Directional statistical graphical modeling of phase-based connectivity

Issey Sukeda[1,2,*], Takeru Matsuda[1,2],

1 Department of Mathematical Informatics, Graduate school of Information Science and Technology, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo, Japan
2 RIKEN Center for Brain Science, RIKEN, 2-1, Hirosawa, Wako, Saitama, Japan

* sukeda-issei006@g.ecc.u-tokyo.ac.jp

Abstract

Identifying phase coupling recorded from multiple electrodes during electrophysiological diagnostics, such as electroencephalogram (EEG) and electrocorticography (ECoG), helps neuroscientists and clinicians understand the underlying brain structures or mechanisms. From a statistical perspective, these signals are multi-dimensional circular measurements that are correlated with one another and can be effectively modeled using a torus graph model designed for circular random variables. Using the torus graph model avoids the issue of detecting pseudo correlations. However, the naive estimation of this model tends to lead to a dense network structure, which is difficult to interpret. Therefore, to enhance the interpretability of the brain network structure, we propose to induce a sparse solution by implementing a regularized score matching estimation for the torus graph model based on the information criteria. In numerical simulations, our method successfully recovered the true dependence structure of the brain, from a synthetic dataset sampled from a pre-given torus graph model distribution. Furthermore, we present analyses of two real datasets, one involving human EEG and the other marmoset ECoG, demonstrating that our method can be widely applied to phase-coupling analysis across different types of neural data. Using our proposed method, the modularity of the estimated network structure revealed more resolved brain structures and demonstrated differences in trends among individuals.

Author summary

Current EEG and ECoG analyses are limited because relatively simplistic statistical methods, such as linear correlation or coherence measures, are often used, which fail to capture the full complexity of phase relationships. In this study we attempted to use a superior data analysis technique to capture the multidimensional dependence structure in brain that can impede clinical interpretations from participant data. Specifically, we utilized a modified torus graph model based on directional statistics. With this approach, certain brain structures could be resolved into a sparse graphical network with fewer edges, implying modular and small-world characteristics.

Introduction

The phase analysis of brain waves, as measured by electroencephalography (EEG) or electrocorticogram (ECoG), has gained significant attention in neuroscience research

due to its potential to reveal intricate dynamics of brain connectivity and function. Such phase-based approaches can uncover neural mechanisms that are not readily visible through conventional power-based analyses, which focus primarily on the amplitude of oscillatory activity [1]. In some literature, phase-based connectivity is thought to reveal the temporal information within neural populations, whereas power-based connectivity is thought to reveal the spatial information of them [2].

Phase synchronization and phase locking between different brain regions have been widely studied as they are thought to underlie communication between neural networks [2–6]. Thatcher et al. [7] examined the relationship between EEG phases and intelligence, and revealed that the increased neural efficiency and brain complexity are positively related to IQ scores. Busch et al. [8] investigated the relationship between the phase oscillations and our visual perceptions, showing that oscillatory phase at stimulus onset can predict visual detection performance. Moreover, they have also become important from the perspective of research on brain diseases. For example, the early gamma band phase locking in schizophrenia [9] and the phase synchronization of epilepsy participants [10] have been reported. In addition, various metrics have been developed to quantitatively measure phase relationships, such as the phase locking value [11] and the phase lag index [12]. However, many of the existing studies employ relatively simplistic statistical methods, such as linear correlation or coherence measures, which often fail to account for the full complexity of phase relationships.

The analysis of phase data is inherently circular in nature, meaning that the data reside on a unit circle. To handle such circular data effectively, specialized statistical methods are required. Here, we introduce the directional statistics, a branch of statistics specifically designed to analyze angular or circular data, to offer a robust framework for analyzing phase relationships (e.g., [13–15]). Directional statistics involves the application of circular distributions such as the von Mises distribution and techniques such as circular correlation and angular regression, which are well-suited to analyzing the angular characteristics of phase data. Despite the suitability of these methods, they have not been widely applied in EEG phase analysis, except for Klein et al. [16], who applied the torus graph, a high-dimensional circular model, to analyze local field potentials collected from the prefrontal cortex and hippocampus of a macaque monkey. Most studies still rely on traditional methods like linear correlation, which can be inadequate for understanding the true dependencies between brain wave phases.

This study aims to address this gap by employing directional statistics to provide a more accurate and refined analysis of phase dependencies in EEG data. Our method, which involves statistical estimation of the torus graph, simultaneously handles multiple phases of EEG and naturally produces a sparse and interpretable graph network structure, without relying on arbitrary binarization or thresholding — practices that often compromise result reliability in neuroscience. The resulting graph network can be interpreted as the brain's functional connectivity. To analyze these structures, we examined key graph properties such as modularity and small-worldness. Our method is applied to both synthetic data and real-world electrophysiological data, including 61-dimensional human EEG and 96-dimensional marmoset ECoG, with a focus on modularity differences across participant groups and frequency bands. By utilizing circular statistical techniques, we hope to offer deeper insights into the phase dynamics of brain waves and their potential implications for understanding neural communication and brain function.

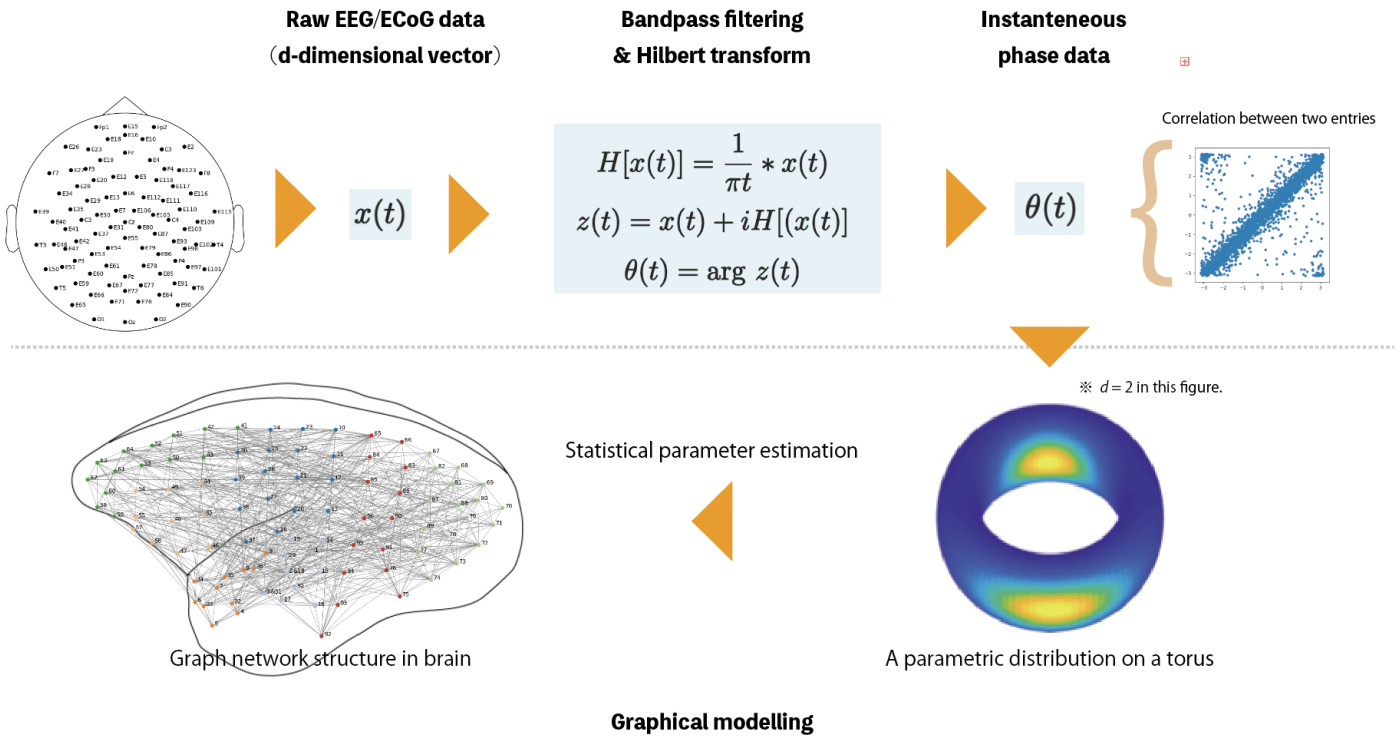


Fig 1. Overview of graphical modeling of EEG/ECoG phase data using a distribution on a torus.

Materials and Method

Data acquisition

The data used in our paper are all publicly available datasets from existing studies, and we have not collected any new data independently. Firstly, we used human EEG data with 91 measurement channels published by Chennu et al. [17] and extracted 61 channels with a length of 25000 (5 seconds at 500Hz, repeated for 10 epochs) for analysis. According to Chennu et al. [17], these data were collected from 17 participants. Among those participants, seven were labeled as a "drowsy" group and the rest as a "responsive" group, depending on the perceptual hit rate after administration of propofol. Each participant's EEG was recorded during the following four states: baseline, mild, moderate, and recovery. We also used two marmoset ECoG datasets provided by Komatsu and Ichinohe [18, 19] with all 96 measurement channels. The marmoset IDs are R03_0024_CM622M and R03_0035_CM814F. The types of conditions in these data were as follows: AD-Awake, AD-K10, AD-K30, and AD-K150, where K10, K30, and K150 mean 10, 30, and 150 minutes after ketamine administrations, respectively. We extracted each data such that its length is 25000 starting from the beginning. According to Komatsu and Ichinohe [19], these data were collected from two marmosets (one from the left hemisphere and another from the right hemisphere of their brains).

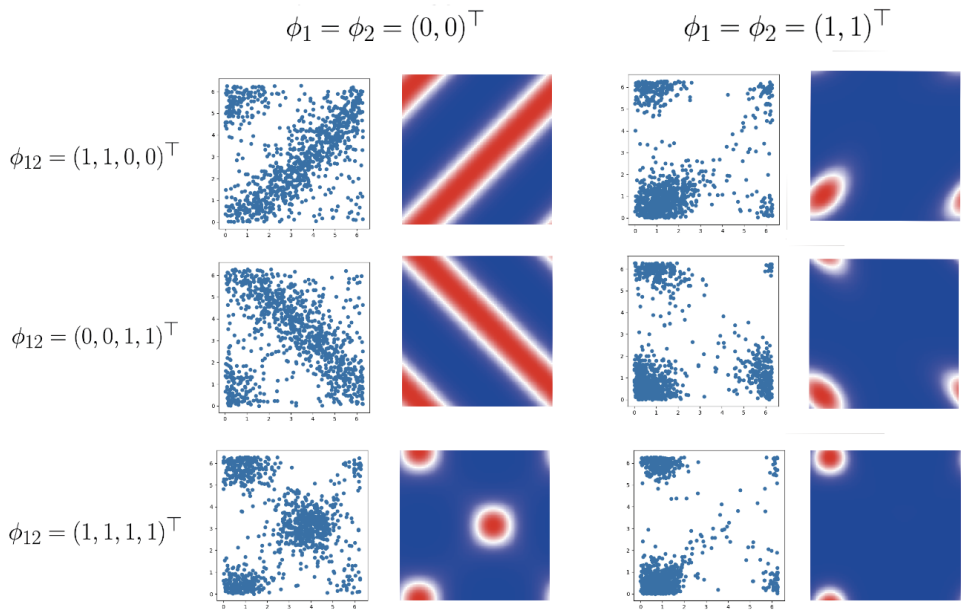


Fig 2. Submodels of torus graph model. 1000 samples and the density function are plotted for each submodel of the torus graph model.

Torus graph model for high-dimensional circular data

The overall methodological process for our data analysis is illustrated in Fig. 1. Previously, Klein et al. [16] proposed the torus graph model[1]. Suppose X is a d-dimensional circular random vector. The torus graph model is defined by the following exponential family parametrized by ϕ:

$$p(x;\phi) \propto \exp\left(\sum_{i=1}^{d} \phi_i^\top \begin{pmatrix}\cos x_i \\ \sin x_i\end{pmatrix} + \sum_{j<k} \phi_{jk}^\top \begin{pmatrix}\cos(x_j - x_k) \\ \sin(x_j - x_k) \\ \cos(x_j + x_k) \\ \sin(x_j + x_k)\end{pmatrix}\right) \tag{1}$$

where $\phi = [\{\phi_i\}_i, \{\phi_{jk}\}_{(j,k)}]$. We refer to this as the full torus graph model, in contrast to the submodels including rotational model where $\phi_{jk} = (0, 0, \cdot\,, \cdot\,)^\top$ and reflectional model where $\phi_{jk} = (\,\cdot\,, \cdot\,, 0, 0)^\top$ (see Fig. 2).

This model is a parametric non-normalized model, and its parameters can be easily estimated using score matching since it belongs to an exponential family [20]. The model incorporates the interactions between different channels through pairwise terms, making it naturally adaptable to graphical modeling. Although the originally proposed torus graph model is quite flexible and applicable when applied to high-dimensional circular data, an estimated result by naive score matching yields a dense graph, which makes it challenging to directly interpret the estimated functional connectivity between different brain regions.

To understand which parts of the brain are relatively active or which connections are stronger than others, it is preferable to obtain a sparse graphical structure. In practice,

[1]Note that the term "model" hereafter refers to the statistical model that expresses the data generation probability distribution, which is in contrast to mathematical models, which are typically described using differential equations.

a common approach to obtaining a sparse graphical structures is to apply a threshold — the connections stronger than the threshold are retained, whereas weaker ones are discarded. However, this method introduces an arbitrary error into the final result, as the outcome is highly sensitive to where the choice of threshold levels are set. Such sensitivity can significantly affect the final result, thereby reducing the reliability of the overall estimation. To address this issue, Klein et al. [16] performed statistical testing on each edge individually based on the dense estimator to determine which edges should be retained. Bonferroni correction analysis was performed to mitigate the effects of multiple testing.

Regularized score matching method for the torus graph model

While these aforementioned approaches rely on dense estimators, another direct approach is to allow the model to induce a sparse estimator, which includes many zero entries. Regularization is the typical method used to achieve this in machine learning. Here we propose to incorporate the regularization to the torus graph model [16].

In the torus graph model, $\phi_{jk} = (0, 0, 0, 0)^\top$ is equivalent to the conditional independence of x_j and x_k; x_j and x_k are independent given all other variables (Corollary 2.1-2 of Klein et al. [16]). Hence, this model can be used for graphical modeling. Suppose $G = G(V, E), V = \{1, \cdots, d\}$. The natural parameters in the torus graph model (1) can be interpreted as follows: ϕ_i is a two-dimensional vector corresponding to the node i, whereas ϕ_{jk} is a four-dimensional vector representing the interaction between the node j and k. Then, the problem of identifying the dependence between different nodes reduces to estimating the parameters ϕ statistically. Note that the normalization constant of (1) is intractable, the usual maximum likelihood estimation is also intractable. In such cases, the score matching estimation [20] is known to be useful. Specifically, the score matching objective of the exponential family takes the quadratic form [20]. For the torus graph model, the objective function is known to be

$$J(\phi) = E_{x\sim p(x;\phi)}[\frac{1}{2}\phi^\top \Gamma(x)\phi - \phi^\top H(x)]$$

where

$$\begin{cases} S_j^1(x) & = [\cos x_j, \sin x_j]^\top \\ S_{jk}^2 & = [\cos(x_j - x_k), \sin(x_j - x_k), \cos(x_j + x_k), \sin(x_j + x_k)]^\top \end{cases}$$

$$S(x) = [S^1(x), S^2(x)]^\top,$$

$$H(x) = [S^1(x), 2S^2(x)]^\top \in \mathbb{R}^{2d^2},$$

and

$$\Gamma(x) = \{\nabla_x S(x)\}\{\nabla_x S(x)\}^\top \in \mathbb{R}^{2d^2\times 2d^2}.$$

On the basis of above formula, the empirical objective function using N observed samples becomes

$$\tilde{J}(\phi) = \frac{1}{2}\phi^\top \hat{\Gamma}\phi - \phi^\top \hat{H}$$

where

$$\hat{\Gamma} = \frac{1}{N}\sum_{i=1}^{N}\Gamma(x^{(i)}),\ \hat{H} = \frac{1}{N}\sum_{i=1}^{N}H(x^{(i)}).$$

As a result, the score matching estimator of ϕ, which is the minimizer of $\tilde{J}(\phi)$, is easily calculated by using the formula below:

$$\hat{\phi}_{SM} = \hat{\Gamma}^{-1}\hat{H}. \tag{2}$$

See Klein et al. [16] for the explicit expressions for $\hat{\Gamma}$ and $\hat{H}$.

The approach for determining the graph structure via naive score matching of the torus graph model entails the issue of arbitrariness in binarization. Since the full torus graph (resp. rotational model) assigns four (resp. two) estimated values for each of the $d(d-1)/2$ edges, this approach tends to complicate the interpretation of the estimated results. Therefore, the binarization process is conducted to recover the underlying graph structure and make the results easier to interpret. The easiest method is to calculate the norm of estimated vectors corresponding to each edge, and then pull an edge if only the norm surpasses a certain threshold. However, there remains an arbitrary choice of determining an exact threshold value. To avoid this issue, Klein et al. [16] conducted hypothesis testing that examines the null hypothesis $\phi_{jk} = 0$ for each (j, k) to determine the dependence of each edge and recover an interpretable graph structure. Although this approach is statistically acceptable, the number of hypothetical tests grows exponentially under general settings. This issue was not a crucial problem in their work as they used the pre-given community structure of the network from the well-known domain knowledge, however, such prior knowledge is not always available.

To address this issue, the l_1-regularization methods can be applied, especially for higher dimensional problems, leading to a sparse solution.[2] Similar approaches for the general exponential families were studied by Lin et al. [21]. Following these studies, we propose the use of l_1-regularization for estimating the torus graph parameters by the score matching. This estimator is defined as the minimizer of

$$\tilde{J}_\lambda(\phi) = \tilde{J}(\phi) + \lambda\|\phi\|_1, \tag{3}$$

where λ is a tuning parameter which controls the sparsity of the resulting solution. Since this objective function being minimized is convex, we may utilize normal optimization methods, such as gradient descent. However, the solution may not remain sparse. Therefore, to address this, a LASSO algorithm [22] such as LARS [23] can be implemented. Yet, due to computational efficiency, the alternating direction method of multipliers (ADMM) is a more generally accepted method [24]. Our modified objective function, which is required to be minimized with respect to x, is

$$\frac{1}{2}\phi^\top\hat{\Gamma}\phi - \phi^\top\hat{H} + \lambda\|\phi\|_1.$$

In ADMM, we instead minimize $\frac{1}{2}\phi^\top\hat{\Gamma}\phi - \phi^\top\hat{H} + \lambda\|z\|_1 + \nu^\top(\phi - z) + \frac{\mu}{2}\|\phi - z\|_1$, coupled with the dual variable z by iterating the following three steps until convergence:

$$\begin{aligned}
\phi_{k+1} &= (\mu I + \hat{\Gamma})^{-1}(\mu(z_k - u_k) + \hat{H}),\\
z_{k+1} &= S_{\lambda/\mu}(\phi_k + u_k),\\
u_{k+1} &= u_k + \phi_{k+1} - z_{k+1}.
\end{aligned}$$

In this process, I denotes the identity matrix, S denotes the soft threshold function, z_k is the auxiliary parameter, and μ is the arbitrary constant.

One of the primary motivations for the use of LASSO algorithms is to obtain a regularization path (or sometimes called the solution path), which is a sequence of solutions along with various λ values. Formally, the regularization path is denoted as $\hat{\phi}^\lambda = \{\text{argmin } \tilde{J}_\lambda(\phi) \mid \lambda \in \mathbb{R}_{\geq 0}\}$, or practically $\hat{\phi}^\lambda = \{\text{argmin } \tilde{J}_\lambda(\phi) \mid \lambda \in \Lambda\}$ where Λ is the set of candidate values of λ. Note that the former becomes piecewise linear, i.e., there exists the change points $0 = \lambda_0 < \lambda_1 < \cdots < \lambda_R = \infty$ and $\xi_0, \ldots, \xi_{R-1}$ such that $\phi^\lambda = \phi^{\lambda_r} + (\lambda - \lambda_r)\xi_r$ for $\lambda \in [\lambda_r, \lambda_{r+1}]$, but the latter may not.

[2] Klein et al. [16] has also mentioned adding a group l_1 penalty in their paper, however practices were not presented.

In practice, we consider the variant problem called Group LASSO [25] instead, where the regularization terms are group-wise. In this case, the estimator becomes the minimizer of

$$\tilde{J}_\lambda(\phi) = \tilde{J}(\phi) + \lambda \sum_{j<k} c_{jk} \|\phi_{jk}\|_2$$

where c_{jk} is a weight constant assigned to each edge between nodes j and k. For convenience, we set $c_{jk} = 1$ for any corresponding edge (j, k). For the optimization process in Group LASSO, the ADMM algorithm requires a slight modification so that the update of z becomes group-wise:

$$z^g_{k+1} = R_{\lambda/\mu}(\phi^g_k + u^g_k)$$

where g indicates the partition of parameters into groups, i.e., in our case two dimensional vectors for each node and four dimensional vectors for each edge, and R is a shrinkage operator. See S1 Appendix for details.

Model selection with score matching information criteria (SMIC)

Our method for recovering the graph structure generally follows the approach outlined in Section 8.2 of Matsuda et al. [26], comparing the truncated Gaussian graphical model and the log-Gaussian graphical model for the RNAseq dataset. In our case, since we have multidimensional phase (circular) data, first, we solve the l_1-regularized score matching problem, described in previous section, for various λs. Here we recommend to set λ in 100 logarithmic steps from 10^{-3} to $10^{0.5}$, depending on the computer specification for parallel computing. Given that we obtain a sparse solution due to our modified LASSO algorithm, specifying one regularization parameter λ corresponds to deciding the edges of the recovered graph structure, resulting in one of the candidate models. For each of these models, we recalculate the score matching estimator, this time without the regularization term, and then calculate the information criteria specifically proposed for the score matching estimator. Finally, the estimated graph structure is obtained by selecting the best model with the minimum value of the information criteria. [3]

Although the l_1-regularization is a useful method for edge selection in graphical modeling (or variable selection in regression as well), the tuning parameter λ is given arbitrarily. To choose an appropriate λ, cross validation is widely used but often time consuming, especially when using leave-one-out cross validation. An alternative method is to consider the information criteria, such as the Akaike information criteria (AIC) for the maximum likelihood estimator, and select the model with the smallest AIC. Recently, the score matching information criteria (SMIC) [26] was also proposed for the score matching estimator. Model selection via SMIC minimization performs as well as the cross validation, yet with less computational time taken to complete the analysis.

In our case, the SMIC can be used to decide the regularization parameter λ, or equivalently the appropriate graph structure. The torus graph model belongs to an exponential family, hence the SMIC is calculated as

$$\text{SMIC} = 2N\tilde{J}(\hat{\phi}_{SM}) + 2\text{tr}(\hat{I}\hat{J}^{-1}), \tag{4}$$

$$\hat{I} = \frac{1}{N}\sum_{i=1}^{N}(\Gamma(x^{(i)})\hat{\phi}_{SM} - H(x^{(i)}))(\Gamma(x^{(i)})\hat{\phi}_{SM} - H(x^{(i)}))^\top,\ \hat{J} = \frac{1}{N}\sum_{i=1}^{N}\Gamma(x^{(i)}).$$

[3] The model with the minimum information criteria is generally considered to fit the data best while suppressing the model complexity. We do not argue the validity of the information criteria itself in this paper.

Briefly, the first term $N\tilde{J}(\hat{\phi}_{SM})$ has the negative bias of $\mathrm{tr}(I(\phi)J(\phi)^{-1})$ as the estimator of the predictive distribution $J(\phi)$; thus, the bias correction is conducted by adding the second term. To perform model selection via SMIC minimization, we should calculate the value of SMIC for all candidate models and finally select the one with minimum SMIC value. However, since λ can vary continuously, we discretize it and select a limited number of points in practice. To effectively identify the best λ, we first pick 10 equally spaced discrete points and select an interval that includes the minimum SMIC value. We then repeat the same calculation for this interval recursively. In each recursive iteration, parallel computing can be applied since the calculation of SMIC is independent for each choice of λ, which further accelerates the entire procedure.

Graphical properties of the estimated network structure

Once the appropriate graphical model is specified, the representative graphical properties can be observed. Indeed, we obtained the following four metrics: (1) number of edges (2) average clustering coefficient (3) average shortest path length, and (4) modularity.

Average clustering [27] measures how often the neighbors of a node are connected to each other. For an undirected, unweighted graph $G = (V, E)$, let $m(v) = |\{(u, w) \in E \mid (v, w) \in E \text{ and } (v, u) \in E\}|$ and $t(v) = \frac{d(d-1)}{2}$ where d is the degree of node v. Then, the average clustering coefficient is defined as

$$C = \frac{1}{|V'|} \sum_{v \in V'} \frac{m(v)}{t(v)}$$

where V' is the set of nodes v with $d \geq 2$. As the value of C ($\in [0, 1]$) becomes larger, the more densely connected the network becomes.

Average shortest path length measures the length of paths between any two nodes in the network. For a connected undirected unweighted graph $G = (V, E)$,

$$L = \sum_{s,t \in V, s \neq t} \frac{D(s, t)}{|V|(|V| - 1)}$$

where $D(s, t)$ denotes the shortest path length between node s and t. The smaller L (≥ 1) is, the more densely connected the network is.

In addition, community-detection-based measures are also considered useful. Newman and Girvan [28] introduced the modularity maximization problem for community detection. For an undirected, unweighted graph $G = (V, E), |E| = m$, modularity is defined as

$$\max_{\mathcal{C}} \sum_{C \in \mathcal{C}} \left(\frac{m_C}{m} - \left(\frac{D_C}{2m} \right)^2 \right)$$

where $\mathcal{C}$ is a clustering (a set of mutually exclusive sets of V), C denotes each cluster, m_C denotes the number of edges in C, D_C denotes the sum of degree of nodes in C. The larger the modularity is, the denser the connections within each cluster become, and the sparser the connections become between the different clusters in the network. To overcome the drawbacks of the modularity [29, 30], Li et al. [31] proposed the modularity density maximization problem. Other groups have tested this method and have developed mathematical approaches to improve the computation burden and heuristic methods [32, 33]. Among them, calculating modularity by the Louvain method [34] is mostly useful in practice because it operates in time complexity $O(|V| \log |V|)$ [35, 36]; hence, we utilize it in the following experiments.

Results and Discussion

Regularized score matching of synthetic data

To validate our method, we performed it on synthetic data containing 5, 19, and 61 dimensions. The dimensions were chosen arbitrarily to represent different levels of granularity: a coarse, a moderate, and a finer resolution for analysis. In doing so, we generated an artificial dataset from a known torus graph model and attempted to recover a true graph structure. We then sampled from a five-dimensional torus graph using rejection sampling to generate 1000 data points. Firstly, 5 dimensional case is examined. The true graph structure here is set to $G = G(V, E), V = \{1, 2, 3, 4, 5\}, E = \{(1, 3), (1, 4), (2, 4), (2, 5), (3, 5)\}$. Moreover, the true natural parameter ϕ of the torus graph model is set such that the corresponding parameters to E are all set to 0.3 and 0 otherwise for instance. Then, we applied our regularized score matching with various λ values. For each λ, we calculate SMIC, which is plotted in Fig. 3. As depicted in Fig. 4, the SMIC takes the minimum value at the model with five edges, which correctly corresponds to E. Conversely, calculating the phase locking value of each edge pair overdetects the edge, suffering from false pseudo correlations.

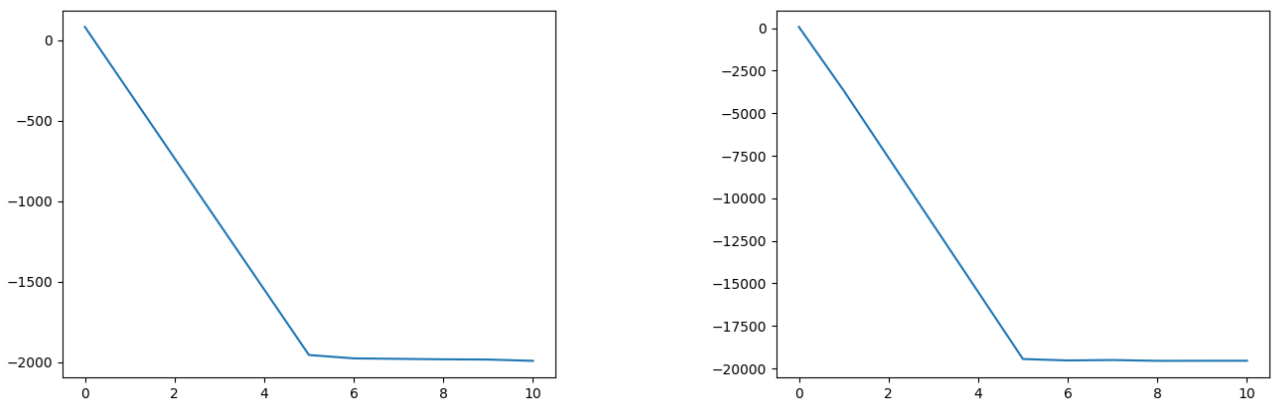


Fig 3. The recovery of graph structure with 5-dimensional synthetic data. The axes are #edges vs SMIC. The left figure shows the result with 1000 samples and right with 10000 samples.

Next, we sample from a higher dimensional torus graph model to generate artificial samples. Since the sampling from a high dimensional torus graph model is computationally prohibitive, we use Gibbs sampling[4] instead. This is possible because the univariate conditional distributions of one variable given the rest $p(x_k \mid x_{-k}; \boldsymbol{\phi}), k = 1, 2, \ldots, d$ are von Mises distribution (see Theorem 2.2 of Klein et al. [16] for the formulation). In this setting, the true model parameter is set as follows. First, we generate a random graph structure named Erdös-Rényi model $G(V, E) = G(|V|, p)$. In this model, each edge is included in the graph with probability $p = \frac{1}{2}$. Then we assign

$$\begin{cases} \phi_i = [0.0, 0.0] & (i \in V) \\ \phi_{jk} = [0.3, 0.3, 0.0, 0.0] & ((j, k) \in E) \\ \phi_{jk} = [0.0, 0.0, 0.0, 0.0] & ((j, k) \notin E) \end{cases}$$

[4]Gibbs Sampling is a popular Markov Chain Monte Carlo (MCMC) algorithm used for generating samples from a joint probability distribution.

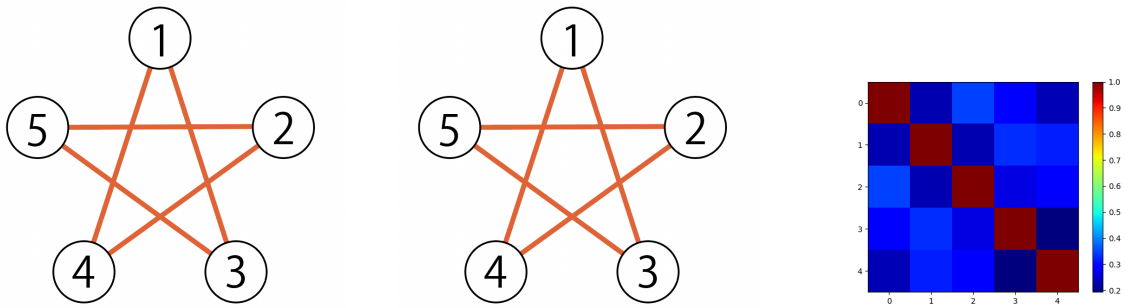


Fig 4. Results on 5 dimensional synthetic data sampled from a predefined torus graph model. The first image shows the ground truth structure. The second image shows the recovered structure by our method. The third image shows the adjacency matrix representing the phase locking values between nodes. While the naive phase locking value calculation overdetects the edges, our proposed method correctly recovers the dependence structure.

d	# true edges	# estimated edges	TP	FP	TN	FN	best λ
19	80	122	80	42	0	49	0.034
61	889	1263	889	374	0	567	1.616

Table 1. Detection probability for high-dimensional synthetic dataset from a random graph structure. Example results for artificially generated dataset sampled from a torus graph where the graph structures are randomly determined by the Erdös-Rényi random graph.

as the ground truth. Note that since this distribution is not multimodal, the simple Gibbs sampling works well. The burn-in period[5] was set to 10000.

As a result, we found that our method based on SMIC minimization tends to select a denser graph structure than the true one. Tab. 1 and Fig. 5 show the results for $d = 19$ and $d = 61$. This tendency is consistent with the precedent results, where the model selection methods based on AIC-alike information criteria tend to increase the false positive rates. Contrarily, the decline in SMIC drastically changes around the true number of edges: 80 edges in 19 dimensional example and around 889 edges in 61 dimensional example.

Comparison with cross-validation and information criteria

We further compare the performance of SMIC minimization against the typical cross validation method to select the regularization parameter λ. Here we use the leave-one-out-cross-validation (LOOCV). The counterpart of information criteria, which we call SMCV herein, is calculated as follows:

$$\hat{\Gamma}^{(-i)} = \frac{1}{N-1} \sum_{j \in \{1,\ldots,N\} \setminus \{i\}} \Gamma(x^{(j)}),$$

$$\hat{H}^{(-i)} = \frac{1}{N-1} \sum_{j \in \{1,\ldots,N\} \setminus \{i\}} H(x^{(j)}),$$

$$\mathrm{SMCV} = \sum_{i=1}^{N} \frac{1}{2} \hat{\phi}_{SM}^{\top} \hat{\Gamma}^{(-i)} \hat{\phi}_{SM} - \hat{\phi}_{SM}^{\top} \hat{H}^{(-i)}.$$

[5]The burn-in period of Markov Chain Monte Carlo (MCMC) refers to the initial set of iterations during which the chain is allowed to warm up and approach the target distribution.

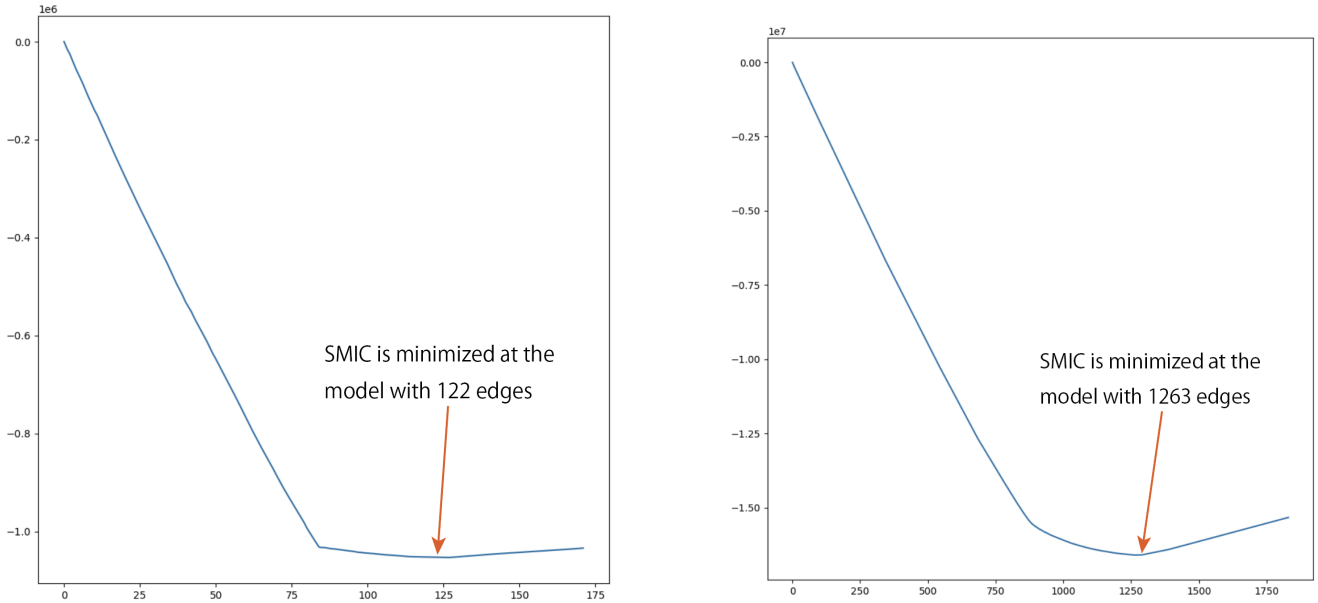


Fig 5. The number of edges vs SMIC for generated dataset sampled from a torus graph. (left: $d = 19$, right: $d = 61$)

Here, i denotes the index of the data, which is left out from the dataset one by one. To compare the performance of SMIC minimization with SMCV minimization, we conducted two simulation experiments using three- and five-dimensional synthetic data, respectively. These simulations show that minimizing SMIC exhibits performance comparable to the standard method of choosing the minimum SMCV model when measured by the edge detection probability, while achieving a significant acceleration as shown in Fig. 6.

First, we sample $N = 100$ samples from 3-dimensional torus graph using rejection sampling. Suppose the true graph structure is $G = G(V, E), V = \{1, 2, 3\}$ where $E = \{(1, 2), (2, 3)\}$ and $\phi = (0, 0, 0, 0, 0, 0, c, c, c, c, 0, 0, 0, 0, c, c, c, c)^\top, c \in \mathbb{R}$. Since G consists of 3 edges, the number of candidate models is $2^3 = 8$. We select the model with minimum SMCV or SMIC as the optimal model. Tab. 2 shows the detection probabilities of each edge in the torus graph model for $N = 100$ when $c = 0.1, 0.2, 0.3, 0.5$. As shown in Tab. 3, the similar experiment is also conducted for

$$\phi = (c, c, c, c, c, c, c, c, c, c, 0, 0, 0, 0, c, c, c, c)^\top, c \in \mathbb{R}.$$

Next, we sample $N = 100$ samples from 5-dimensional torus graph using rejection sampling. Suppose the true graph structure is $G = G(V, E)$,$V = \{1, 2, 3, 4, 5\}$, $E = \{(1, 3), (1, 4), (2, 4), (2, 5), (3, 5)\}$. The true natural parameter ϕ of the torus graph model is set such that the corresponding parameters to E are all set to 0.3 and 0 otherwise, in align with previous experiments. Since the number of full candidate models 2^{10} is too large to explore, we instead employ the group LASSO method to restrict the candidate models into less than 100, which are automatically exploited as the regularization path. The model with least SMCV or SMIC is selected as the optimal model as well. As a result, Tab. 4 shows the detection probability of each edge.

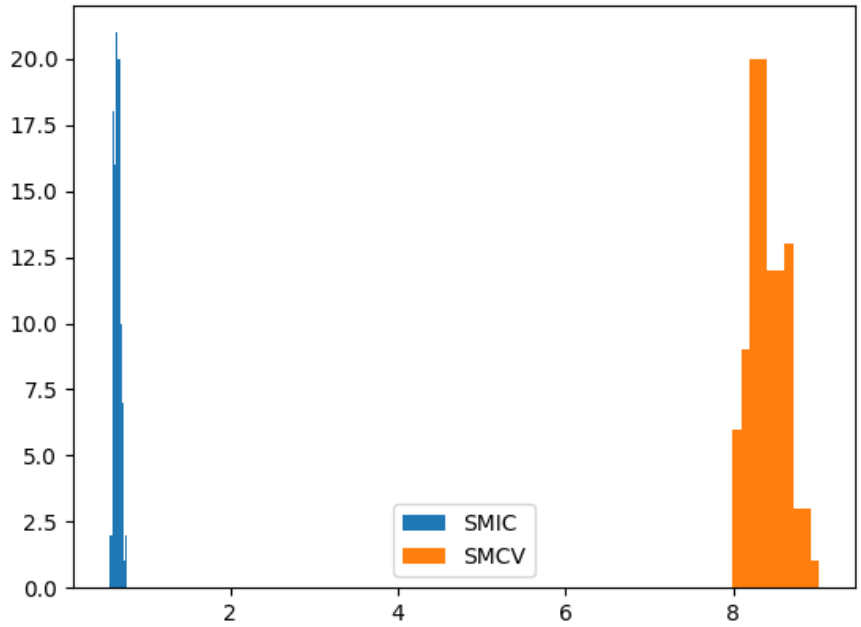


Fig 6. Average computation times in seconds of 30 runs. In average, 8.41 seconds for the SMCV calculation and 0.649 seconds for the SMIC calculation.

Real data analysis with human EEG

By using a 61-dimensional human EEG dataset (50 seconds) collected from 17 participants, with seven labeled as the "drowsy" group and the rest as the "responsive" group based on their perceptual hit rate after administration of propofol, we investigated how differences in wakefulness modulate functional connectivity across different frequency bands. To investigate the dependence structures for each individual, we applied the rotational model with regularized score matching to each frequency band. This choice is supported by the experimental results that even when the full torus graph is applied to the real EEG data, the latter half of ϕ_{jk} in Eq (1) are always near zero, implying that the rotational model without these terms is sufficient for modeling these data. Fig. 7 shows one example. This is also similar to the model selection performed in the analyses using local field potentials as described by Klein et al. [16].

The preprocessing is composed of the Hilbert transform and the bandpass filter extracting α-band (8Hz – 15Hz), β-band (12Hz – 25Hz), or γ-band (25Hz – 40Hz). These thresholds of frequency follow Chennu et al. [17]. In practice, the Hilbert transform is implemented by Python Scipy library. After the preprocessing (Hilbert transform and band pass filtering), we merge 10 epochs (0∼9) together to prepare 25000 samples. Then, we extract 61 dimensional data from it following the 10-10 system of electrodes [37]. The position of electrodes is visualized in Fig. 8. The final data results in the shape of $(N, d) = (25000, 61)$.

Fig. 9 displays one representative example from the drowsy group. The results are shown for the four participant state (baseline $\rightarrow$ mild $\rightarrow$ moderate $\rightarrow$ recovery), where the sparse solution is obtained for each state. For this participant, the number of edges varies at approximately $400 \sim 600$ edges, which is significantly smaller than the

Table 2. Detection probability. The 3-dimensional synthetic data were sampled from the torus graph model with $\phi = (0, 0, 0, 0, 0, 0, c, c, c, c, 0, 0, 0, 0, c, c, c, c)^\top, c \in \mathbb{R}$, $c = 0.1, 0.2, 0.3, 0.5$.

edge	Ground truth	SMCV	SMIC
(1,2)	true	0.426	0.430
(1,3)	false	0.201	0.203
(2,3)	true	0.436	0.437
edge	Ground truth	SMCV	SMIC
(1,2)	true	0.680	0.686
(1,3)	false	0.080	0.079
(2,3)	true	0.686	0.703
edge	Ground truth	SMCV	SMIC
(1,2)	true	0.935	0.938
(1,3)	false	0.055	0.057
(2,3)	true	0.940	0.948
edge	Ground truth	SMCV	SMIC
(1,2)	true	1.000	1.000
(1,3)	false	0.043	0.078
(2,3)	true	1.000	1.000

Table 3. Detection probability The 3-dimensional synthetic data were sampled from the torus graph model with
$\phi = (c, c, c, c, c, c, c, c, c, c, 0, 0, 0, 0, c, c, c, c)^\top, c = 0.1, 0.2, 0.3, 0.5$.

edge	Ground truth	SMCV	SMIC
(1,2)	true	0.409	0.415
(1,3)	false	0.203	0.202
(2,3)	true	0.433	0.434
edge	Ground truth	SMCV	SMIC
(1,2)	true	0.667	0.744
(1,3)	false	0.067	0.052
(2,3)	true	0.606	0.670
edge	Ground truth	SMCV	SMIC
(1,2)	true	0.843	0.950
(1,3)	false	0.041	0.060
(2,3)	true	0.834	0.952
edge	Ground truth	SMCV	SMIC
(1,2)	true	0.922	0.995
(1,3)	false	0.030	0.169
(2,3)	true	0.929	0.991

maximum edge number $61 \cdot 60/2 = 1830$. It is observed that the number of edges drops by the propofol injection between the baseline state and the mild state, while it recovers the network structure towards the recovery state gradually. To further enhance the interpretability of these networks, observing the underlying community structure of them is an useful approach. Here, the color of each node represents the optimal community structure obtained by the modularity maximization [28], where the number of communities is not specified in advance but determined through the optimization process. Overall, it is observed that the close nodes forms a cluster in general, while some clusters are dispersed among the whole brain, e.g., the gray-colored cluster in the mild state. These community structures depict the occurrence of neural oscillations

Table 4. Detection Probability. The 5-dimensional synthetic data were sampled from the torus graph model with $c = 0.1, 0.3$.

edge	Ground truth	SMCV	SMIC
(1,2)	false	0.04	0.04
(1,3)	true	0.11	0.16
(1,4)	true	0.16	0.21
(1,5)	false	0.04	0.06
(2,3)	false	0.07	0.09
(2,4)	true	0.27	0.32
(2,5)	true	0.15	0.16
(3,4)	false	0.07	0.08
(3,5)	true	0.13	0.16
(4,5)	false	0.06	0.06

edge	Ground truth	SMCV	SMIC
(1,2)	false	0.05	0.09
(1,3)	true	0.83	0.92
(1,4)	true	0.85	0.90
(1,5)	false	0.05	0.09
(2,3)	false	0.06	0.07
(2,4)	true	0.86	0.94
(2,5)	true	0.90	0.95
(3,4)	false	0.04	0.08
(3,5)	true	0.80	0.88
(4,5)	false	0.05	0.07

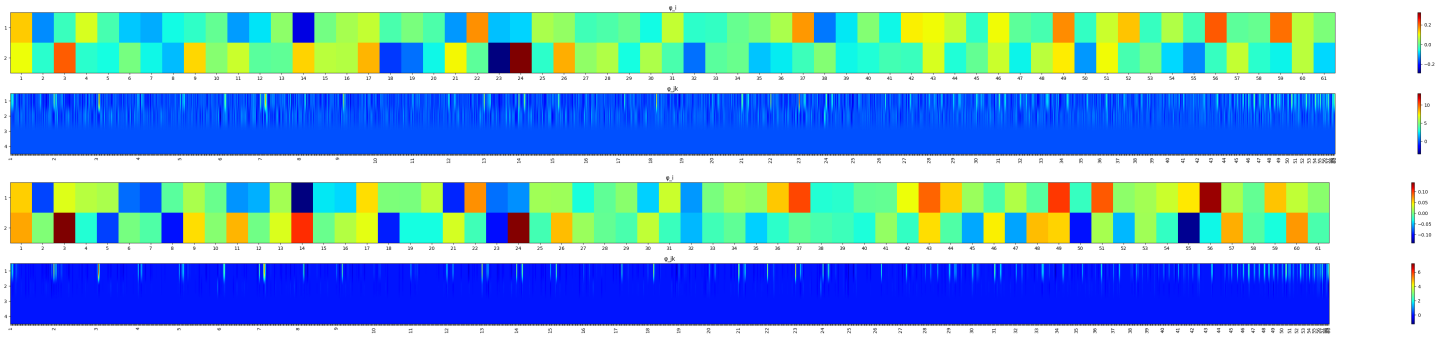


Fig 7. An example of the estimated values of $\hat{\phi}$. The upper half shows the naive score matching and the bottom half shows the regularized score matching with Group LASSO. In each figure, the first image depicts ϕ_i, 61 two-dimensional vectors, and the second image depicts ϕ_{jk}, 1830 four-dimensional vectors, column by column.

with richer information compared to the naive way of calculating the correlation between each node pairs.

For all participants, Tab. 5 and Tab. 6 summarize the changes in the number of edges, modularity, average clustering coefficient, and average shortest path length for the recovered graph structures for each band of EEGs. The shifts of the ratio between average clustering and average shortest path length during each state are shown in Fig. 10. The responsive group clearly exhibited higher values in every frequency band analyzed. A significant distinction between the drowsy group and the responsive group cannot be observed in the alpha band EEGs. Contratily, for the beta bands, the drowsy group tends to exhibit a lower number of edges, higher modularity, lower average clustering coefficient, and higher average shortest path length compared to the

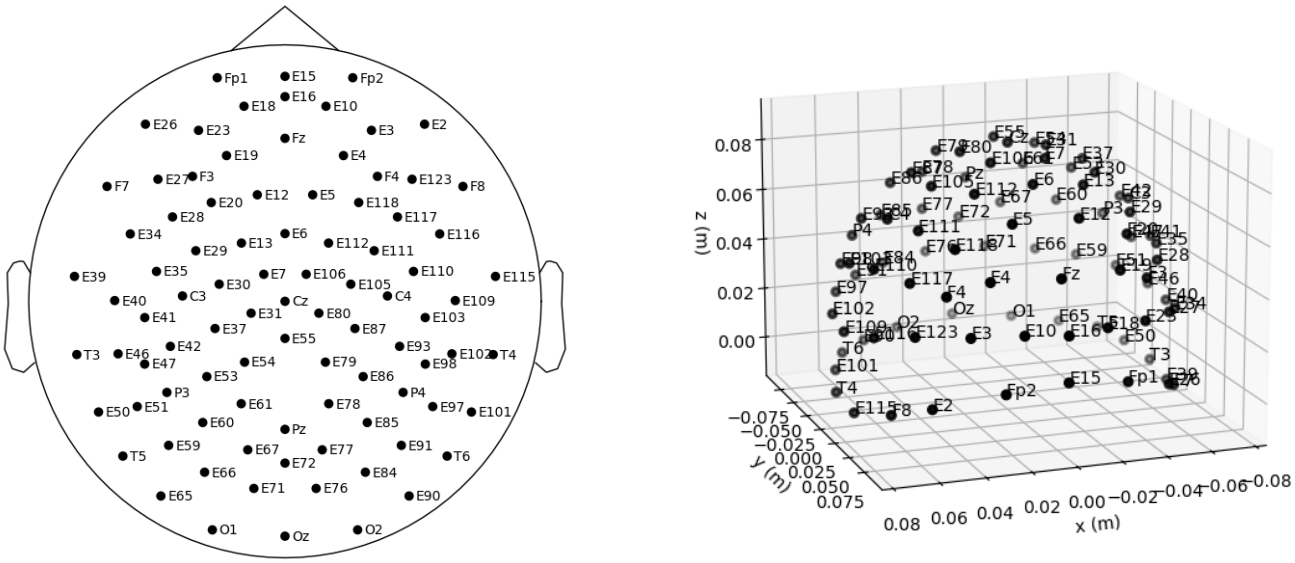


Fig 8. The positions of 91 electrodes. 61 out of 91 electrodes are used in our analysis following the 10-10 system of electrodes [37].

responsive group. Given that a small-world network [38, 39] is primarily characterized by a higher average clustering coefficient and shorter average shortest path [40], we conclude that small-worldness is more pronounced in the beta connectivity of the responsive group, while modularity is not increased in this group. This is not aligned with the previous study on 10 healthy humans [41], which reported that propofol increased both average path length, clustering coefficient, and modularity calculated from phase lag index (PLI) network. This discrepancy could be due to several factors, including differences in the methodologies used, the specific experimental conditions, or the statistical method for reconstructing the graph network structure from the EEG data.

Table 5. Summary of graphical properties (the number of edges and modularity). The average is taken among all the participants in each participant group. Refer to S4 Appendix for full individual results.

	participant group	baseline	mild	moderate	recovery
α band (8-15Hz)	drowsy	(762, 0.388)	(855, 0.353)	(749, 0.394)	(863, 0.351)
α band (8-15Hz)	responsive	(863, 0.364)	(865, 0.346)	(840, 0.358)	(886, 0.348)
β band (12-25Hz)	drowsy	(844, 0.411)	(857, 0.393)	(788, 0.421)	(944, 0.376)
β band (12-25Hz)	responsive	(870, 0.399)	(980, 0.361)	(967, 0.358)	(953, 0.368)
γ band (25-40Hz)	drowsy	(836, 0.424)	(795, 0.423)	(799, 0.428)	(772, 0.444)
γ band (25-40Hz)	responsive	(851, 0.402)	(995, 0.359)	(985, 0.353)	(962, 0.366)

Additionally, we observe the relationship between different frequency bands since the cross-frequency coupling has been crucial in EEG analysis. Modularity is plotted in Fig. 11, with the responsive group exhibiting a significantly strong correlation between β band and γ band compared to the drowsy group. This difference in patterns may potentially serve as a feature for predicting the participant group, without the need to conduct the injection experiment.

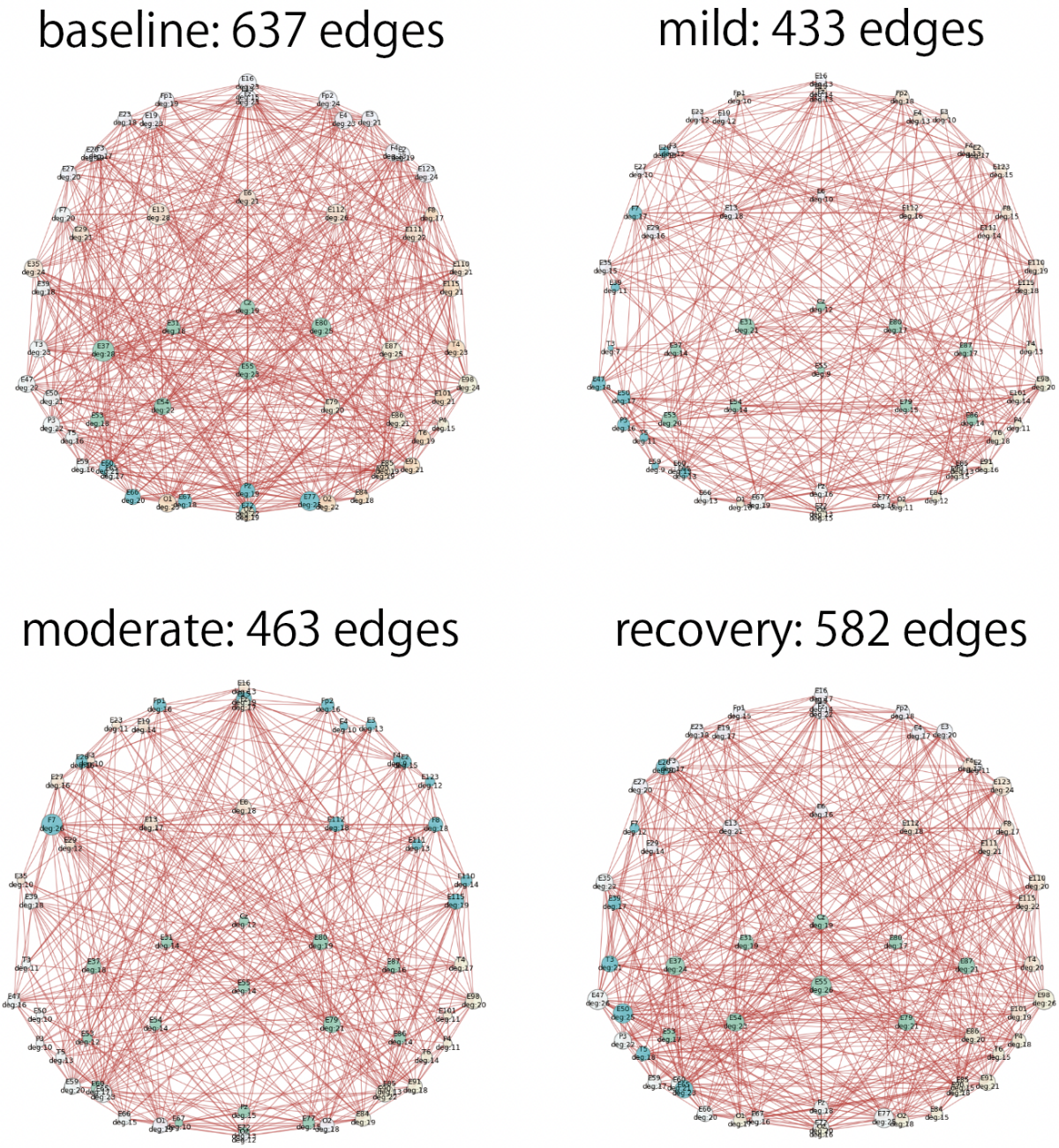


Fig 9. Change in reconstructed graph structures.(baseline → mild → moderate → recovery, participant_ID = 7 from the drowsy group). The color of each node indicates the community structure obtained via modularity maximization.

Real data analysis with marmoset ECoG

To further apply the proposed method for visualizing the phase-based connectivity, we used the 96 dimensional ECoG series data of marmosets. For example, the α bands of resting-state ECoG from marmosets were extracted and depicted in Fig. 12. For the data taken from the left hemisphere, the SMIC is minimized at the model with 667 edges. Contrarily, for the data taken from the right hemisphere, on the other hand, the

Table 6. Summary of graphical properties (average clustering coefficient and average shortest path length). The average is taken among all the participants in each participant group. Refer to S4 Appendix for full individual results.

	participant group	baseline	mild	moderate	recovery
α band (8-15Hz)	drowsy	(0.427, 1.583)	(0.272, 0.881)	(0.156, 0.574)	(0.125, 0.403)
α band (8-15Hz)	responsive	(0.478, 1.529)	(0.259, 0.823)	(0.162, 0.535)	(0.126, 0.389)
β band (12-25Hz)	drowsy	(0.475, 1.539)	(0.275, 0.876)	(0.166, 0.565)	(0.137, 0.391)
β band (12-25Hz)	responsive	(0.487, 1.524)	(0.289, 0.791)	(0.185, 0.511)	(0.136, 0.380)
γ band (25-40Hz)	drowsy	(0.467, 1.544)	(0.262, 0.894)	(0.166, 0.564)	(0.120, 0.421)
γ band (25-40Hz)	responsive	(0.478, 1.535)	(0.293, 0.787)	(0.189, 0.507)	(0.138, 0.378)

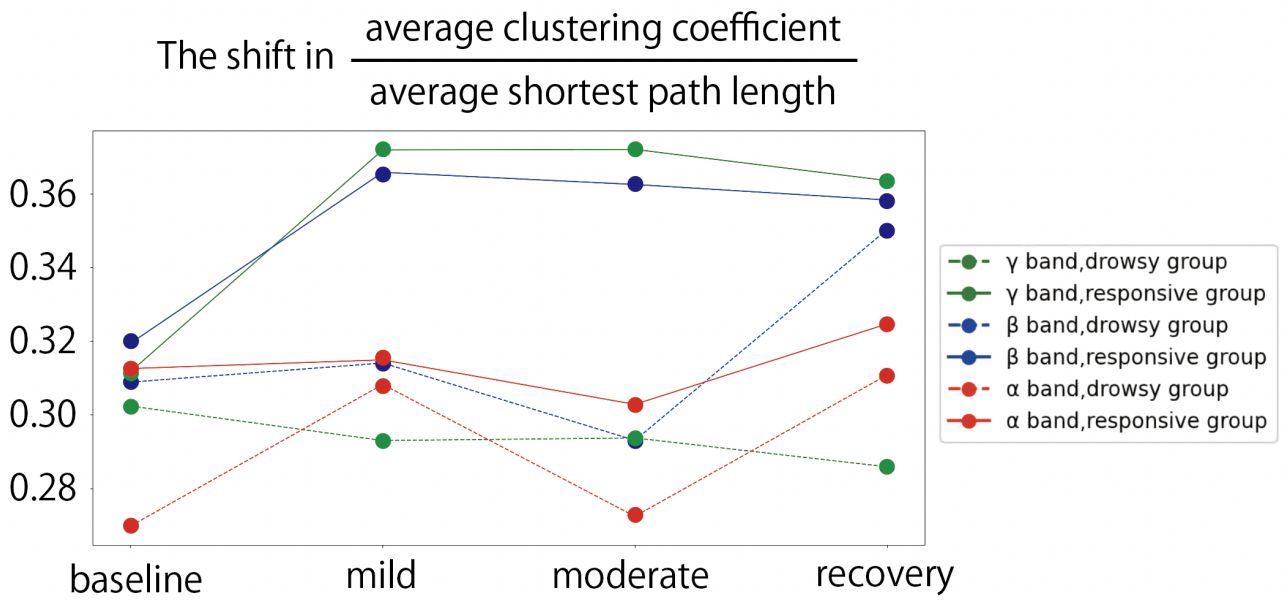


Fig 10. The shift in graphical properties during each participant state. The ratio between the average clustering coefficient and the average shortest path length in each participant state.

SMIC is minimized at the model with edges. In both cases, considering that the maximum number of edges can be as high as ${}_{96}C_2 = 4560$, the edges in the resulting graph are kept relatively sparse. The color of nodes represents the optimal community structure obtained by modularity maximization. Thus, due to the regularization effect of the proposed method, a sparse network structure with clearly defined community structure is obtained as a result.

To observe the shift in the data between before and after the ketamin injection, the number of edges and modularity in each state are summarized in Tab. 7. The network structure becomes less modular for the first marmoset, whereas the opposite is true for the second marmoset. These observations, based on the comparison of network modularity, are interesting in the context of previous studies that have explored the effects of anesthetics, including ketamine, on brain network architecture.

Many prior studies based on EEG signals in humans have consistently shown that from the awake to anesthetized state, the brain's functional connectivity is reduced [42]. In contrast, some ECoG studies in monkeys report that the overall functional connectivity shows an increasing trend in anesthetized state [43]. For instance, using phase-amplitude coupling to build functional connectivity, propofol-induced anesthesia enhances frontal and parietal interactions in monkeys [44]. These contrasting findings could reflect the complex and dynamic nature of ketamine's interaction with neural

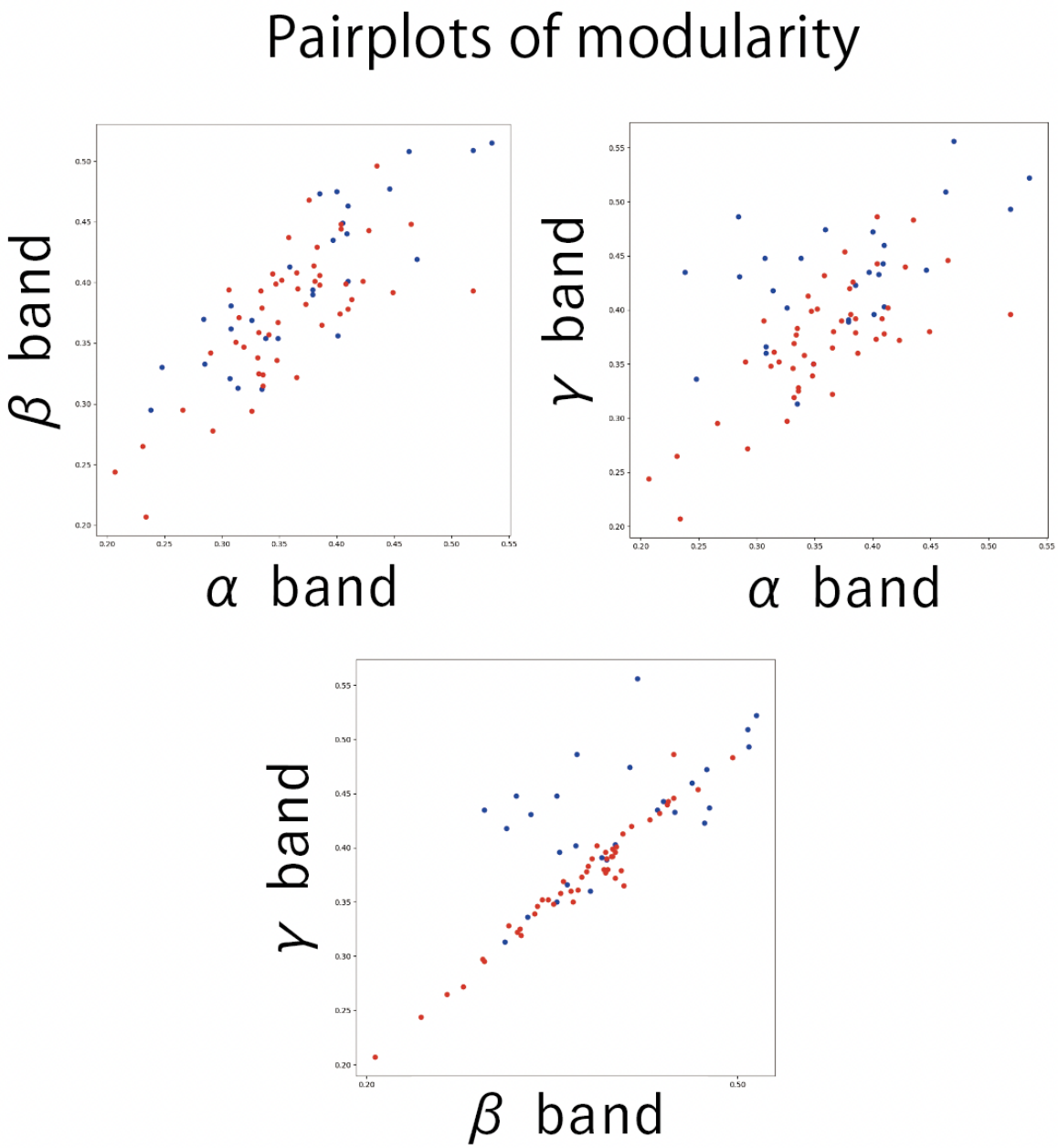


Fig 11. Pairplots of modularity from different frequency bands. The blue points are from the drowsy group whereas the red points are from the responsive group.

networks, where individual differences in brain structure or the specific nature of the intervention could yield divergent outcomes in terms of network topology. However, compared to modularity analysis based on EEG signals [45], modularity analysis based on ECoG signals is still limited. This is an important area for future work, as the network topology outcomes could depend on the target animal and the type of signal used.

The changes in network modularity we observed were consistent across all frequency bands, adding an important layer of insight. Previous studies (e.g., Yan et al. [46]) have

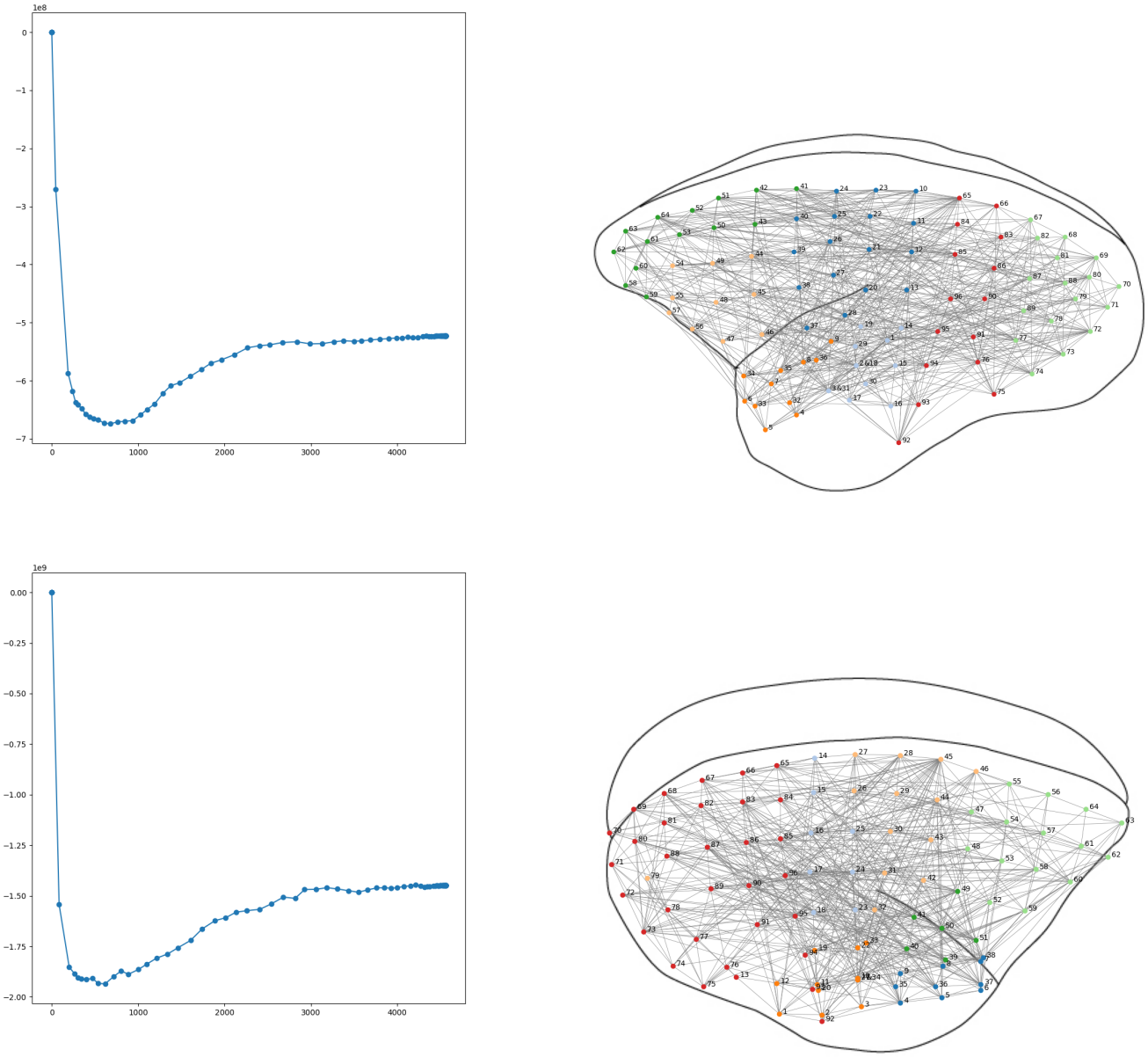


Fig 12. Results of the proposed method applied to the α band of two resting-state marmoset ECoG datasets. The number of edges is 667 and 603, respectively.

demonstrated that different frequency bands of ECoG in monkeys respond in opposite ways to acute ketamine injection. Our results, showing that the modularity changes occurred across all frequency bands, suggest that ketamine has a global effect on brain network topology, influencing both low and high-frequency oscillations. This finding contrasts with previous research that has highlighted frequency-specific effects of ketamine, indicating that the drug's impact on network dynamics may involve more complex, widespread alterations in brain activity.

Although the number of marmoset samples is too low to draw any solid conclusions, our results suggest the possibility of predicting an individual's state based on network

Marmoset ID Target brain	R03_0024_CM622M left hemisphere	R03_0035_CM814F right hemisphere
Resting-state (α band)	#edge = 680, modularity = 0.634	#edge = 618, modularity = 0.594
after 10 min. (α band)	#edge = 695, modularity = 0.600	#edge = 505, modularity = 0.622
Resting-state (β band)	#edge = 435, modularity = 0.666	#edge = 379, modularity = 0.623
after 10 min. (β band)	#edge = 798, modularity = 0.595	#edge = 368, modularity = 0.648
Resting-state (γ band)	#edge = 370, modularity = 0.673	#edge = 360, modularity = 0.655
after 10 min. (γ band)	#edge = 576, modularity = 0.631	#edge = 321, modularity = 0.657

Table 7. The shift in the recovered network structure from marmoset ECoG.

modularity extracted from ECoG. Considering the increasing use of ketamine in clinical contexts, understanding the individual variability in ketamine's effects on brain network topology could have significant implications for personalized treatment strategies. It also emphasizes the need for further studies with larger sample sizes to better understand ketamine's impact on brain networks.

Conclusion

We extended the score matching estimation method for the torus graph model proposed by Klein et al. [16] by applying Group LASSO regularization to achieve a sparse solution, enhancing the interpretation of EEG/ECoG phase-based connectivity analysis. Our method, implemented via the ADMM algorithm and SMIC minimization, was validated through simulation studies using synthetic data sampled from a pre-given torus graph model. We also applied our method to real EEG/ECoG data analysis and found that it contributes to a more sophisticated approach for analyzing high-dimensional phase-based functional connectivity, despite certain limitations, such as the disregard for time series structure including autocorrelations. From the human EEG dataset, we found that the modularity of the reconstructed brain network structure exhibits different tendencies when the β and γ bands are used; the modularity of β band connectivity is nearly the same as that of γ band connectivity in the responsive group, whereas the latter tends to have a larger value than the former in the drowsy group. Similar analyses were applied to marmoset ECoG data with ketamin injection, as our method is domain-agnostic.

Future research could focus on spatio-temporal analysis by further incorporating time series features and the relative positions of each electrode. Additionally, exploring the integration of our method with the source localization problem may yield valuable insights into brain connectivity.

Author Contributions

Conceived and designed the analysis: Takeru Matsuda. Performed the analysis: Issey Sukeda. Contributed reagents/materials/analysis tools: Issey Sukeda. Wrote and improved the paper: Issey Sukeda, Takeru Matsuda.

Supporting Information

S1 Appendix. Computational time of naive score matching (top) / Group LASSO (bottom) in seconds. In Tab. 8, the latter includes the former once per run because calculating SMIC requires the naive score matching estimator. Since our method is based on LASSO algorithm, it is expected to work in reasonable

computational time even for high-dimensional problems. Specifically, when the dimension of data is 61, the full torus graph model includes 7442 parameters and the rotational model includes 3782 parameters.

Table 8. Computational time of naive score matching (top) / Group LASSO (bottom) in seconds.

dimension of data	full torus graph model	rotational model
5	3.71	2.15
19	75.13	40.61
61	2012.72	939.60

dimension of data	full torus graph model	rotational model
5	142.67	57.87
19	1037.39	285.18
61	5383.70	1874.62

S2 Appendix. ADMM : the optimization algorithm for calculating score matching estimators. Since the score matching estimator with regularization term have the similar objective function to the LASSO problem, ADMM algorithm can be applied to calculate it. The following Algorithm 1 $\sim$ 2 describe the whole process.

Algorithm 1 ADMM (for Group LASSO)

Input: $\lambda(\in \mathbb{R}_+), \hat{\Gamma}, \hat{H}$, regularization weight c^g
Output: binary vector $\boldsymbol{\phi}^{bin}$

```
Initialize φ, z, u.
while converge do
    φ = (μI + Γ̂)^{-1}(μ(z − u) + Ĥ)
    for each group g do
        z^g ← R_{c^g λ/μ}(φ^g + u^g)
    end for
    u ← u + φ − z
end while
φ^bin ← sign(φ)
```

Algorithm 2 Regularized score matching of the torus graph via SMIC minimization

```
S ← [ ]
Calculate Γ̂, Ĥ.
φ* ← Γ̂^{-1} Ĥ
for λ ∈ logspace(−2, 1, 30) do
    φ^bin ← ADMM(λ)
    S ← calculate_SMIC(φ*, φ)
end for
return argmin_φ S
```

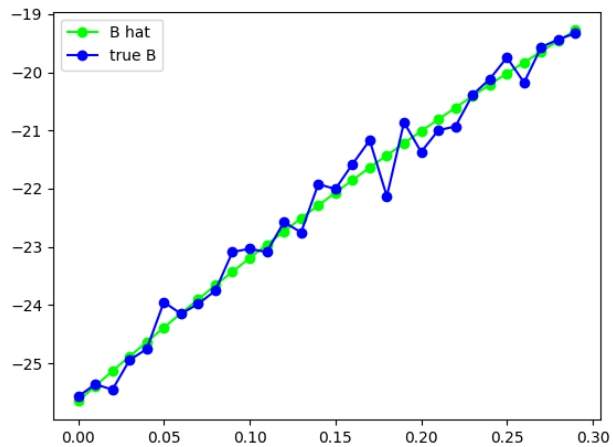


Fig 13. Comparison of B and $\hat{B}$ by the Monte Carlo simulation. The axes of the figure are B (or $\hat{B}$) versus ε.

S3 Appendix. Bias correction role of the SMIC. In general, information criteria is constructed by correcting the negative bias of the discrepancy between the empirical distribution using samples and the plug-in model (predictive distribution) when estimating the discrepancy between the true distribution and the model. The predictive distribution via score matching estimator is obtained as $p_{\hat{\phi}_{SM}} = p(x|\hat{\phi}_{SM}) \propto \exp(\hat{\phi}_{SM}^\top S(x))$. The score matching objective $\frac{1}{2}\phi^\top \hat{\Gamma}\phi - \phi^\top \hat{H}$ is an unbiased estimator of $d(q, p_\phi) = \int (\frac{1}{2}\phi^\top \Gamma(x)\phi - H(x)\phi)q(x)dx$, where $q(x)$ is the unknown true distribution. However, $\frac{1}{2}\hat{\phi}_{SM}^\top \hat{\Gamma}\hat{\phi}_{SM} - \hat{\phi}_{SM}^\top \hat{H}$ has an inherit bias when used for estimating $d(q, p_{\hat{\phi}_{SM}})$. The true bias can be calculated as

$$B = 2N\mathbb{E}_{x\sim q}[\frac{1}{2}\hat{\phi}_{SM}^\top \hat{\Gamma}\hat{\phi}_{SM} - \hat{\phi}_{SM}^\top \hat{H} - d(q, p_{\hat{\phi}_{SM}})].$$

On the other hand, the bias correction term of SMIC is an unbiased estimator of this bias:

$$\hat{B} = -2\mathrm{tr}(\hat{I}\hat{J}^{-1})$$

as in (4).

Here we confirm that $\hat{B}$ is truly an unbiased estimator of B by numerical simulations. We use the mixture of two bivariate von Mises distribution [13] (2-dimensional torus graph) as the true distribution:

$$\phi^{(1)} = (0, 0, 0, 0, 1, 1, 1, 1)^\top,\ \phi^{(2)} = (0, 0, 0, 0, 0.1, 0.1, 0.1, 0.1)^\top$$

$$q(x, y) = (1 - \epsilon)\frac{\exp((\phi^{(1)})^\top S(x, y))}{\int \exp((\phi^{(1)})^\top S(x, y))dxdy} + \epsilon\frac{\exp((\phi^{(2)})^\top S(x, y))}{\int \exp((\phi^{(2)})^\top S(x, y))dxdy}.$$

The first distribution is more concentrated to $(0, 0)$, whereas the second distribution has larger variance and functions as adding a noise.

We set $N = 1000$. The values B and $\hat{B}$ were computed by the Monte Carlo simulation with 10^5 repetitions. During the simulation, the expectation in $d(q, p_{\hat{\phi}_{SM}})$ was directly calculated by the numerical integration since the support of q is compact. Fig. 13 shows that B and $\hat{B}$ coincides well for various ϵ.

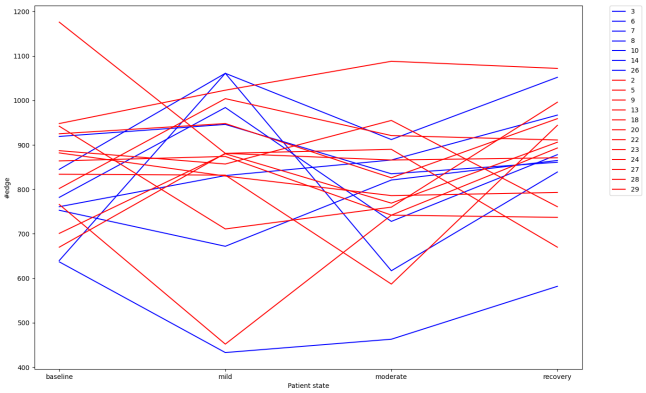
S4 Appendix. Full results for human EEG analysis.

Table 9. The changes in the number of edges, modularity, average clustering coefficient, and average shortest path length for the recovered graph structures in each state. The first, second, and third table corresponds to α, β, and γ band, respectively.

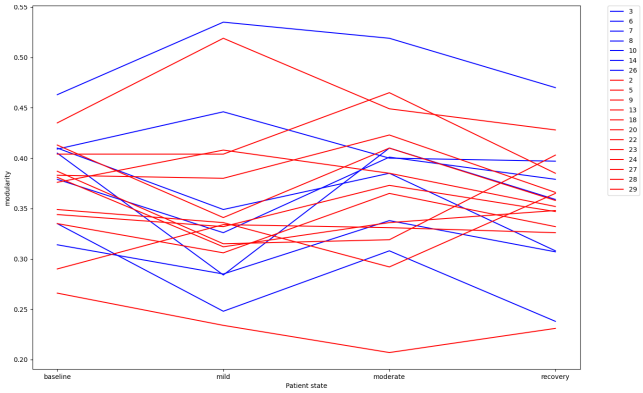
participant_ID	group	baseline	mild	moderate	recovery
3	drowsy	(640,0.405,0.380,1.650)	(1061,0.284,0.585,1.420)	(617,0.410,0.369,1.664)	(839,0.359,0.465,1.542)
6	drowsy	(845,0.335,0.475,1.538)	(1061,0.248,0.583,1.420)	(912,0.308,0.507,1.502)	(1052,0.238,0.583,1.425)
7	drowsy	(637,0.463,0.351,1.652)	(433,0.535,0.296,1.792)	(463,0.519,0.318,1.774)	(582,0.470,0.336,1.682)
8	drowsy	(919,0.314,0.507,1.498)	(946,0.285,0.514,1.483)	(835,0.338,0.459,1.544)	(861,0.307,0.467,1.530)
10	drowsy	(782,0.379,0.435,1.573)	(984,0.326,0.538,1.462)	(728,0.401,0.406,1.602)	(878,0.379,0.487,1.520)
14	drowsy	(761,0.410,0.422,1.584)	(831,0.349,0.464,1.546)	(866,0.385,0.482,1.527)	(967,0.308,0.526,1.472)
26	drowsy	(753,0.409,0.421,1.589)	(672,0.446,0.396,1.633)	(821,0.400,0.469,1.551)	(865,0.397,0.483,1.527)
2	responsive	(1176,0.290,0.648,1.357)	(881,0.334,0.499,1.519)	(866,0.331,0.487,1.527)	(871,0.326,0.503,1.524)
5	responsive	(925,0.335,0.508,1.495)	(948,0.306,0.526,1.482)	(826,0.365,0.452,1.549)	(959,0.332,0.522,1.476)
9	responsive	(864,0.383,0.471,1.528)	(874,0.380,0.474,1.522)	(742,0.423,0.413,1.595)	(893,0.366,0.494,1.512)
13	responsive	(834,0.404,0.461,1.544)	(832,0.404,0.461,1.545)	(587,0.465,0.342,1.681)	(944,0.385,0.524,1.484)
18	responsive	(701,0.413,0.400,1.617)	(880,0.341,0.481,1.519)	(769,0.410,0.431,1.580)	(906,0.358,0.507,1.505)
20	responsive	(948,0.266,0.524,1.482)	(1023,0.234,0.562,1.441)	(1088,0.207,0.598,1.405)	(1072,0.231,0.578,1.414)
22	responsive	(802,0.381,0.445,1.562)	(1004,0.312,0.556,1.451)	(921,0.336,0.505,1.497)	(911,0.348,0.504,1.502)
23	responsive	(887,0.349,0.491,1.515)	(857,0.336,0.474,1.532)	(955,0.292,0.526,1.478)	(761,0.365,0.419,1.584)
24	responsive	(942,0.376,0.518,1.485)	(711,0.408,0.397,1.612)	(760,0.385,0.421,1.585)	(996,0.352,0.545,1.456)
25	responsive	(817,0.368,0.457,1.554)	(1077,0.273,0.592,1.411)	(984,0.305,0.541,1.462)	(1009,0.281,0.559,1.449)
27	responsive	(882,0.344,0.481,1.518)	(830,0.332,0.459,1.546)	(786,0.373,0.439,1.570)	(793,0.347,0.443,1.567)
28	responsive	(670,0.387,0.398,1.635)	(881,0.315,0.489,1.519)	(890,0.319,0.494,1.514)	(670,0.403,0.396,1.634)
29	responsive	(766,0.435,0.409,1.581)	(452,0.519,0.292,1.776)	(742,0.449,0.409,1.595)	(737,0.428,0.414,1.597)

participant_ID	group	baseline	mild	moderate	recovery
3	drowsy	(694,0.449,0.406,1.621)	(901,0.370,0.508,1.508)	(836,0.401,0.468,1.543)	(760,0.413,0.450,1.585)
6	drowsy	(1125,0.312,0.627,1.385)	(943,0.330,0.526,1.485)	(794,0.381,0.453,1.566)	(1021,0.295,0.564,1.442)
7	drowsy	(562,0.508,0.339,1.694)	(569,0.515,0.333,1.691)	(530,0.509,0.373,1.725)	(950,0.419,0.521,1.481)
8	drowsy	(1117,0.313,0.611,1.390)	(998,0.333,0.549,1.455)	(963,0.354,0.539,1.474)	(988,0.321,0.541,1.460)
10	drowsy	(838,0.394,0.463,1.542)	(965,0.369,0.534,1.473)	(1125,0.356,0.617,1.385)	(971,0.390,0.531,1.469)
14	drowsy	(731,0.463,0.402,1.601)	(979,0.354,0.537,1.465)	(617,0.473,0.357,1.663)	(999,0.362,0.548,1.454)
26	drowsy	(844,0.440,0.479,1.539)	(646,0.477,0.389,1.648)	(654,0.475,0.398,1.643)	(922,0.435,0.515,1.496)
2	responsive	(979,0.342,0.548,1.465)	(799,0.393,0.450,1.563)	(971,0.338,0.542,1.469)	(1218,0.294,0.673,1.334)
5	responsive	(959,0.379,0.524,1.476)	(910,0.394,0.498,1.503)	(889,0.408,0.488,1.514)	(1066,0.359,0.584,1.417)
9	responsive	(832,0.429,0.467,1.545)	(1011,0.414,0.554,1.448)	(925,0.401,0.512,1.495)	(993,0.395,0.546,1.457)
13	responsive	(838,0.448,0.470,1.542)	(861,0.444,0.476,1.530)	(773,0.448,0.428,1.578)	(1005,0.406,0.547,1.451)
18	responsive	(1005,0.386,0.558,1.451)	(1078,0.357,0.595,1.411)	(999,0.378,0.550,1.454)	(741,0.437,0.426,1.595)
20	responsive	(1032,0.295,0.569,1.436)	(1286,0.207,0.709,1.297)	(1122,0.244,0.621,1.387)	(1066,0.265,0.590,1.417)
22	responsive	(819,0.401,0.459,1.552)	(856,0.351,0.473,1.532)	(1094,0.315,0.601,1.402)	(1086,0.336,0.599,1.407)
23	responsive	(906,0.367,0.508,1.505)	(1017,0.324,0.563,1.444)	(1072,0.278,0.590,1.414)	(1031,0.322,0.568,1.437)
24	responsive	(675,0.468,0.394,1.631)	(819,0.399,0.456,1.552)	(844,0.398,0.461,1.539)	(913,0.402,0.500,1.501)
25	responsive	(852,0.402,0.472,1.534)	(1096,0.320,0.605,1.401)	(1061,0.319,0.580,1.420)	(862,0.353,0.483,1.529)
27	responsive	(823,0.407,0.457,1.550)	(1044,0.325,0.574,1.430)	(882,0.382,0.488,1.518)	(783,0.399,0.449,1.572)
28	responsive	(922,0.365,0.516,1.496)	(880,0.371,0.494,1.519)	(962,0.347,0.537,1.474)	(881,0.374,0.494,1.519)
29	responsive	(674,0.496,0.387,1.632)	(1078,0.393,0.588,1.411)	(982,0.392,0.532,1.463)	(738,0.443,0.429,1.597)

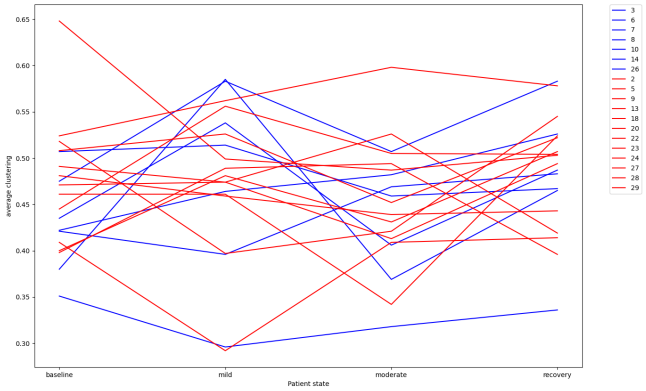
participant_ID	group	baseline	mild	moderate	recovery
3	drowsy	(918,0.433,0.503,1.498)	(535,0.486,0.381,1.717)	(828,0.403,0.465,1.548)	(565,0.474,0.383,1.697)
6	drowsy	(1117,0.313,0.622,1.390)	(911,0.336,0.509,1.502)	(901,0.360,0.503,1.508)	(941,0.435,0.522,1.486)
7	drowsy	(559,0.509,0.338,1.697)	(541,0.522,0.322,1.709)	(610,0.493,0.397,1.672)	(349,0.556,0.311,1.977)
8	drowsy	(830,0.418,0.456,1.546)	(788,0.431,0.443,1.569)	(760,0.448,0.436,1.585)	(680,0.448,0.402,1.628)
10	drowsy	(852,0.389,0.469,1.534)	(926,0.402,0.513,1.494)	(948,0.396,0.523,1.482)	(967,0.391,0.529,1.472)
14	drowsy	(741,0.460,0.406,1.595)	(996,0.350,0.545,1.456)	(876,0.423,0.490,1.521)	(986,0.366,0.542,1.461)
26	drowsy	(833,0.443,0.473,1.545)	(865,0.437,0.488,1.527)	(668,0.472,0.403,1.635)	(919,0.435,0.514,1.498)
2	responsive	(940,0.352,0.526,1.486)	(907,0.377,0.503,1.504)	(931,0.346,0.523,1.491)	(1204,0.297,0.665,1.342)
5	responsive	(950,0.383,0.518,1.481)	(928,0.390,0.508,1.493)	(1070,0.365,0.590,1.415)	(1023,0.369,0.562,1.441)
9	responsive	(841,0.426,0.472,1.540)	(996,0.420,0.547,1.456)	(1038,0.372,0.572,1.433)	(1050,0.380,0.574,1.426)
13	responsive	(629,0.486,0.377,1.656)	(865,0.443,0.478,1.527)	(779,0.446,0.432,1.574)	(1146,0.379,0.627,1.374)
18	responsive	(919,0.402,0.517,1.498)	(1076,0.358,0.593,1.412)	(1000,0.378,0.550,1.454)	(756,0.432,0.433,1.587)
20	responsive	(1032,0.295,0.569,1.436)	(1286,0.207,0.709,1.297)	(1122,0.244,0.621,1.387)	(1066,0.265,0.590,1.417)
22	responsive	(836,0.396,0.468,1.543)	(872,0.348,0.480,1.523)	(1051,0.328,0.579,1.426)	(1064,0.339,0.588,1.419)
23	responsive	(987,0.350,0.554,1.461)	(1014,0.325,0.562,1.446)	(1091,0.272,0.600,1.404)	(1028,0.322,0.566,1.438)
24	responsive	(706,0.454,0.407,1.614)	(857,0.392,0.474,1.532)	(863,0.392,0.470,1.528)	(924,0.401,0.507,1.495)
25	responsive	(758,0.421,0.431,1.586)	(1065,0.329,0.589,1.418)	(1055,0.319,0.577,1.423)	(838,0.361,0.471,1.542)
27	responsive	(788,0.413,0.439,1.569)	(1078,0.319,0.590,1.411)	(841,0.390,0.468,1.540)	(782,0.399,0.449,1.573)
28	responsive	(943,0.360,0.527,1.485)	(925,0.361,0.519,1.495)	(951,0.352,0.531,1.480)	(886,0.373,0.497,1.516)
29	responsive	(739,0.483,0.414,1.596)	(1063,0.396,0.580,1.419)	(1018,0.380,0.554,1.444)	(743,0.440,0.433,1.594)



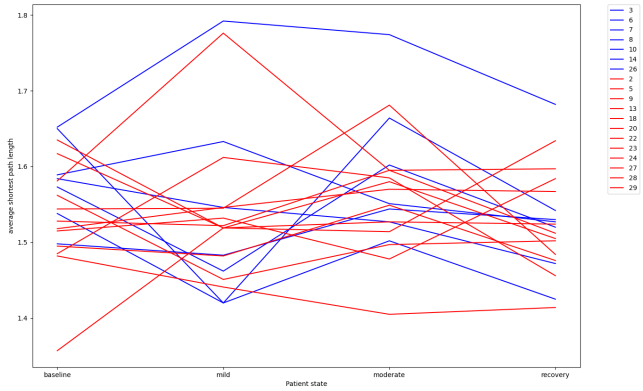
(a) The number of edges



(b) Modularity

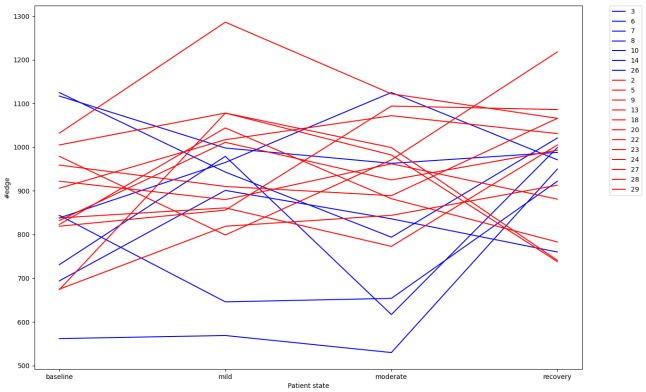


(c) Average clustering coefficient

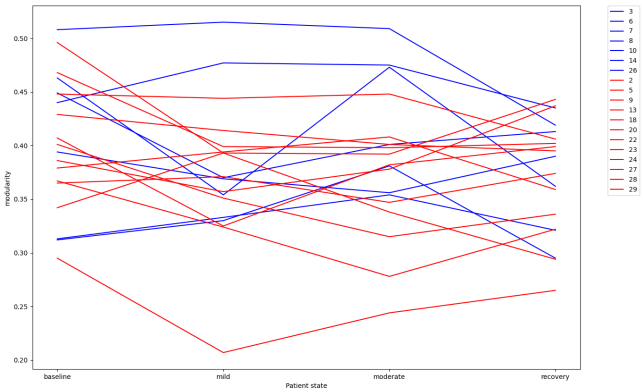


(d) Average shortest path length

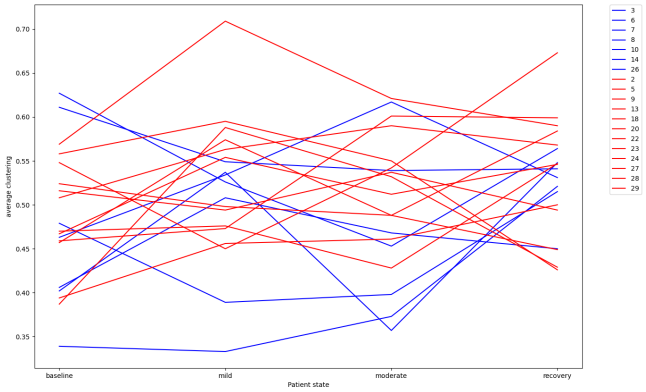
Fig 14. The shift in the recovered graph structures between participant states for α band EEGs.



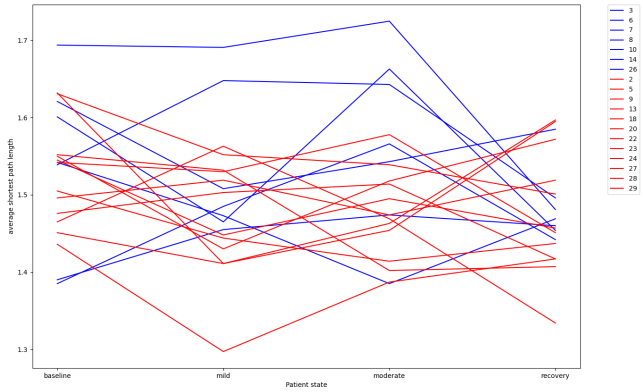
(a) The number of edges



(b) Modularity

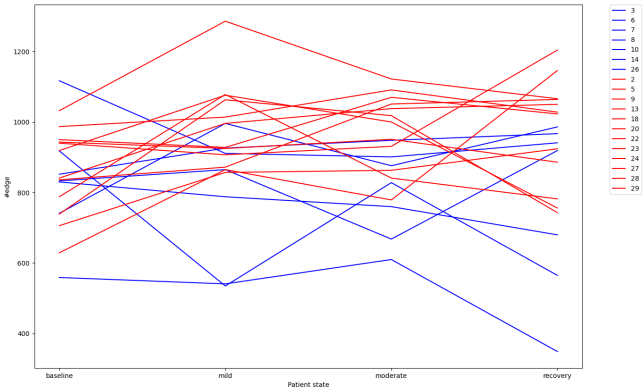


(c) Average clustering coefficient

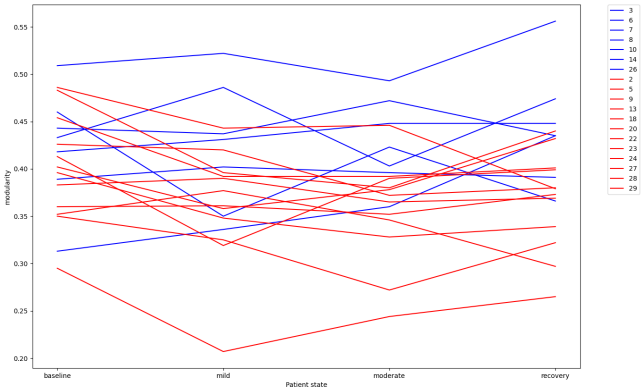


(d) Average shortest path length

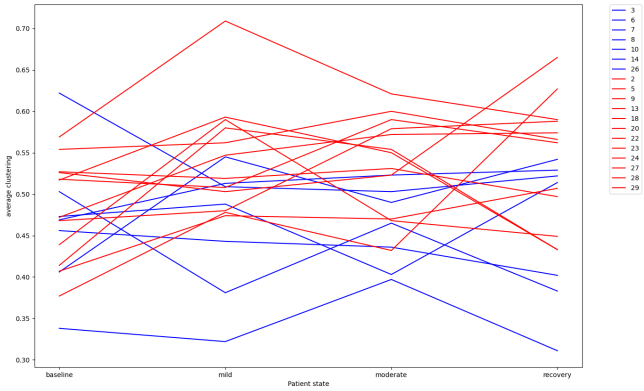
Fig 15. The shift in the recovered graph structures between participant states for β band EEGs.



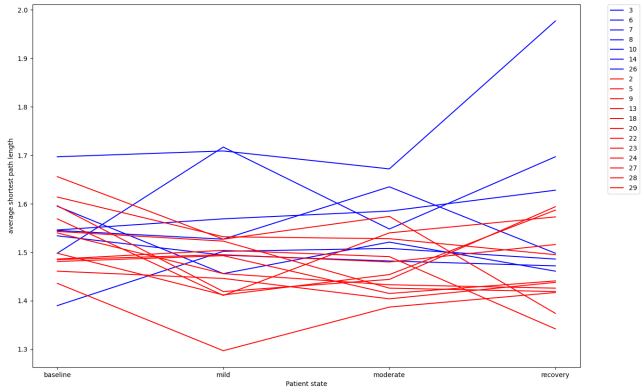
(a) The number of edges



(b) Modularity



(c) Average clustering coefficient



(d) Average shortest path length

Fig 16. The shift in the recovered graph structures between participant states for γ band EEGs.

Acknowledgments

Issey Sukeda was supported by RIKEN Junior Research Associate program. Takeru Matsuda was supported by JSPS KAKENHI Grant Numbers 22K17865, 24K02951, and JST Moonshot Grant Number JPMJMS2024. We also thank Dr. Yoshihito Saito and Dr. Sai Tanimoto for supportive comments.

References

1. Scheeringa R, Petersson KM, Kleinschmidt A, Jensen O, Bastiaansen MC. EEG alpha power modulation of fMRI resting-state connectivity. Brain connectivity. 2012;2(5):254–264.
2. Cohen MX. Analyzing neural time series data: theory and practice. MIT press; 2014.
3. Singer W, et al. Synchronization of cortical activity and its putative role in information processing and learning. Annual review of physiology. 1993;55(1):349–374.
4. Fries P. A mechanism for cognitive dynamics: neuronal communication through neuronal coherence. Trends in cognitive sciences. 2005;9(10):474–480.
5. Fries P. Neuronal gamma-band synchronization as a fundamental process in cortical computation. Annual review of neuroscience. 2009;32(1):209–224.
6. Salinas E, Sejnowski TJ. Correlated neuronal activity and the flow of neural information. Nature reviews neuroscience. 2001;2(8):539–550.
7. Thatcher RW, North D, Biver C. EEG and intelligence: relations between EEG coherence, EEG phase delay and power. Clinical neurophysiology. 2005;116(9):2129–2141.
8. Busch NA, Dubois J, VanRullen R. The phase of ongoing EEG oscillations predicts visual perception. Journal of neuroscience. 2009;29(24):7869–7876.
9. Roach BJ, Mathalon DH. Event-related EEG time-frequency analysis: an overview of measures and an analysis of early gamma band phase locking in schizophrenia. Schizophrenia bulletin. 2008;34(5):907–926.
10. Mormann F, Lehnertz K, David P, Elger CE. Mean phase coherence as a measure for phase synchronization and its application to the EEG of epilepsy patients. Physica D: Nonlinear Phenomena. 2000;144(3-4):358–369.
11. Lachaux JP, Rodriguez E, Martinerie J, Varela FJ. Measuring phase synchrony in brain signals. Human brain mapping. 1999;8(4):194–208.
12. Stam CJ, Nolte G, Daffertshofer A. Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. Human brain mapping. 2007;28(11):1178–1193.
13. Mardia KV. Statistics of directional data. Journal of the Royal Statistical Society Series B: Statistical Methodology. 1975;37(3):349–371.
14. Mardia KV, Jupp PE. Directional statistics. John Wiley & Sons; 2009.

15. Ley C, Verdebout T. Applied directional statistics: modern methods and case studies. CRC Press; 2018.

16. Klein N, Orellana J, Brincat SL, Miller EK, Kass RE. Torus graphs for multivariate phase coupling analysis. The Annals of Applied Statistics. 2020;14(2):635 – 660. doi:10.1214/19-AOAS1300.

17. Chennu S, O'Connor S, Adapa R, Menon DK, Bekinschtein TA. Brain connectivity dissociates responsiveness from drug exposure during propofol-induced transitions of consciousness. PLoS computational biology. 2016;12(1):e1004669.

18. Komatsu M, Ichinohe N. Effects of ketamine administration on auditory information processing in the neocortex of nonhuman primates. Frontiers in Psychiatry. 2020;11:826.

19. Komatsu M, Ichinohe N. Brain/MINDS Marmoset Brain ECoG Auditory Dataset 01 (DataID: 4924); 2021. https://dataportal.brainminds.jp/ecog-auditory-01.

20. Hyvärinen A. Estimation of non-normalized statistical models by score matching. Journal of Machine Learning Research. 2005;6(4).

21. Lin L, Drton M, Shojaie A. Estimation of high-dimensional graphical models using regularized score matching. Electronic Journal of Statistics. 2016;10(1):806 – 854. doi:10.1214/16-EJS1126.

22. Tibshirani R. Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B: Statistical Methodology. 1996;58(1):267–288.

23. Efron B, Hastie T, Johnstone I, Tibshirani R. Least angle regression. Ann Statist. 2004;32(1):407–499.

24. Boyd S, Parikh N, Chu E, Peleato B, Eckstein J, et al. Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends® in Machine learning. 2011;3(1):1–122.

25. Yuan M, Lin Y. Model selection and estimation in regression with grouped variables. Journal of the Royal Statistical Society Series B: Statistical Methodology. 2006;68(1):49–67.

26. Matsuda T, Uehara M, Hyvärinen A. Information criteria for non-normalized models. Journal of Machine Learning Research. 2021;22:1–33.

27. Schank T, Wagner D. Approximating clustering coefficient and transitivity. Journal of Graph Algorithms and Applications. 2005;9(2):265–275.

28. Newman MEJ, Girvan M. Finding and evaluating community structure in networks. Phys Rev E. 2004;69:026113. doi:10.1103/PhysRevE.69.026113.

29. Fortunato S, Barthelemy M. Resolution limit in community detection. Proceedings of the national academy of sciences. 2007;104(1):36–41.

30. Good BH, De Montjoye YA, Clauset A. Performance of modularity maximization in practical contexts. Physical Review E—Statistical, Nonlinear, and Soft Matter Physics. 2010;81(4):046106.

31. Li Z, Zhang S, Wang RS, Zhang XS, Chen L. Quantitative function for community detection. Phys Rev E. 2008;77:036109. doi:10.1103/PhysRevE.77.036109.

32. Costa A. MILP formulations for the modularity density maximization problem. European Journal of Operational Research. 2015;245(1):14–21.

33. Sukeda I, Miyauchi A, Takeda A. A study on modularity density maximization: Column generation acceleration and computational complexity analysis. European Journal of Operational Research. 2023;309(2):516–528.

34. Blondel VD, Guillaume JL, Lambiotte R, Lefebvre E. Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment. 2008;2008(10):P10008.

35. Chennu S, Annen J, Wannez S, Thibaut A, Chatelle C, Cassol H, et al. Brain networks predict metabolism, diagnosis and prognosis at the bedside in disorders of consciousness. Brain. 2017;140(8):2120–2132.

36. Lui KFH, Lo JCM, Ho CSH, McBride C, Maurer U. Resting state EEG network modularity predicts literacy skills in L1 Chinese but not in L2 English. Brain and Language. 2021;220:104984.

37. Koessler L, Maillard L, Benhadid A, Vignal JP, Felblinger J, Vespignani H, et al. Automated cortical projection of EEG sensors: Anatomical correlation via the international 10–10 system. NeuroImage. 2009;46(1):64–72. doi:https://doi.org/10.1016/j.neuroimage.2009.02.006.

38. Watts DJ, Strogatz SH. Collective dynamics of 'small-world'networks. nature. 1998;393(6684):440–442.

39. Muldoon SF, Bridgeford EW, Bassett DS. Small-world propensity and weighted brain networks. Scientific reports. 2016;6(1):22057.

40. Humphries MD, Gurney K. Network 'small-world-ness': a quantitative method for determining canonical network equivalence. PloS one. 2008;3(4):e0002051.

41. Lee H, Mashour GA, Noh GJ, Kim S, Lee U. Reconfiguration of network hub structure after propofol-induced unconsciousness. Anesthesiology. 2013;119(6):10–1097.

42. Hudetz AG. General anesthesia and human brain connectivity. Brain connectivity. 2012;2(6):291–302.

43. Xie T, Chen K, Ma L, Ai Q, Liu Q, Hudson AE. Brain connectivity analysis in anesthetized and awake states: An ECoG study in monkeys. In: 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). IEEE; 2021. p. 117–120.

44. Ma L, Liu W, Hudson AE. Propofol anesthesia increases long-range frontoparietal corticocortical interaction in the oculomotor circuit in macaque monkeys. Anesthesiology. 2019;130(4):560.

45. Kim H, Hudetz AG, Lee J, Mashour GA, Lee U, Group RS. Estimating the integrated information measure phi from high-density electroencephalography during states of consciousness in humans. Frontiers in human neuroscience. 2018;12:42.

46. Yan T, Suzuki K, Kameda S, Maeda M, Mihara T, Hirata M. Electrocorticographic effects of acute ketamine on non-human primate brains. Journal of Neural Engineering. 2022;19(2):026034.