synaptic connections, information flow) as its edges [39]. *Structural connectivity* represents large-scale anatomical connections between cortical regions. *Functional* and *effective* connectivities are generally estimated from the time-series of brain signals. A fundamental distinction between the two is that the edge weights in *functional networks* correspond to cross-correlation coefficients, while those in *effective* networks correspond to patterns of causal interactions. Quantifying (correlative or causative) interactions between time-series is of particular interest in studying complex network systems such as the brain [40]. This study estimates and analyzes functional connectivity and effective connectivity between the electrodes and the regions. The effective networks were constructed using two methods: directed information (DI) and convergent cross-mapping (CCM). A review of these two methods is provided in the Supplementary Information (see Sections S1 and S2).

Instead of assuming parametric models such as the auto-regressive model used by Granger causality, DI is based on information theory [41]. It can therefore measure nonlinear interactions and is not dependent on accurate estimation of model parameters. A benefit of using CCM for causal inference is that, unlike Granger-related causality measures, CCM does not assume ‘separability’ between variables [24]. The separability assumption is usually valid in linear and strongly coupled non-linear systems. Complex subsystems such as the brain are characterized by moderate to weak coupling and are affected by unobserved/external variables.

F. Maximal Variance Node Merging (MVNM) and Region-Level Networks

A significant challenge in this specific application space, read-out of cognition from distributed electrodes, is an imbalance in the number of measurements between nodes/edges. A given brain region (node) may have anywhere from 1 to 5 physical electrodes measuring it, depending on the size of the region and the specific clinical placement of the electrode. This imbalance makes it harder to interpret channel-level networks, where each node corresponds to a specific electrode contact. Similarly, since electrode positions vary between subjects, it is unclear which channels can be safely averaged/combined. That combination is achievable if we can identify channels as belonging to specific brain regions, such that we can work in terms of the dominant signal within each region and the inter-region interactions.

We introduce maximal variance node merging (MVNM) as an approach to combine nodes in a channel-level network that were mapped to the same brain region to generate a region-level network. Each node in a region-level network is associated with a brain region. This enables us to interpret the network connections better and, consequently, detect specific task engagement regions.

First, the channels were organized into their corresponding regions based on electrode localization results. Then, region-level networks were constructed from channel-level networks using MVNM. For each pair of regions, all the network connections between the regions (inter-region connections) were replaced by one representative connection in the case of undirected networks. Note that intra-region connections were discarded. There would be two resultant connections in directed networks: one representing all the edges from region-A to region-B and the other for the edges from region-B to region-A. Fig. 3(a) and Fig. 3(b) depict an example network before and after the node merging process, respectively.

These new inter-region edges were estimated by computing the optimal linear combination of the original edges to maximize the variance over time. This process is equivalent to finding the first principal component of the original edges between the two regions using principal component analysis (PCA) – a dimensionality reduction method that attempts to reduce the number of variables while preserving as much information as possible. This approach is beneficial since the goal is to use these edges as features for classification. The MVNM algorithm's rationale is that a larger variance implies a broader spread in the feature space, which helps find the decision boundary.

Fig. 3(c) illustrates this process for two regions A and B with m and n channels, respectively. There can be mn edges between the two regions in an undirected network (functional) and $2mn$ edges in a directed network (effective). Intra-region connections can be ignored here. If there are N such networks in the dataset, it can be represented by a matrix of dimensions $N \times mn$, denoted by $\mathbf{X}$. Note that N is the number of time-series segments (task and non-task) extracted for a given subject. There would be two such matrices, $\mathbf{X_{in}}$ and $\mathbf{X_{out}}$, for effective networks. The optimal linear combination of the columns of $\mathbf{X}$ that maximizes its variance is given by $\mathbf{Xw}$, where the $mn \times 1$ weight vector $\mathbf{w}$ is computed as:

$$\mathbf{w} = \underset{\|w\| = 1}{\arg\max} \left\{ \mathbf{w}^T \mathbf{X}^\mathbf{T} \mathbf{X} \mathbf{w} \right\}. \quad (1)$$

The optimal linear combination of these connections, $\mathbf{Xw}$, provides the inter-region edge strengths for the N networks. Note that $\mathbf{Xw}$ is the first principal component of $\mathbf{X}$. In effective networks, the incoming and outgoing connections were processed separately to determine the two inter-region connections. It can be noticed that the adjacency matrices in effective networks (Fig. 4(b) and Fig. 4(c)) are not symmetric, indicating the directionality of the networks.

G. Edge Importance Score and Task Engagement Networks (TENs)

Once all the networks were computed, PCA of the edges was used for feature extraction. Task engagement networks were constructed from the most significant connections that were effective in identifying MSIT states. Note that principal components of the network edges were used as the features for classification. The PCA assigns a weight to each network connection to compute optimal principal components. These PCA coefficients were used as an indicator of the significance of the connections. For a given connection between two regions, the sum of its coefficients linked to the top p features (principal components) is

defined as its *importance score*. The edges were then ranked based on their importance scores and the q highest-ranked edges were deemed significant edges – represented by the set S_{sub} for a given subject *sub*. Parameters $p = 10$ and $q = 60$ were used in our analysis. Note that p and q were chosen such that the choice of these parameters leads to high classification accuracy (see subsection below). This algorithm is depicted in Fig. 5.

A 10-fold cross-validation split was performed on the data before the PCA as shown in Fig. 5. Since the cross validation results in ten models for each subject, S_{sub} is the intersection of significant edges from the ten models. This is given by the expression,

$$S_{sub} = \bigcap_{fold} S_{sub,\, fold}.$$

This ensures that the edges are not specific to a subset of the data and generalize across the dataset. The mean network of all S_{sub} constitutes the task engagement network S_{TEN}. That is, the edge strength e_{AB} from regions A to B is defined as

$$e_{AB} = \frac{\Sigma_{sub} e_{AB,\, sub}}{|sub|}, \tag{2}$$

where $e_{AB,sub} \in S_{sub}$ and |sub| is the number of subjects. Therefore, the edges in the TEN signify their prevalence among multiple subjects.

H. Task vs. Non-Task Classification

The PCA features were used as inputs to *linear* support vector machine (SVM) classifiers to distinguish task data from non-task data. For a given subject, training and testing sets for each of the ten folds were selected via sequential sub-sampling to ensure that test samples are not in the temporal vicinity of the training samples. Essentially, data from each class are sequentially partitioned into ten subsets (folds). In each iteration, one of the ten folds is used for testing, while the other nine are used for training. This setup ensures that all the data are tested, and the classifiers are less prone to overfitting. Fig. 5 depicts the classification process. Classifiers with a varying number of inputs were trained with a maximum limit of 100 features and the testing accuracy with the optimal number of features is reported.

I. Subband Networks

Neurophysiological studies indicate that theta activity increases with the need for cognitive control [42–45]. We constructed subband networks to analyze the network interactions of frequency-specific neural activity. First, the neural recordings were bandpass filtered into five pre-defined subbands to quantify the role of individual frequency bands on task engagement. These bandpass filtered signals were then used to construct band-specific networks, named *subband networks*. The frequency bands considered in the study are theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), gamma-1 (30-55 Hz), and gamma-2 (65-100 Hz). The bandpass filtering was implemented using 6^{th} order Butterworth IIR filters. All signals were filtered bidirectionally to avoid undesired phase shifts introduced by the filtering, which can affect causal inference.

III. Results

A. Identifying Task States

The SVM models were evaluated based on how accurately they could distinguish between task and non-task states for all subjects. Their classification accuracy, sensitivity (true-positive rate), and specificity (true-negative rate) were calculated; the task data was considered as the positive class. Fig. 6 presents the SVM classification accuracy for all the subjects evaluated based on the three network construction approaches. The reported values represent the mean accuracy over 10-fold cross-validation. The plot illustrates that all network types attain accuracies substantially higher than the baseline of 50%, suggesting that there are patterns of interaction in the neural activity that can be used to identify task engagement. CCM networks perform the best with 85.2±5.0% accuracy. The high accuracy indicates that inter-region interactions contain useful task-related information. The median accuracy and interquartile range values across the ten subjects are summarized in Table I. The median accuracy, sensitivity, and specificity of all MVNM-based networks exceed 80%. The interquartile range values are also low, indicating the algorithm's reliability and robustness across multiple human subjects. All three methods outperform fixed canonical coherence analysis (FCHA) presented in [19].

B. Task Engagement Networks

Fig. 7 depicts the TENs from functional (R) and effective networks (DI and CCM). Each edge in the graphs represents the number of times a specific connection appears in S_{TEN}. Out of the fourteen regions of interest that emerged from the analysis, connections between some regions are more prominent than the others. More interestingly, the graph visualizations in Fig. 7(c) showcase a close resemblance to the results in Fig. 7(b), which are based on an independent causal approach. Task engagement networks for individual subjects are presented in the Supplementary Information (see Section S4).

To quantify the importance of each of the regions in task engagement, we measured the node centralities of the regions in each TEN (see Fig. 8). Node degree and outdegree were computed for functional and (two) effective networks, respectively. The dorsolateral PFC (dlPFC) and temporal lobe show more centrality in all three cases, although the distinction is more prominent in effective networks. In correlation networks, the difference between regions is less noticeable, making it harder to discriminate key hubs using just correlation analysis. However, dlPFC has a considerably higher outdegree in effective networks estimated from causal interactions, followed by temporal lobe and ventrolateral PFC (vlPFC).

C. Analysis of Networks for Specific Frequency Bands

1) Increased Theta Band Activity During Task Performance: Greater theta band activity is known to be associated with cognitive control [27], [42–46]. Using bandpass filtering, the neural activity in the task and non-task periods was divided into five subbands: theta (4-8 Hz), alpha (8-13 Hz), beta (13-30 Hz), gamma-1 (30-55 Hz), and gamma-2 (65-100 Hz). Relative power in the theta band was computed as a ratio of the spectral power in the 4-8 Hz frequency range to the total power spectral density in 4-100 Hz. Fig.

9 shows a distribution of relative power in theta band in the left dlPFC of subject-2 in task vs. non-task segments. Theta power during the task and non-task states has a log-normal distribution, with means 0.28 and 0.21, respectively, demonstrating an increased theta band power during this cognitive control task. We also observed an enhanced dlPFC theta activity in seven out of the ten subjects. The increased mean relative power in the seven subjects is also associated with p-values <0.05 and Bayes factor >9 (in 6 of the 7 cases) according to the two-sample t-test. This is consistent with a wide range of prior reports implicating modulated theta-band activity during cognitive control [42–47].

D. Theta Band Network Interactions Differentiate Task and Non-Task States

To evaluate the role of inter-region subband activity of the five subbands in the MSIT, we decode the task states using the subband networks. The aim is to discern and compare the discriminative ability of network dynamics associated with each of the five subbands. Since CCM networks have the best performance, CCM based subband networks are used. We observe that the median classification accuracy of the subband networks is 79.5% – marginally less than the networks without frequency filtering. This implies that subband dynamics also encode task-related information. By measuring the number of significant edges corresponding to each subband, we observe that 4-8 Hz activity distinguishes task states better than other frequencies. This is illustrated in Fig. 10.

TENs built using only significant theta subband edges also highlight the influence of dlPFC and temporal lobes in the MSIT, as shown in Fig. 11. The dorsomedial PFC (dmPFC) attains the third highest centrality. This shows that theta oscillations (especially in dlPFC) and their network-level interactions can act as biomarkers for task engagement.

IV. Discussion

Despite its pressing need, devising an effective neurological mechanism-based treatment to enhance cognitive control is still a major challenge. This might be achieved through adaptive brain stimulation that intervenes when cognitive control lapses [10], [11], [48]. Developing such a treatment requires a deeper understanding of cognitive control that may be provided by network analysis.

Authors in [19] describe a unique approach to construct inter-region networks using canonical correlation analysis, called fixed canonical coherence analysis (FCHA). Even though both FCHA and MVNM involve estimating a linear combination of multiple variables, there are fundamental distinctions between them. For a given pair of regions X and Y, the FCHA maximizes the coherence between multivariate time-series in X and Y. This optimization is accomplished by estimating the optimal linear combination of multiple channels within each region. MVNM is a novel method that optimizes the variance of the network interactions over time. Unlike FCHA, MVNM converts (possibly informative) channel-level networks into more interpretable region-level networks. This method can be applied to correlational or causal networks.

Averaging time-series recordings is a common approach to combine neural signals from the same anatomical region [23], [49]. Such techniques may discard valuable network

dynamics. Classification results presented in the Supplementary Information (see Section S3) demonstrate that the averaged connectivity is not well-suited to uncover the optimal causal interactions between the regions. Table S1 shows that MVNM attains a 10% increase in accuracy compared to the time-series averaging. The proposed methods help detect task engagement with approximately 85% median accuracy, a notable improvement over [19]. In addition, the classifiers are characterized by low variation across cross-validation folds and across subjects indicating the robustness of the approach. However, it should be noted that these classifiers are subject-specific. A generalized decoder for all subjects may lead to lower accuracy.

A high classification accuracy may not always lead to advances in understanding underlying brain mechanisms. Hence, we use graph analysis of the networks to determine TENs presented in Fig. 7. TENs highlight changes in network interactions across the subjects due to task-related effort. The causal TENs highlight dorsolateral PFC, temporal lobe, and ventrolateral PFC as the most influential regions of interest with high outdegree. The observations from CCM and DI networks – independent causality measures – corroborate each other, supporting the validity of our approach. The involvement of dlPFC in cognitive control, especially in tasks involving conflict or inhibition of irrelevant information, is reported in prior research [10], [25], [50]. Schizophrenia and OCD have been linked to dysfunctional modulation in dlPFC and vlPFC [51], [52]. Our TENs emphasize the activation of dlPFC better than [19]. dlPFC was also used as a stimulation site to enhance cognitive control successfully [53]. Several studies suggest that increased frontal theta activity is associated with cognitive control [10], [42], [43], [45], [54]. Fig. 10 shows that the majority of the discriminative features used in the classifiers originate from theta band activity in all the subjects. Fig. 11 implies that theta interactions in dlPFC, dmPFC and the temporal lobe play an important role in cognition. The alignment of our data-driven findings with prior studies validates the ability of our approach to discover true mechanistic network structure.

The role of the temporal lobe, which is outside the canonical frontoparietal cognitive control networks, is not entirely understood [55], [56]. This is partly because most cognitive control studies are focused on the frontal regions instead of observing a global brain network. It is shown in [57] that the temporal lobe had the largest response to a cognitive task among non-frontal regions. Alternately, the temporal lobe weighting could be related to the participants' epilepsy. Patients with temporal lobe epilepsy have been reported to suffer from dysfunctional control characterized by interactions between the epileptogenic temporal lobe and the PFC [58], [59].

Even though the MSIT has been shown to activate cognition/attention networks [29–31], clinical translation towards a viable treatment for psychiatric disorders presents significant challenges. First, any task performance can include several other behavioral and physiological mechanisms that may not be related to cognitive effort. For example, task engagement can provoke anxiety, leading to engagement of emotional arousal networks. Therefore, further research involving multiple tasks and a larger cohort of subjects is needed to form robust conclusions about the neural encoding of cognitive control. Another challenge is that the efficacy of adaptive DBS on psychiatric disorders is not sufficiently

understood. We believe that objective network biomarkers at an electrophysiological level can help detect and rectify cognitive dysfunction [3], [11]. Future studies may also explore non-invasive recording/stimulating modalities

V. Conclusion

Network representations can decode cognitive control and other mental functions by identifying relevant cross-region interactions. We show that network connections between dorsolateral PFC, temporal lobe and ventrolateral PFC were discriminative through graph analysis of causal task engagement networks. Moreover, independent causal inference techniques (DI and CCM) indicate a higher outdegree in those regions, supporting the potential for dlPFC as a stimulation site [3], [12], [53], [60]. Subband network analysis reveals that enhanced cognitive control is associated with modulated theta band activity.

There is substantial evidence supporting stimulation-induced modulation of pathological network activity as a therapeutic mechanism of treatments such as deep brain stimulation [61–63]. The methods developed in this paper enable the discovery of objective biomarkers, localized regions of interest, and band oscillatory activity associated with task engagement. This knowledge can facilitate a transition from symptom-based treatments to more effective mechanism-based treatments for mental illness.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The collection of data was supported by the Defense Advanced Research Projects Agency (DARPA) under Cooperative Agreement W911NF-14-2-0045 issued by the ARO Contracting Office through DARPA's SUBNETS Program. This research was supported in part by the National Science Foundation under Grant CCF-1954749. The work of Sandeep Avvaru was supported by the MNDrive Neuromodulation Research Fellowship. The work of Alik S. Widge was supported in part by the MNDrive Brain Conditions Program, in part by the UMN Medical Discovery Team on Addiction, and in part by the National Institutes of Health under Grant R01NS113804 and Grant R01MH123634.

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by Massachusetts General Hospital.

References

[1]. Patel Vet al., "Addressing the burden of mental, neurological, and substance use disorders: Key messages from disease control priorities, 3rd edition," Lancet, vol. 387, no. 10028, pp. 1672–1685, 4. 2016. [PubMed: 26454360]

[2]. Widge ASet al., "Treating refractory mental illness with closed-loop brain stimulation: Progress towards a patient-specific transdiagnostic approach," Exp. Neurol, vol. 287, pp. 461–472, 1. 2017. [PubMed: 27485972]

[3]. Sullivan CRP, Olsen S, and Widge AS, "Deep brain stimulation for psychiatric disorders: From focal brain targets to cognitive networks," NeuroImage, vol. 225, 1. 2021, Art. no. 117515.

[4]. Huys QJM, Maia TV, and Frank MJ, "Computational psychiatry as a bridge from neuroscience to clinical applications," Nature Neurosci, vol. 19, no. 3, pp. 404–413, 3. 2016. [PubMed: 26906507]

[5]. Goschke T, "Dysfunctions of decision-making and cognitive control as transdiagnostic mechanisms of mental disorders: Advances, gaps, and needs in current research," Int. J. Methods Psychiatric Res, vol. 23, no. S1, pp. 41–57, 1. 2014.

[6]. Arnsten AFT and Rubia K, "Neurobiological circuits regulating attention, cognitive control, motivation, and emotion: Disruptions in neurodevelopmental psychiatric disorders," J. Amer. Acad. Child Adolescent Psychiatry, vol. 51, no. 4, pp. 356–367, 4. 2012

[7]. McTeague LM, Huemer J, Carreon DM, Jiang Y, Eickhoff SB, and Etkin A, "Identification of common neural circuit disruptions in cognitive control across psychiatric disorders," Amer. J. Psychiatry, vol. 174, no. 7, pp. 676–685, 7. 2017. [PubMed: 28320224]

[8]. Gratton G, Cooper P, Fabiani M, Carter CS, and Karayanidis F, "Dynamics of cognitive control: Theoretical bases, paradigms, and a view for the future," Psychophysiology, vol. 55, no. 3, 3. 2018, Art. no. e13016.

[9]. Shi Let al., "Brain networks of happiness: Dynamic functional connectivity among the default, cognitive and salience networks relates to subjective well-being," Social Cogn. Affect. Neurosci, vol. 13, no. 8, pp. 851–862, 8. 2018.

[10]. Widge ASet al., "Deep brain stimulation of the internal capsule enhances human cognitive control and prefrontal cortex function," Nature Commun, vol. 10, no. 1, p. 1536, 12. 2019. [PubMed: 30948727]

[11]. Basu Iet al., "Closed loop enhancement and neural decoding of human cognitive control," bioRxiv, 4. 2020, doi: 10.1101/2020.04.24.059964.

[12]. Dubreuil-Vall L, Chau P, Ruffini G, Widge AS, and Camprodon JA, "TDCS to the left DLPFC modulates cognitive and physiological correlates of executive function in a state-dependent manner," Brain Stimulation, vol. 12, no. 6, pp. 1456–1463, 11. 2019. [PubMed: 31221553]

[13]. Bullmore E and Sporns O, "The economy of brain network organization," Nature Rev. Neurosci, vol. 13, pp. 336–349, 4. 2012. [PubMed: 22498897]

[14]. Sen B, Bernstein GA, Mueller BA, Cullen KR, and Parhi KK, "Sub-graph entropy based network approaches for classifying adolescent obsessive-compulsive disorder from resting-state functional MRI," NeuroImage: Clin, vol. 26, 1. 2020, Art. no. 102208.

[15]. Xu Tet al., "Network analysis of functional brain connectivity in borderline personality disorder using resting-state fMRI," NeuroImage: Clin, vol. 11, pp. 302–315, 1. 2016. [PubMed: 26977400]

[16]. Marek S, Hwang K, Foran W, Hallquist MN, and Luna B, "The contribution of network organization and integration to the development of cognitive control," PLoS Biol, vol. 13, no. 12, 12. 2016, Art. no. e1002328.

[17]. Ham T, Leff A, de Boissezon X, Joffe A, and Sharp DJ, "Cognitive control and the salience network: An investigation of error processing and effective connectivity," J. Neurosci, vol. 33, no. 16, pp. 7091–7098, 4. 2013. [PubMed: 23595766]

[18]. Dobryakova E, Rocca MA, Valsasina P, DeLuca J, and Filippi M, "Altered neural mechanisms of cognitive control in patients with primary progressive multiple sclerosis: An effective connectivity study," Hum. Brain Mapping, vol. 38, no. 5, pp. 2580–2588, 5 2017.

[19]. Provenza NRet al., "Decoding task engagement from distributed network electrophysiology in humans," J. Neural Eng, vol. 16, no. 5, 8. 2019, Art. no. 056015.

[20]. Avvaru S, Provenza N, Widge A, and Parhi KK, "Decoding human cognitive control using functional connectivity of local field potentials," in Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Oct./Nov. 2021.

[21]. Granger CWJ, "Investigating causal relations by econometric models and cross-spectral methods," Econometrica, vol. 37, no. 3, pp. 424–438, 8. 1969.

[22]. Stokes PA and Purdon PL, "A study of problems encountered in Granger causality analysis from a neuroscience perspective," Proc. Nat. Acad. Sci. USA, vol. 114, no. 34, pp. E7063–E7072, 8. 2017. [PubMed: 28778996]

[23]. Stanley ML, Moussa MN, Paolini BM, Lyday RG, Burdette JH, and Laurienti PJ, "Defining nodes in complex brain networks," Frontiers Comput. Neurosci, vol. 7, p. 169, 11. 2013.

[24]. Sugihara Get al., "Detecting causality in complex ecosystems," Science, vol. 338, no. 6106, pp. 496–500, 10. 2012. [PubMed: 22997134]

[25]. Smith Eet al., “Widespread temporal coding of cognitive control in the human prefrontal cortex,” Nature Neurosci, vol. 22, pp. 1883–1891, 9. 2019. [PubMed: 31570859]

[26]. Ezzyat Yet al., “Closed-loop stimulation of temporal cortex rescues functional networks and improves memory,” Nature Commun, vol. 9, no. 1, p. 365, 12. 2018. [PubMed: 29410414]

[27]. Smith EHet al., “Frequency-dependent representation of reinforcement-related information in the human medial and lateral prefrontal cortex,” J. Neurosci, vol. 35, no. 48, pp. 15827–15836, 12. 2015. [PubMed: 26631465]

[28]. Oehrn CRet al., “Neural communication patterns underlying conflict detection, resolution, and adaptation,” J. Neurosci, vol. 34, no. 31, pp. 10438–10452, 7. 2014. [PubMed: 25080602]

[29]. Bush G, Shin LM, Holmes J, Rosen BR, and Vogt BA, “The multi-source interference task: Validation study with fMRI in individual subjects,” Mol. Psychiatry, vol. 8, no. 1, pp. 60–70, 1. 2003. [PubMed: 12556909]

[30]. Davey CG, Yücel M, Allen NB, and Harrison BJ, “Task-related deactivation and functional connectivity of the subgenual cingulate cortex in major depressive disorder,” Frontiers Psychiatry, vol. 3, p. 14, 2. 2012.

[31]. Cocchi Let al., “Functional alterations of large-scale brain networks related to cognitive control in obsessive-compulsive disorder,” Hum. Brain Mapping, vol. 33, pp. 1089–1106, 5 2012.

[32]. Heckers S, Weiss AP, Deckersbach T, Goff DC, Morecraft RJ, and Bush G, “Anterior cingulate cortex activation during cognitive interference in schizophrenia,” Amer. J. Psychiatry, vol. 161, no. 4, pp. 707–715, 4. 2004. [PubMed: 15056518]

[33]. Felsenstein Oet al., “Multi-modal neuroimaging analysis and visualization tool (MMVT),” 2019, arXiv:1912.10079. [Online]. Available: https://arxiv.org/abs/1912.10079

[34]. Bastos AM and Schoffelen J-M, “A tutorial review of functional connectivity analysis methods and their interpretational pitfalls,” Frontiers Syst. Neurosci, vol. 9, p. 175, 1. 2016.

[35]. Dykstra ARet al., “Individualized localization and cortical surface-based registration of intracranial electrodes,” NeuroImage, vol. 59, no. 4, pp. 3563–3570, 2. 2012. [PubMed: 22155045]

[36]. Freesurfer. Accessed: Jun. 2020. [Online]. Available: http://surfer.nmr.mgh.harvard.edu/

[37]. Invasive Electrodes Identification and Labeling. Accessed: Jul. 2020. [Online]. Available: https://github.com/pelednoam/ieil

[38]. Desikan RSet al., “An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest,” NeuroImage, vol. 31, no. 3, pp. 968–980, 7. 2006. [PubMed: 16530430]

[39]. Yu Qet al., “Application of graph theory to assess static and dynamic brain connectivity: Approaches for building brain graphs,” Proc. IEEE, vol. 106, no. 5, pp. 886–906, 5 2018.

[40]. Bullmore E and Sporns O, “Complex brain networks: Graph theoretical analysis of structural and functional systems,” Nature Rev. Neurosci, vol. 10, pp. 98–186, 2. 2009.

[41]. Malladi R, Kalamangalam G, Tandon N, and Aazhang B, “Identifying seizure onset zone from the causal connectivity inferred using directed information,” IEEE J. Sel. Topics Signal Process, vol. 10, no. 7, pp. 1267–1283, 10. 2016.

[42]. Cooper PSet al., “Frontal theta predicts specific cognitive control-induced behavioural changes beyond general reaction time slowing,” NeuroImage, vol. 189, pp. 130–140, 4. 2019. [PubMed: 30639331]

[43]. Cavanagh JF and Frank MJ, “Frontal theta as a mechanism for cognitive control,” Trends Cogn. Sci, vol. 18, no. 8, pp. 414–421, 8. 2014. [PubMed: 24835663]

[44]. Tang etal H., “Cascade of neural processing orchestrates cognitive control in human frontal cortex,” eLife, vol. 5, 2. 2016, Art. no. e12352.

[45]. Angelidis A, Hagenaars M, van Son D, van der Does W, and Putman P, “Do not look away! Spontaneous frontal EEG theta/beta ratio as a marker for cognitive control over attention to mild and high threat,” Biol. Psychol, vol. 135, pp. 8–17, 5 2018. [PubMed: 29518523]

[46]. Cooper PSet al., “Theta frontoparietal connectivity associated with proactive and reactive cognitive control processes,” NeuroImage, vol. 108, pp. 354–363, 3. 2015. [PubMed: 25528657]

[47]. Avvaru S, Provenza N, Widge A, and Parhi KK, “Spectral features based decoding of task engagement: The role of theta and high gamma bands in cognitive control,” in Proc. 43rd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC), Oct./Nov. 2021.

[48]. Koenig Aet al., “Real-time closed-loop control of cognitive load in neurological patients during robot-assisted gait training,” IEEE Trans. Neural Syst. Rehabil. Eng, vol. 19, no. 4, pp. 453–464, 8. 2011. [PubMed: 21827971]

[49]. Chu S-H, Parhi KK, and Lenglet C, “Function-specific and enhanced brain structural connectivity mapping via joint modeling of diffusion and functional MRI,” Sci. Rep, vol. 8, no. 1, p. 4741, 12. 2018. [PubMed: 29549287]

[50]. MacDonald AW, “Dissociating the role of the dorsolateral prefrontal and anterior cingulate cortex in cognitive control,” Science, vol. 288, no. 5472, pp. 1835–1838, 6. 2000. [PubMed: 10846167]

[51]. Vaghi MM. et al., “Specific frontostriatal circuits for impaired cognitive flexibility and goal-directed planning in obsessive-compulsive disorder: Evidence from resting-state functional connectivity,” Biol. Psychiatry, vol. 81, no. 8, pp. 708–717, 4. 2017. [PubMed: 27769568]

[52]. Lesh TAet al., “Proactive and reactive cognitive control and dorsolateral prefrontal cortex dysfunction in first episode schizophrenia,” NeuroImage: Clin, vol. 2, pp. 590–599, 1. 2013. [PubMed: 24179809]

[53]. Metuki N, Sela T, and Lavidor M, “Enhancing cognitive control components of insight problems solving by anodal tDCS of the left dorsolateral prefrontal cortex,” Brain Stimulation, vol. 5, no. 2, pp. 110–115, 4. 2012. [PubMed: 22483547]

[54]. Dvorak D, Shang A, Abdel-Baki S, Suzuki W, and Fenton AA, “Cognitive behavior classification from scalp EEG signals,” IEEE Trans. Neural Syst. Rehabil. Eng, vol. 26, no. 4, pp. 729–739, 4. 2018. [PubMed: 29641377]

[55]. Zanto TP and Gazzaley A, “Fronto-parietal network: Flexible hub of cognitive control,” Trends Cognit. Sci, vol. 17, no. 12, pp. 602–603, 12. 2013. [PubMed: 24129332]

[56]. Harding IH, Yücel M, Harrison BJ, Pantelis C, and Breakspear M, “Effective connectivity within the frontoparietal control network differentiates cognitive control and working memory,” NeuroImage, vol. 106, pp. 144–153, 2. 2015. [PubMed: 25463464]

[57]. Kerns J, Cohen J, MacDonald A, Cho R, Stenger V, and Carter C, “Anterior cingulate conflict monitoring and adjustments in control,” Science, vol. 303, pp. 1023–1026, 2. 2004. [PubMed: 14963333]

[58]. Stretton J and Thompson PJ, “Frontal lobe function in temporal lobe epilepsy,” Epilepsy Res, vol. 98, no. 1, pp. 1–13, 1. 2012. [PubMed: 22100147]

[59]. Hudson JM, Flowers KA, and Walster KL, “Attentional control in patients with temporal lobe epilepsy,” J. Neuropsychol, vol. 8, no. 1, pp. 140–146, 3. 2014. [PubMed: 23323970]

[60]. Widge AS and Miller EK, “Targeting cognition and networks through neural oscillations: Next-generation clinical brain stimulation,” JAMA Psychiatry, vol. 76, no. 2, pp. 671–672, 7. 2019. [PubMed: 31116372]

[61]. Anidi Cet al., “Neuromodulation targets pathological not physiological beta bursts during gait in Parkinson’s disease,” Neurobiol. Disease, vol. 120, pp. 107–117, 12. 2018.

[62]. Dippel G, Mückschel M, Ziemssen T, and Beste C, “Demands on response inhibition processes determine modulations of theta band activity in superior frontal areas and correlations with pupillometry–Implications for the norepinephrine system during inhibitory control,” NeuroImage, vol. 157, pp. 575–585, 8. 2017. [PubMed: 28647483]

[63]. Mcintyre C, Grill W, Sherman D, and Thakor N, “Cellular effects of deep brain stimulation: Model-based analysis of activation and inhibition,” J. Neurophysiol, vol. 91, no. 4, pp. 1457–1469, 5 2004. [PubMed: 14668299]

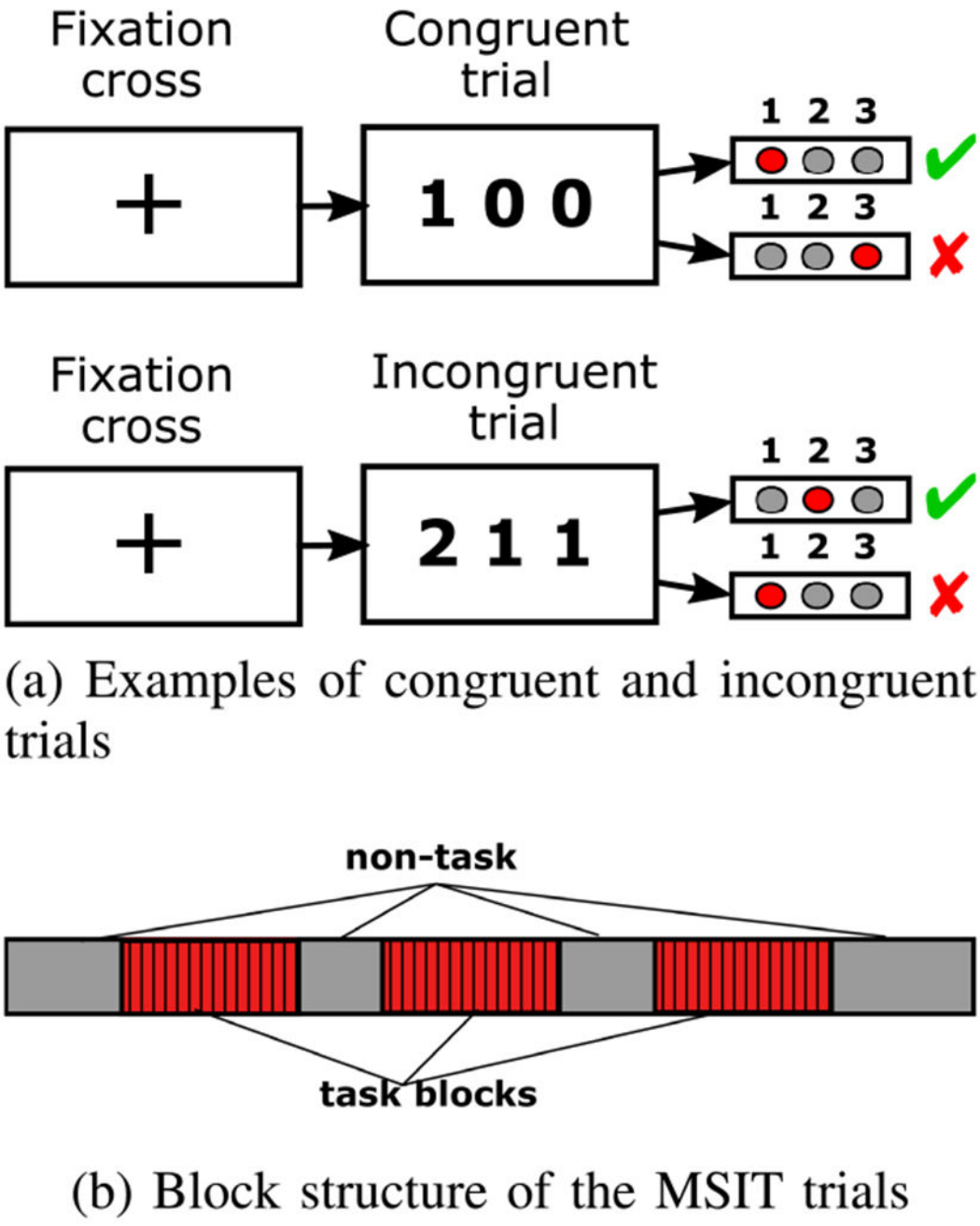


Fig. 1.
The multi-source interference task.

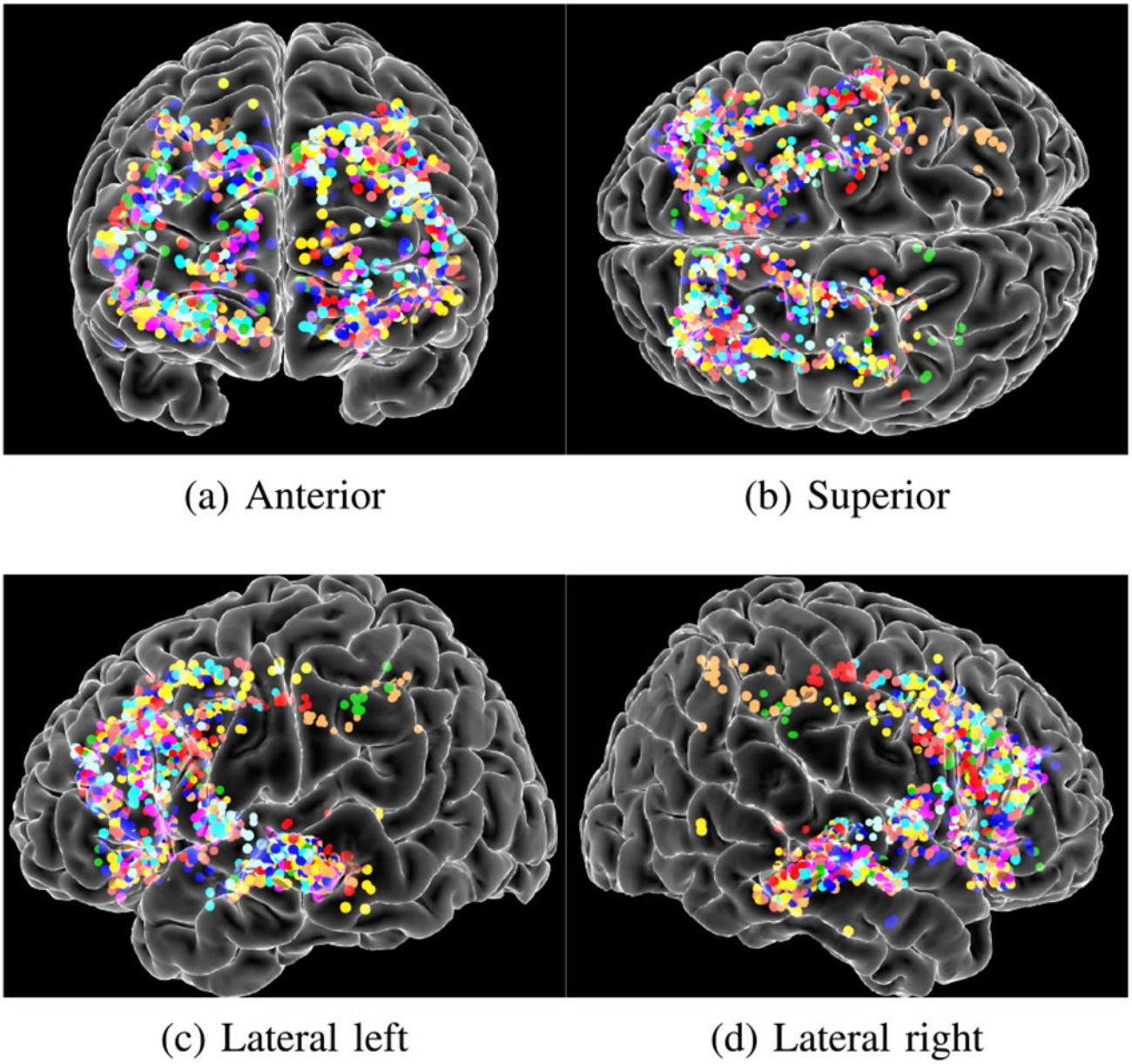


Fig. 2.
Glass brain models showing the electrode locations. Colors represent different subjects.

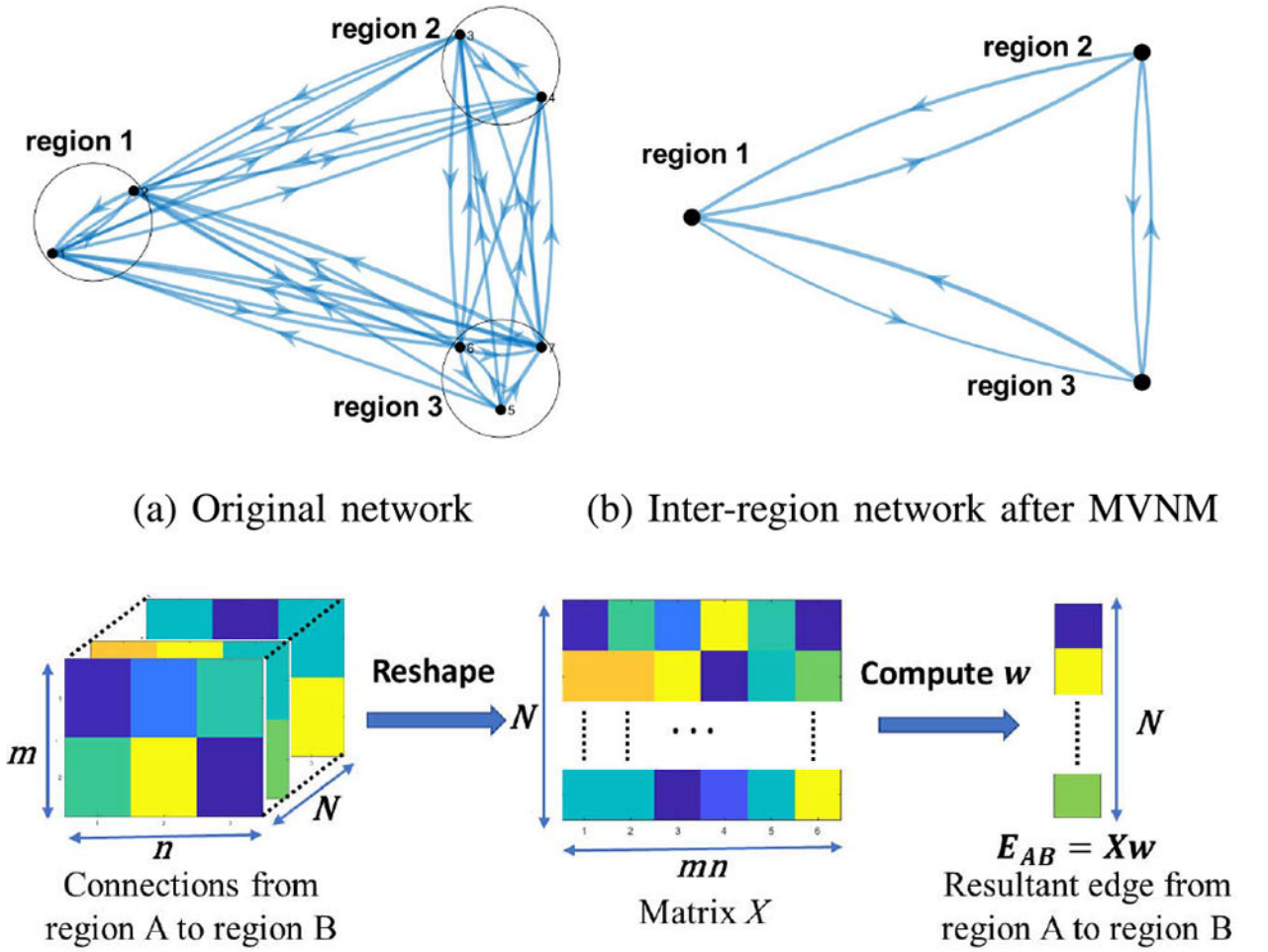
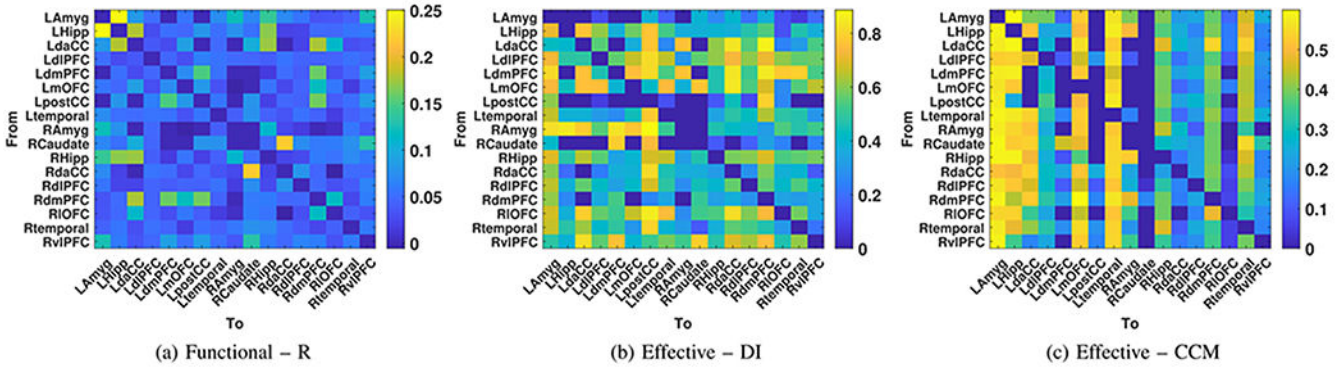


Fig. 3.

The MVNM algorithm and graph visualization of a sample causal network before and after MVNM. The causal network has three regions with two channels (each) in regions 1 and 2, and three channels in region 3.



(a) Functional – R (b) Effective – DI (c) Effective – CCM

Fig. 4.

Functional (correlative) and effective (causal) networks of subject-1 constructed from a randomly chosen task segment.

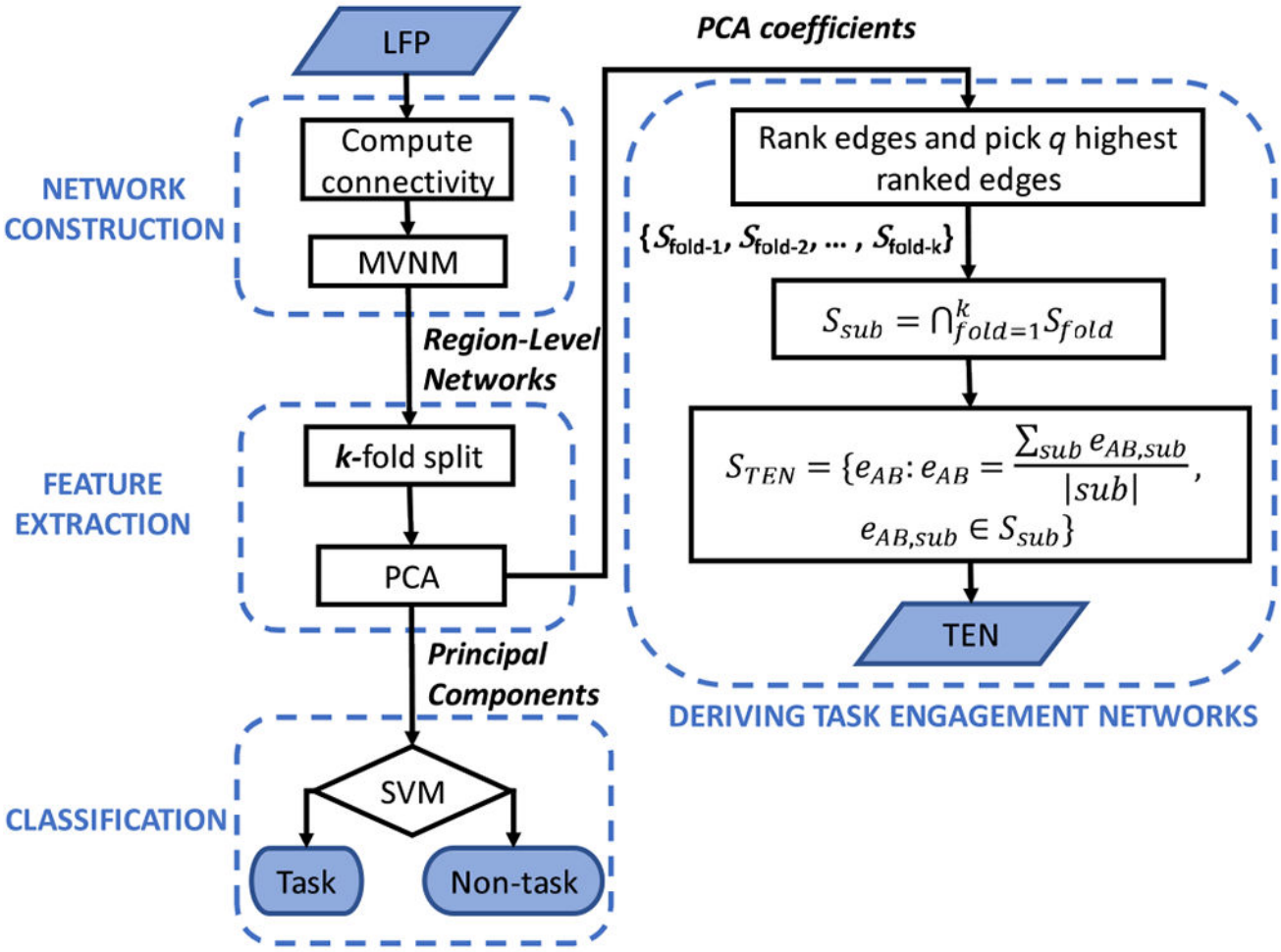


Fig. 5.
Flowchart showing the key steps of ‘task’ vs. ‘non-task’ classification process and determination of task engagement networks.

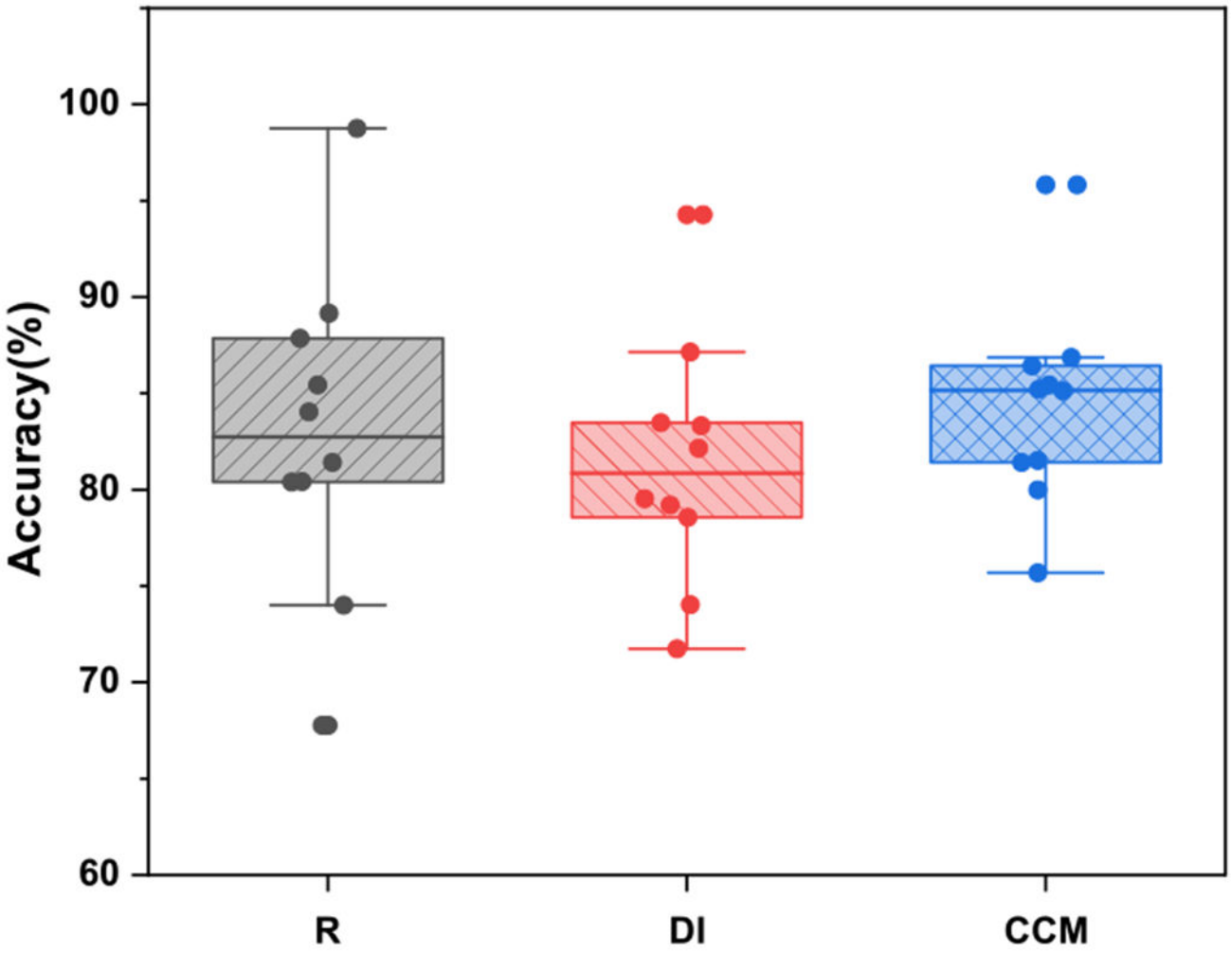
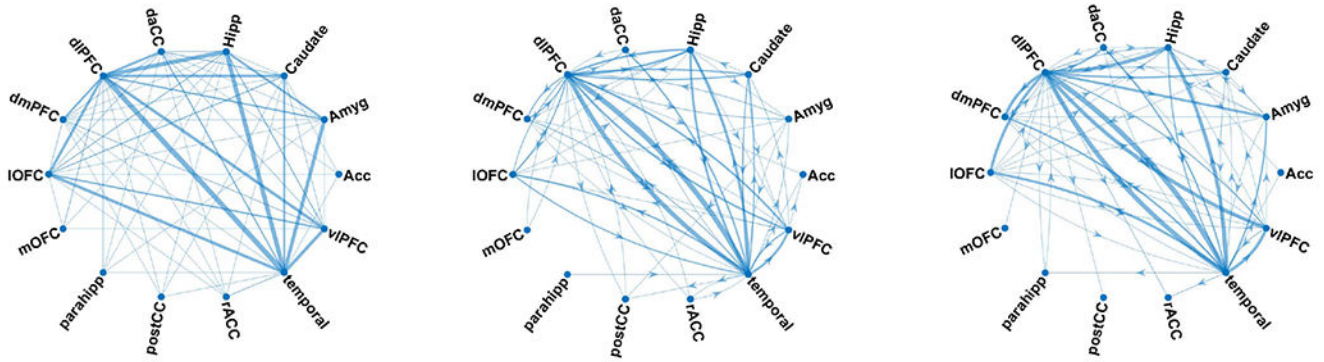


Fig. 6.

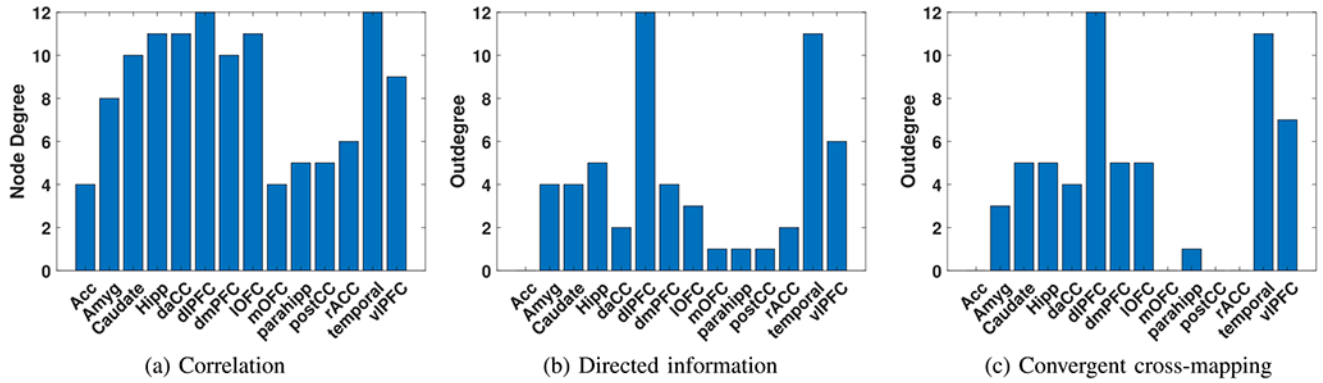
Task vs. non-task classification accuracy using networks constructed using three connectivity measures: correlation (R), directed information (DI), convergent cross-mapping (CCM). Each point within the boxplots represents 10-fold cross-validation accuracy of a participant.



(a) Correlation (b) Directed information (c) Convergent cross-mapping

Fig. 7.

Task engagement networks generated from the three network construction methods. Each node represents one of the 14 regions of interest (Acc: accumbens, Amyg: amygdala, caudate, Hipp: hippocampus, dACC: dorsal anterior cingulate cortex, dlPFC: dorsolateral prefrontal cortex, dlPFC: dorsomedial prefrontal cortex, lOFC: lateral orbitofrontal cortex, mOFC: medial orbitofrontal cortex, parahipp: parahippocampus, postCC: posterior cingulate cortex, rACC: rostal anterior cingulate cortex, temporal lobe, vlPFC: ventral lateral prefrontal cortex). The thickness of the edges represent edge strength, as described by (2).



(a) Correlation (b) Directed information (c) Convergent cross-mapping

Fig. 8.

Node centrality of each region in the task engagement networks. The bar plots represent node degree for undirected (R) networks and outdegree for directed networks (DI and CCM).

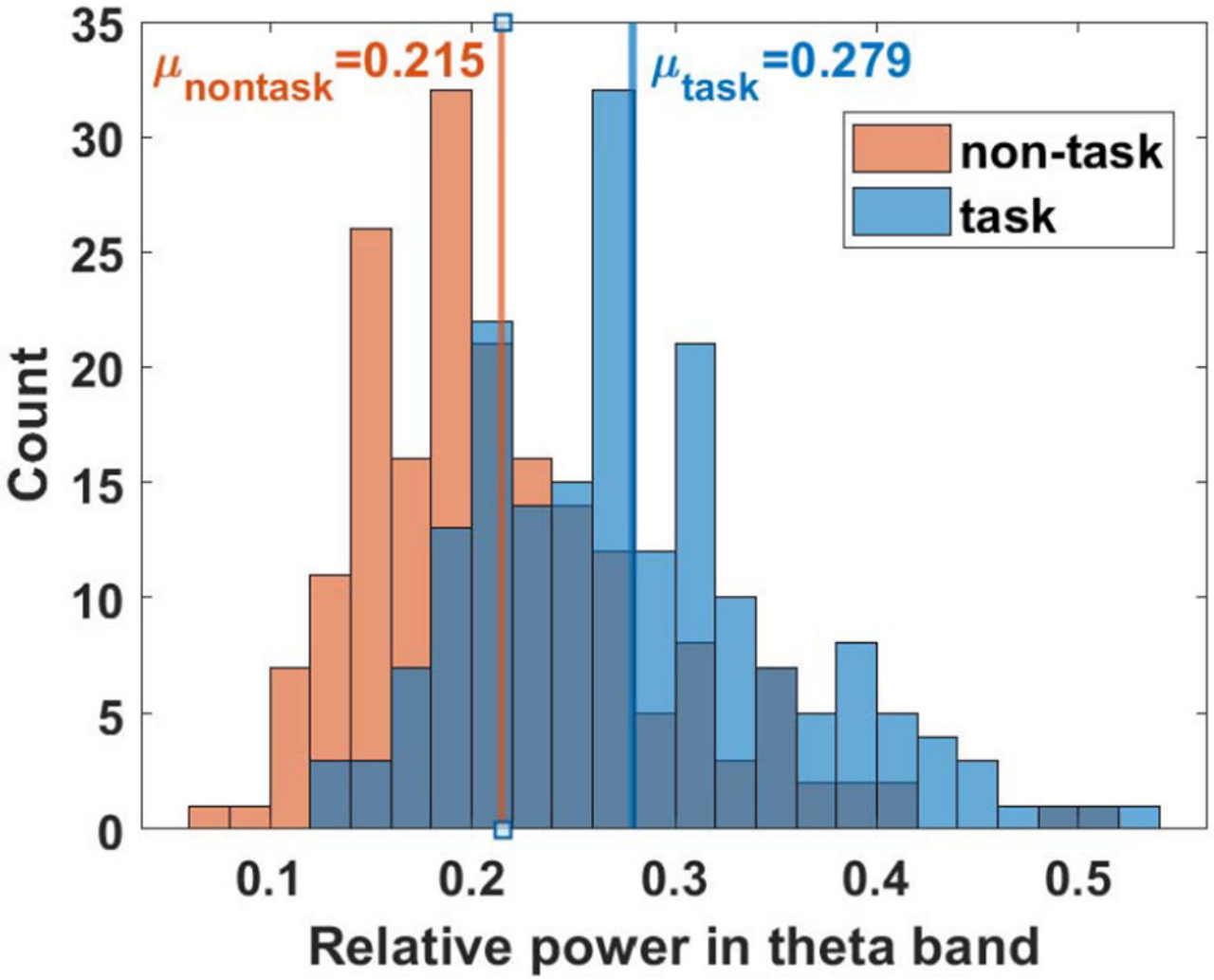


Fig. 9.

Histogram of relative theta band power in left dorsolateral PFC of subject-2 for task and non-task periods.

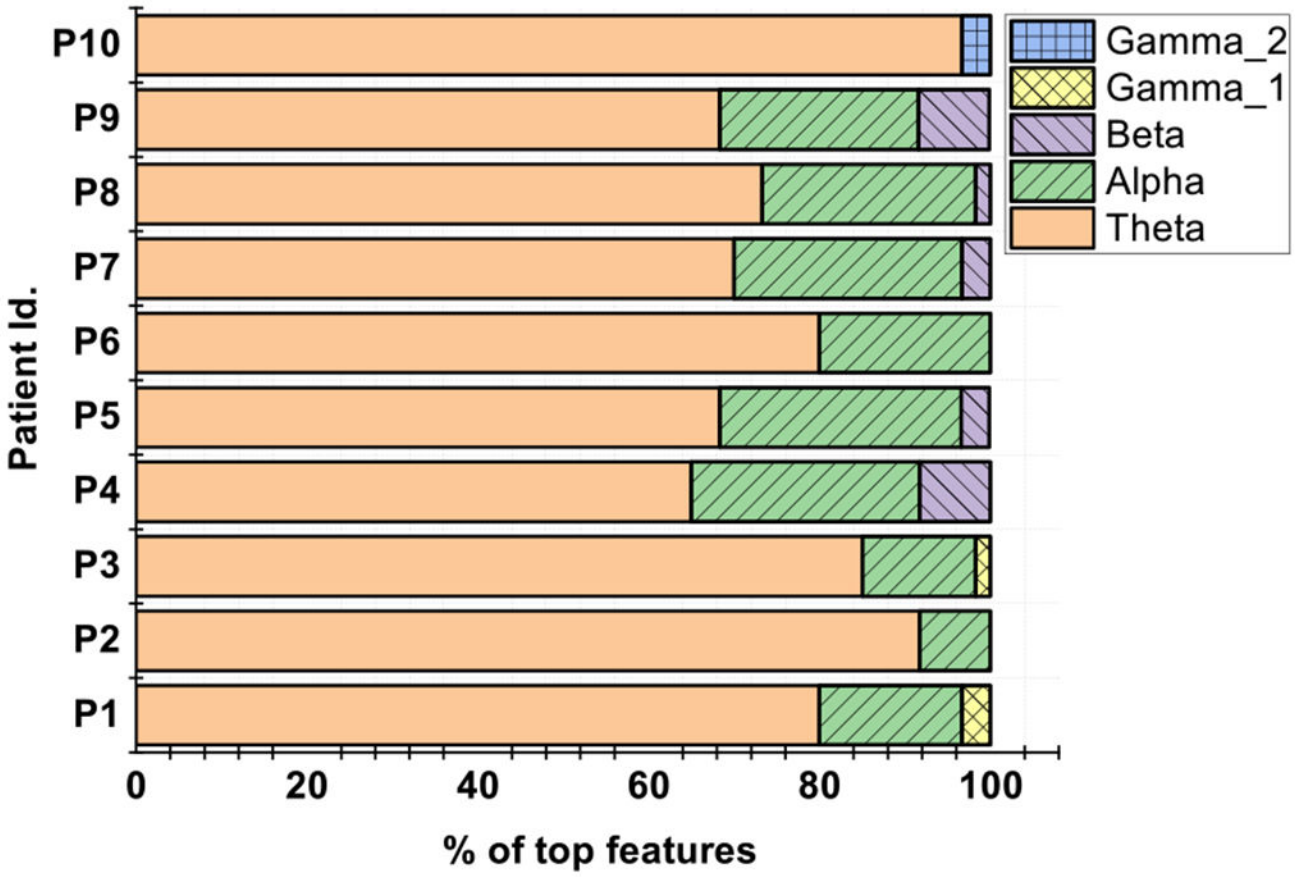


Fig. 10.
Proportion of optimal features from each subband.

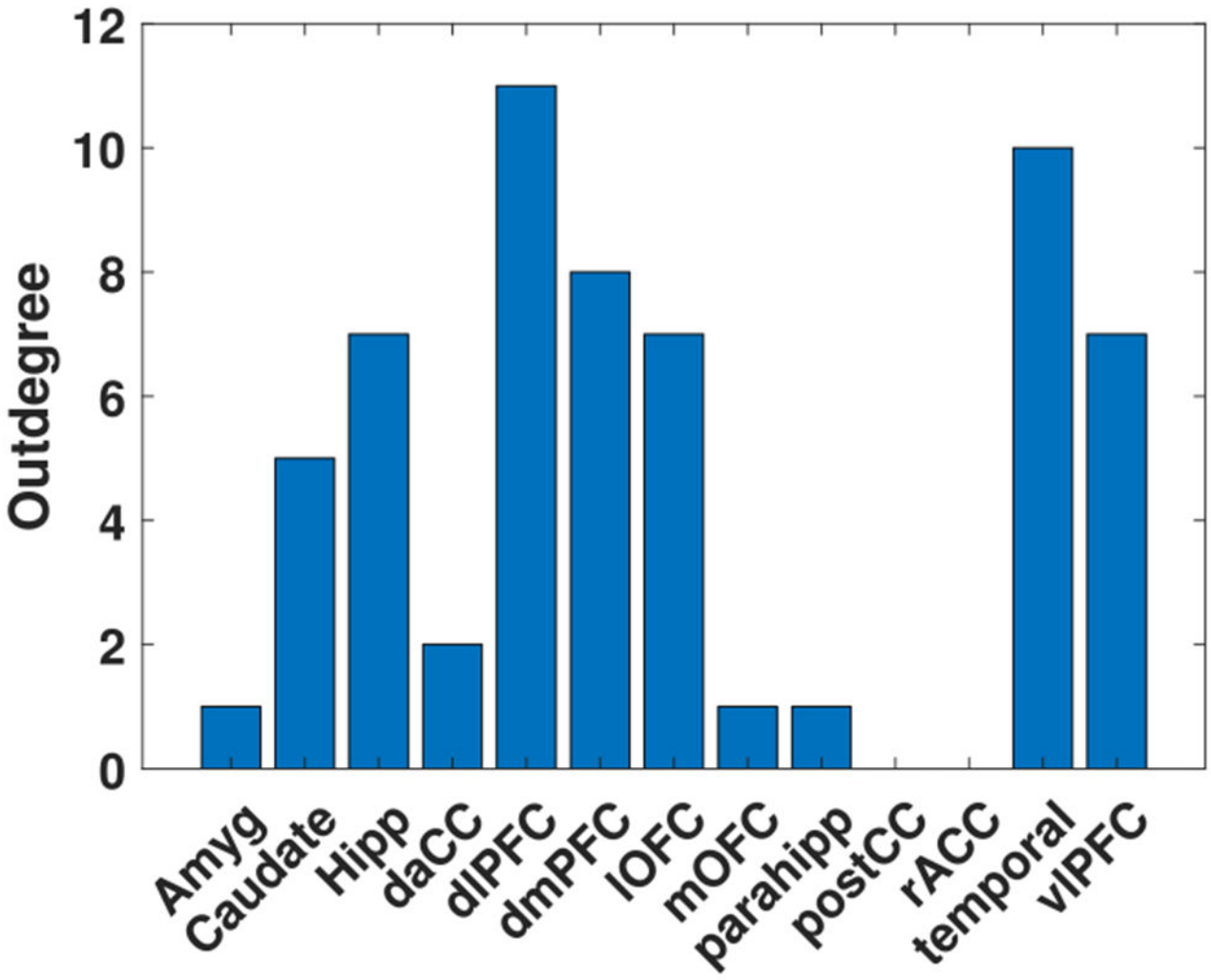


Fig. 11.
Outdegree of the regions in theta band TEN.

TABLE I

SUMMARY OF TASK VS. NON-TASK CLASSIFICATION RESULTS. THE CLASSIFIERS ARE SUBJECT SPECIFIC. THE TABLE PRESENTS MEDIAN AND INTERQUARTILE RANGE VALUES OF CLASSIFICATION ACCURACY, SENSITIVITY (TASK ACCURACY) AND SPECIFICITY (NON-TASK ACCURACY) FOR EACH OF THE THREE NETWORK CONSTRUCTION METHODS. RANDOM LABEL-ASSIGNMENT WOULD RESULT IN A BASELINE ACCURACY OF 50%. THE HIGHEST ACCURACY IN EACH ROW IS PRESENTED IN BOLD

Network	**Functional**		**Effective**	
Method	**FCHA [19]**	**R+MVNM**	**DI+MVNM**	**CCM+MVNM**
Acc.	78.1 ± 7.39	82.74 ± 7.45	80.85 ± 4.9	**85.17 ± 5.0**
Sens.	71.0 ± 10.3	84.86 ± 8.57	82.58 ± 8.57	**87.49 ± 8.77**
Spec.	79.2 ± 7.7	**83.7 ± 9.31**	79.74 ± 6.38	82.12 ± 13.2