Small-world scale-free brain networks from EEG with application to motor imagery decoding and brain fingerprinting

Mohammad Davood Khalili[a], Vahid Abootalebi[a], Hamid Saeedi-Sourck[a], Andrea Santoro[b,c], Harry H. Behjat[b,d,e,*]

[a]*Department of Electrical Engineering, Yazd University, Yazd, Iran*
[b]*Neuro-X Institute, École Polytechnique Fédérale de Lausanne, Geneva, Switzerland*
[c]*CENTAI Institute, Turin, Italy*
[d]*Department of Biomedical Engineering, Lund University, Lund, Sweden*
[e]*Department of Clinical Sciences Malmö, Lund University, Lund, Sweden*

Abstract

Developing individualized spatial models that capture the complex dynamics of multi-electrode EEG data is essential for accurately decoding global neural activity. A widely used approach is network modeling, where electrodes are represented as nodes. A key challenge lies in defining the network edges and weights, as precise connectivity estimation is critical for enhancing neural characterization and extracting discriminative features, such as those needed for task decoding. In this work, we propose a method for inferring subject-specific brain graphs from EEG data, explicitly designed to exhibit small-world and scale-free network properties. Our approach begins by computing phase-locking values between EEG channel pairs to build a backbone graph, which is then refined into an individualized small-world and scale-free network. To reduce computational complexity while preserving subject-specific characteristics, we apply Kron reduction to the resulting graph. We evaluated the proposed method on motor imagery decoding and brain fingerprinting tasks using two EEG datasets. Results show that our model consistently outperforms other benchmark graph models. Furthermore, we show that integrating classical EEG features with those derived using graph signal processing principles significantly improves performance. Overall, our findings highlight the potential of the proposed graph construction framework to enhance EEG analysis, with promising implications for a wide range of applications in cognitive neuroscience and brain-computer interface research.

Keywords: EEG, small-world, scale-free, graph reduction, graph signal processing, motor imagery, brain fingerprinting.

1. Introduction

Electroencephalography (EEG), which records neural activity via a network of scalp electrodes, remains a cornerstone technique for probing brain dynamics [1]. Although EEG excels in temporal resolution, its spatial resolution is limited and high correlations between channels pose additional challenges [2]. Network science has emerged as a powerful paradigm for elucidating the organizational principles of brain architecture and its complex functional interactions [3-5]. Within this framework, brain regions or sensors are represented as nodes, while their structural or statistical relationships are encoded as weighted edges [6, 7]. Nevertheless, constructing an optimal graph representation of EEG data remains an open and challenging research problem [8-10], critical for advancing our understanding of brain connectivity and improving downstream analytical tasks.

Empirical evidence shows that the brain has a small-world modular structure [4], comprising canonical functional networks across diverse temporal and spatial scales [11], alongside structural and functional hubs. This has significant biological implications, equipping the brain with an infrastructure to maintain the equilibrium between integration and segregation [4]. Recent studies suggest that this optimal trade-off stems from an evolutionary process aimed at minimizing wiring costs while maximizing global information processing efficiency [12-14]. Theoretically, it is well-accepted that the brain optimizes its operations by minimizing the costs of information processing, maximizing efficiency, and adaptively rerouting its operations to ensure resilience [15-17].

These diverse requirements underscore the need for EEG graphs that balance biological plausibility with computational tractability. In response, we propose a novel class of EEG-based brain graphs, referred to as Small-world Scale-free (SWSF), along with a lightweight variant incorporating Kron reduction, denoted $\text{SWSF}_{\text{Kron}}$. The proposed Small-world Scale-free brain graphs integrate small-world and scale-free characteristics, offering an effective balance between high communication efficiency and minimal wiring cost. In the SWSF graphs, a small-world architecture encodes global connectivity among a set of high-significance nodes, identified through functional clustering, while a scale-free structure models local connectivity among lower-significance nodes surrounding these hubs. The variant denoted $\text{SWSF}_{\text{Kron}}$, incorporates Kron reduction as a post-processing step,

*Corresponding author; email: harry.behjat@med.lu.se.

reducing computational complexity while preserving (1) inter-regional functional connectivity patterns, (2) spectral characteristics, and (3) essential topological features, including small-worldness and scale-freeness, through the Schur complement. Brain graphs exhibiting these features are known for their efficiency and resilience, which are critical in shaping plasticity, cognition, and insights into neurological diseases and disorders. Both the SWSF and $\text{SWSF}_{\text{Kron}}$ models accurately capture the spatial organization of the brain, emphasizing short transmission paths for high efficiency, a high clustering coefficient, and the emergence of hubs that facilitate coordinated information flow. In summary, these models effectively reflect the brain's balance between integration and segregation, establishing themselves as principled and biologically grounded EEG-based brain graphs.

To demonstrate the applicability of the proposed EEG-based brain graphs, we explore two proof-of-concept applications: motor imagery (MI) decoding [18-20] and brain fingerprinting [21-23]. Accurately distinguishing mental states from EEG data remains a persistent challenge, with a continued wide array of methods being proposed to address this issue [24]. Some methods focus on time- and frequency-based feature [25-27], while others focus on deriving spatial features from EEG topographies [28, 29]. Using two open-access EEG datasets, we show that our SWSF and $\text{SWSF}_{\text{Kron}}$ approaches provide superior performance over small-world and scale-free brain graphs in MI decoding and brain fingerprinting.

The remainder of the paper is organized as follows: Section 2 overviews key GSP concepts, details the proposed SWSF brain graph design, describes the datasets, and outlines methods for MI decoding and brain fingerprinting. Section 3 presents and discusses the experimental results. Section 4 provides our concluding remarks.

2. Materials and Methods

2.1. Graphs, their spectra, and graph signals

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \mathbf{W})$ denote an undirected connected graph characterized by the set of nodes $\mathcal{V}$, where $|\mathcal{V}| = N$ is the number of nodes and the set of edges $\mathcal{E}$, where $|\mathcal{E}| = M$ is the number of edges, and an $N \times N$ weighted adjacency matrix $\mathbf{W}$, with elements $W_{i,j} >= 0$ denoting the weight of edge $(i, j) \in \mathcal{E}$. The combinatorial graph Laplacian matrix is defined as $\mathbf{L} = \mathbf{D} - \mathbf{W}$, where $\mathbf{D}$ is the diagonal degree matrix with elements $D_{i,i} = \sum_{k=1}^{N} W_{i,k}$ [30]. The energy of $\mathcal{G}$ is defined as $E = \sum_{k=1}^{N} |\eta_k|$ [31], where $\{\eta_k\}$ denote the eigenvalues of $\mathbf{W}$. The Laplacian energy of $\mathcal{G}$ is defined as $LE = \sum_{k=1}^{N} |\lambda_k - D|$ [32], where D is the average node degree, $\{\lambda_k\}$ denote the eigenvalues of $\mathbf{L}$ obtained as diagonal elements of $\mathbf{\Lambda}$ in $\mathbf{L} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$, with $\mathbf{U}$ denoting the corresponding eigenvectors.

Let $\mathbf{f} \in \mathbb{R}^N$ denote a graph signal on $\mathcal{G}$, whose n-th component represents the signal value at the n-th node of $\mathcal{G}$. The graph Fourier transform (GFT) of $\mathbf{f}$ is given as [33]:

$$\hat{\mathbf{f}} = \mathbf{U}^T \mathbf{f}, \tag{1}$$

which satisfies the Parseval's relation: $\|\mathbf{f}\|_2 = \|\hat{\mathbf{f}}\|_2$ [34], ensuring energy conservation between the node and spectral domain.

The total variation of $\mathbf{f}$ is defined as:

$$TV(\mathbf{f}) = \left\| \mathbf{f} - \frac{1}{|\lambda_{max}|} \mathbf{W}\mathbf{f} \right\|_2^2, \tag{2}$$

where $\|.\|_2$ denotes the Frobenius norm and λ_{max} is the largest-magnitude eigenvalue of $\mathbf{W}$ [35]. Using $TV(\mathbf{f})$, the difference between the TV of the i-th and j-th graph signal is given as [36, 37]:

$$DTV(\mathbf{f}_i, \mathbf{f}_j) = \left| TV(\mathbf{f}_i) - TV(\mathbf{f}_j) \right|. \tag{3}$$

Lastly, the ensemble energy spectral density of a given graph signal set $\mathcal{F} = \{f_t\}_{t=1}^{T}$ is obtained as [34]:

$$\bar{\mathbf{e}}_{\mathcal{F}}(k) = \frac{1}{T} \sum_{t=1}^{T} \left| \hat{\mathbf{f}}_t[k] \right|^2, \quad k = 1, \ldots, N. \tag{4}$$

2.2. Network Models

Multiple network models have been proposed to represent the structural properties found in real-world networks [38]. Erdős-Rényi random graphs are networks wherein each edge has a fixed probability of being present or absent, resulting in a Poisson degree distribution [39]. They have a small clustering coefficient (CC) and characteristic path length (CPL) [40-42], and thus are not suitable to represent most real-world networks characterized by the presence of modules, hubs, and high CC. In contrast, regular networks feature structured link patterns——such as rings or grids——where each node connects to the same number of neighbors. These lattices typically have higher clustering coefficients and longer characteristic path lengths than random graphs [4].

Barabási and Albert [43] introduced a network growth model with preferential attachment, where new nodes connect to existing ones based on their degree, favoring high-degree nodes. Over many iterations, this forms a scale-free (SF) network with a power-law degree distribution (~3) and prominent hubs [44]. In contrast, Watts and Strogatz [45] proposed the seminal small-world (SW) network model, which begins with a regular ring lattice where each node connects to its $D/2$ nearest neighbors clockwise and $D/2$ counterclockwise. This initial configuration exhibits three defining properties: high CC, long CPL, and absence of hubs. The core mechanism of the model is stochastic edge rewiring: each existing edge is randomly reconfigured with probability ρ, producing three topological regimes:

- $\rho = 0$: Purely regular lattice (high CC, long CPL)
- $0 < \rho < 1$: SW network (high CC, short CPL)
- $\rho = 1$: Fully randomized network (low CC, short CPL).

For EEG applications, we generalize this model to 2D scalp-projected electrode arrays. Connections are defined relative to each electrode's radial line to the geometric centroid, with $D/2$ nearest neighbors selected in each of the clockwise and counterclockwise angular half-planes. This

adaptation preserves the SW generation mechanism through angular rewiring, respects neuroanatomical constraints, and maintains functional balance: high CC for local specialization and short CPL for global integration [13].

Let C and L denote the CC and CPL of $\mathcal{G}$, respectively. Let $C^{(\text{rnd})}$ and $C^{(\text{reg})}$ denote the CC of a random and regular graph of size $\mathcal{G}$, respectively. Similarly, let $L^{(\text{rnd})}$ and $L^{(\text{reg})}$ denote the CPL of a random and regular graph of size $\mathcal{G}$, respectively. The extent of small-worldness of $\mathcal{G}$ can be quantified using the SW index (I) measure, defined as [46]:

$$I = \frac{CL^{(\text{rnd})}}{LC^{(\text{rnd})}}, \tag{5}$$

or the SW propensity (P) measure, defined as [47]:

$$P = 1 - \sqrt{\frac{\Delta C^2 + \Delta L^2}{2}}, \tag{6}$$

where ΔC and ΔL are given as:

$$\Delta C = \frac{C^{(\text{reg})} - C}{C^{(\text{reg})} - C^{(\text{rnd})}}, \tag{7}$$

$$\Delta L = \frac{L - L^{(\text{rnd})}}{L^{(\text{reg})} - L^{(\text{rnd})}}. \tag{8}$$

In a SW network [45], $C/C^{(\text{rnd})} > 1$ and $L/L^{(\text{rnd})} \approx 1$, thus, a scalar summary of a small-world graph is $I > 1$ or $P > 0.6$ [46, 47]. The sparsity of the graph, which is a key parameter in its construction, can be quantified using the edge density, denoted $\bar{E}$, and computed as $\bar{E} = 2M/N(N-1)$ [17].

2.3. Proposed Method: SWSF Brain Graph Design

An overview of the proposed pipeline for designing SWSF brain graphs from EEG is illustrated as a block diagram in Fig. 1. The pipeline takes as input an N-channel EEG recording, consisting of K trials, each of length T, represented as a matrix $N \times T$, denoted $\mathbf{X}$. The time courses are assumed to have been suitably bandpass filtered to be confined in a desired frequency range. First, we construct a phase locking value (PLV) graph [48] that quantifies the phase synchronization of the electrodes, a widely used measure in brain-computer interface (BCI) studies [49-51].

The analytic signal for each row n (denoted $\mathbf{X}_{n,:}$) is computed using the Hilbert Transform, denoted $\mathbf{X}^{(a)}$, and is composed of real ($\mathbf{X}^{(r)} = \mathbf{X}$) and imaginary ($\mathbf{X}^{(i)}$) components, i.e., $\mathbf{X}^{(a)} = \mathbf{X}^{(r)} + i\mathbf{X}^{(i)}$ [52].[1] An instantaneous phase matrix, denoted $\mathbf{X}^{(\theta)}$, is then constructed, with elements:

$$\mathbf{X}^{(\theta)}_{n,t} = \texttt{angle}\left(\mathbf{X}^{(a)}_{n,t}\right), \tag{9}$$

representing the phase of electrode n at time point t. Using $\mathbf{X}^{(\theta)}$, we computed a PLV graph per trial, defined by its weighted adjacency matrix, a symmetric $N \times N$ matrix, denoted $\mathbf{W}$, with elements $\mathbf{W}_{n,m} \in [0, 1]$ defined as the ensemble-averaged instantaneous PLV between pairs of electrodes n and m:

$$\mathbf{W}_{n,m} = \frac{1}{T}\sum_{t=1}^{T}\left|\exp(i\Delta_{n,m,t})\right|, \tag{10}$$

where $\Delta_{n,m,t} = \mathbf{X}^{(\theta)}_{n,t} - \mathbf{X}^{(\theta)}_{m,t}$. If electrodes n and m have similar phases across time, $\mathbf{W}_{n,m} \to 1$ whereas if they have orthogonal phases $\mathbf{W}_{n,m} \to 0$. Inspired by prior works [50, 53, 54], the PLV graph was then sparsified by applying a threshold, denoted $\tau \in (0, 1)$, resulting in a thresholded PLV graph, denoted $\mathbf{W}^{(\tau)}$. We treat τ as a free parameter.

The methodologies for constructing small-world (SW), scale-free (SF), and small-world scale-free (SWSF) graph models are detailed in Algorithms 1, 2, and 3, respectively. These algorithms outline the systematic procedures for generating graph structures tailored to capture the topological characteristics critical to modeling brain networks.

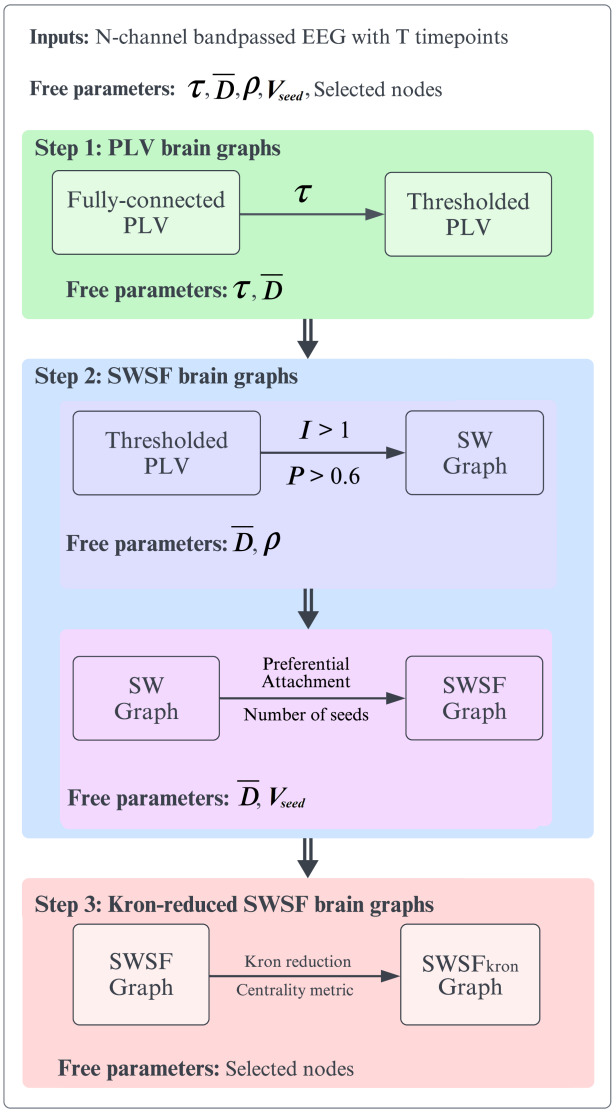

Fig. 1: The block diagram of the proposed method for inferring an SWSF graph from a given set of EEG data.

[1] Different methods can be used to estimate the analytic signal, here, we used the fast Fourier transform-based implementation [52] from MATLAB.

The construction of SWSF graphs follows a two-step procedure. First, SW graphs are constructed by connecting each electrode to its spatially nearest neighbors in the 2D scalp-layout electrode array. The seed vertex set V_{seed} is then determined using application-specific functional clustering, selecting nodes with the highest degree centrality within each cluster. Subsequently, SF graphs are constructed using V_{seed} as the initial seed vertices. The SWSF graphs are derived from the thresholded weight matrix $\mathbf{W}^{(\tau)}$, by selecting a suitable D as the construction parameter to balance computational complexity and network connectivity. To address the stochastic nature of the rewiring process, N_b bootstrap iterations were performed across all analyses: $N_b = 50$ for motor imagery decoding, and $N_b = 30$ for brain fingerprinting.

Random networks serve as a baseline for comparing the properties of SW networks. Thus, we used random networks during the construction of SW and SWSF graphs. Specifically, we constructed random networks with the desired average degree (D) and degree distributions identical to those of the original networks [55]. The parameters $C^{(\text{rnd})}$ and $L^{(\text{rnd})}$ were calculated as the means over N_b random graphs, following established methodologies in neuroscience [46, 56, 57]. Given 118 nodes in MI decoding, 1000 rewiring steps were sufficient to generate random graphs while preserving the nodal degree. This process was essential to establish the null distribution.

Algorithm 1 Small-World Graph Design

Inputs: N: number of vertices/electrodes; D: average degree of each vertex; ρ: probability of rewiring an edge; $\mathbf{W}^{(\tau)}$: thresholded weight matrix; N_b: number of bootstrap iterations.
Output: a set of N_b small-world graphs, $\text{SW}_{\text{graphs}}$; the ensemble average of $\text{SW}_{\text{graphs}}$ is used in both applications.

1: Initialize an empty set $\text{SW}_{\text{graphs}}$.
2: **for** $k = 1$ to N_b **do**
3: Build a binary graph $\mathcal{G}$ with N vertices, where each vertex connects to its $D/2$ nearest neighbors in both clockwise and counterclockwise directions, relative to the radial line from that vertex to the EEG electrode array centroid.
4: **for** each edge (u, v) in $\mathcal{G}$ **do**
5: Generate a random number $r \sim \text{Uniform}[0, 1]$.
6: **if** $r < \rho$ **then**
7: Select a random vertex z in $\mathcal{G}$ such that edge (u, z) does not already exist in $\mathcal{G}$.
8: Remove edge (u, v) from $\mathcal{G}$.
9: Add edge (u, z) to $\mathcal{G}$.
10: **end if** *(line 6)*
11: **end for** *(line 4)*
12: **for** each edge (u, z) in $\mathcal{G}$ **do**
13: Assign the weight from $\mathbf{W}^{(\tau)}$ to edge (u, z).
14: **end for** *(line 12)*
15: Add $\mathcal{G}$ to the list $\text{SW}_{\text{graphs}}$.
16: **end for** *(line 2)*
17: **Return:** $\text{SW}_{\text{graphs}}$.

Note: Lines 4 – 11 are adapted from the MATLAB function `SmallWorld(N, D, ρ)`.

Algorithm 2 Scale-Free Graph Design

Inputs: N: number of vertices; V_{seed}: set of initial vertices; Steplink: number of edges added at each execution step; $\mathbf{W}^{(\tau)}$: thresholded weight matrix; N_b: number of bootstrap iterations.
Output: a set of N_b scale-free graphs, $\text{SF}_{\text{graphs}}$; the ensemble average of $\text{SF}_{\text{graphs}}$ is used in both applications.

1: Initialize an empty set $\text{SF}_{\text{graphs}}$.
2: **for** $k = 1$ to N_b **do**
3: Build a random binary graph $\mathcal{G}$ by selecting a seed set V_{seed} from N vertices, where vertices in V_{seed} have predefined degrees and are interconnected.
4: **for** each vertex $v \in \mathcal{G}$ **do**
5: Calculate the degree of v, denoted $\deg(v)$.
6: **end for** *(line 4)*
7: **for** each vertex $v_{\text{new}} = 1$ to N, where $v_{\text{new}} \notin V_{\text{seed}}$ **do**
8: Add v_{new} to $\mathcal{G}$.
9: **for** $j = 1$ to Steplink **do**
10: Select a vertex $v_i \in \mathcal{G}$ with probability:

$$P(v_i) = \frac{\deg(v_i)}{\sum_{v \in \mathcal{G}} \deg(v)}$$

11: Add an edge between v_{new} and v_i in $\mathcal{G}$.
12: **end for** *(line 9)*
13: **end for** *(line 7)*
14: **for** each edge (v_{new}, v_i) in $\mathcal{G}$ **do**
15: Assign weight $\mathbf{W}^{(\tau)}(v_{\text{new}}, v_i)$ to the edge.
16: **end for** *(line 14)*
17: Add $\mathcal{G}$ to $\text{SF}_{\text{graphs}}$.
18: **end for** *(line 2)*
19: **Return:** $\text{SF}_{\text{graphs}}$.

Note: In each iteration, v_i is selected from vertices currently in $\mathcal{G}$ (includes V_{seed} and newly added vertices) whereas v_{new} is a vertex neither in V_{seed} nor previously added to $\mathcal{G}$.

Algorithm 3 Small-World Scale-Free Graph Design

Inputs: N: number of vertices; D_1: average degree for the small-world (SW) graph; ρ: probability of rewiring an edge; $\mathbf{W}^{(\tau)}$: Thresholded weight matrix; V_{seed}: Set of initial vertices for the scale-free (SF) graph; Steplink: Number of edges added at each execution step; N_b: Number of bootstrap iterations.
Output: a set of N_b small-world scale-free graphs, $\text{SWSF}_{\text{graphs}}$; the ensemble average of $\text{SWSF}_{\text{graphs}}$ is used in both applications.

1: **SW graph** ← **Algorithm 1** (N, D_1, ρ, $\mathbf{W}^{(\tau)}$, N_b).
2: $V_{\text{seed}} \leftarrow N_s$ vertices of the SW graph selected based on application-specific functional clustering and highest degree centrality.
3: **SWSF graph** ← **Algorithm 2** (N, V_{seed}, Steplink, $\mathbf{W}^{(\tau)}$, N_b).
4: **Return:** $\text{SWSF}_{\text{graphs}}$.

Note: N_s is the number of seed vertices selected from the SW graph based on application-specific functional clustering and highest degree centrality within these clusters.

To enhance robustness during the construction of SW, SF, and SWSF graphs while accounting for their stochastic nature, N_b brain graphs were generated for each scheme and each subject. Furthermore, we applied Kron reduction [58, 59] to reduce the size of the SWSF graphs. This reduction involved two steps: first, selecting the top 20% (i.e., 24 nodes) based on weighted nodal degrees, and second, simplifying the graph using Kron reduction with these selected nodes. Cross-validation results validating this node selection process (24 nodes) to construct the Kron-reduced SWSF graphs are shown in Fig. S1 (bottom, right) in the Supplementary Material.

Determining the free parameters. Our method includes several free parameters, which were chosen either based on prior studies or optimized using cross-validation (CV) on the training set. CV was performed on Dataset-1 for the MI decoding task (see Fig. S1 in the Supplementary Material). For the brain fingerprinting application on Dataset-2, parameters from the MI decoding task were reused where applicable, and others optimized independently. A detailed 10-fold CV strategy—using classification accuracy as the selection criterion—ensured that our parameter choices were data-driven, robust, and reproducible. Full details of the CV procedure are provided in the Supplementary Material.

Selection of τ. The threshold parameter τ was determined through CV on Dataset-1 (MI decoding task), evaluating values from 0.2 to 0.8 using classification accuracy as the criterion (see Fig. S1, top left, in the Supplementary Material). The optimal value was determined to be $\tau = 0.4$, which yielded graphs with an average degree of $D = 29.40 \pm 1.52$. This value balances graph sparsity and connectivity, ensuring an accurate representation of the underlying network properties. The same threshold τ was applied to Dataset-1 (MI decoding) and Dataset-2 (brain fingerprinting) for a consistent graph construction.

Joint Selection of ρ and D. To enforce small-world properties, we restricted the rewiring probability (ρ) from the broader range $0.1 < \rho < 0.9$ to $0.1 < \rho < 0.5$, as values above $\rho = 0.5$ did not satisfy the small-world criteria: $I > 1$ and $P > 0.6$ (Fig. 2). This refinement, supported by prior studies [15, 17, 47], provided acceptable I and P values. Detailed results on graph metrics and classification performance for varying values of ρ and D are presented in Fig. S2 and Table S1 of the Supplementary Material.

A joint grid search was conducted over a two-dimensional parameter space: ρ (0.1 to 0.5, step 0.1) and average degree D (12 to 26, step 2), guided by prior studies [16, 17, 46] to balance structural properties and MI decoding accuracy. The optimal values—$D = 20$ ($\bar{E} = 0.17$) and $\rho = 0.3$—achieved the highest training accuracy. Results are shown in Fig. S1 (top, right) in the Supplementary Material, highlighting small-world trade-offs.

Selection of V_{seed} nodes in SWSF graphs. To construct Small-world Scale-free (SWSF) brain graphs for MI decoding, we first created Small-world graphs with $N = 118$ nodes representing all EEG electrode regions. From these graphs, we selected 21 nodes (denoted as V_{seed}) based on an application-specific functional clustering strategy as in previous related work [60]. Specifically, We grouped electrodes into functional clusters related to motor imagery and selected nodes with the highest degree centrality within each cluster, ensuring spatial relevance and structural importance. As central regions like the motor and somatosensory cortices are more effective for MI decoding, we selected a final set of 21 nodes proportionally across clusters to balance performance and computational efficiency. The choice of 21 nodes for V_{seed} is justified in Fig. S1 (bottom, left) of the Supplementary Material, showing its effect on MI classification accuracy. The SWSF graph was then built using all 118 nodes, with V_{seed} defining the scale-free substructure. This maintained a comparable edge count to the SW and SF graphs, enabling fair structural and performance comparisons.

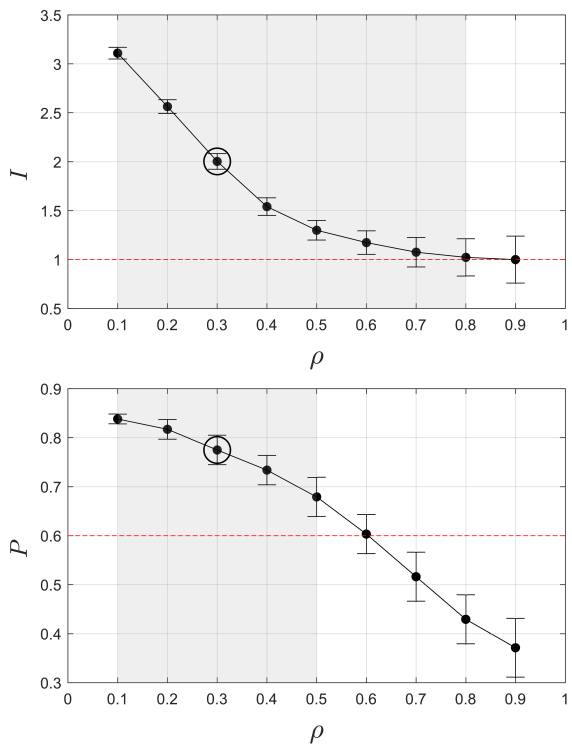

Fig. 2: Effect of ρ on small-world measures I and P; value represent the mean across subjects. The gray area indicates the acceptable values for ρ, and the red dashed line is the threshold for the small-worldness of the graph. Error bars indicate the standard deviation across subjects and bootstraps.

For brain fingerprinting, we used the same identified electrodes as in the motor imagery task, given their efficiency in capturing activity patterns. Following earlier studies on brain fingerprinting [21, 61, 62], the frontal and parietal regions, known for their efficiency in this application, were prioritized during node selection. These regions guided the selection of 21 V_{seed} nodes used to construct the SWSF graph with $N = 64$ nodes for brain fingerprinting. The selected electrodes are visualized in Fig. 3, which presents the aggregate mean and standard deviation of their topographic distributions across subjects for each of the six frequency bands. The mean maps indicate how often each electrode was chosen as a V_{seed}, revealing consistent spatial patterns that

reflect robust functional organization. In contrast, the variability maps highlight inter-subject differences: higher standard deviations point to subject-specific roles, while lower values indicate electrodes that are consistently selected or consistently ignored. Individual subject topographies are shown in Fig. S3 of the Supplementary Material. This frequency-dependent variability aligns with established functional specializations across frequency bands.

SF metrics. One of the frequently used approaches to examine whether a network exhibits scale-free properties is analyzing its degree distribution. If the network follows a power-law, the number of nodes with degree k, denoted by $N(k)$, should follow a distribution of the form $P(k) \sim k^{-\gamma}$, where the scaling exponent γ is typically in the range [2, 3]. In our study, we computed the histogram of node degrees and plotted the degree distribution on a log-log scale. In many cases, this yielded a linear trend in the tail of the distribution, with slope values (i.e., γ) within the expected range (Fig. S4). Although we refer to this behavior as power-law-like, it is more precise to state that the observed distributions are heavy-tailed, approaching power-law behavior asymptotically as $N \to \infty$. This aligns with the inherently finite nature of EEG graphs. Moreover, the coefficient of determination (R^2) of the fitted line to the tail regions was computed to assess the goodness-of-fit, and in most instances, relatively high R^2 values were obtained, indicating a reasonable fit to a heavy-tailed model. For each subject, we calculated the log-log plot, the R-squared value, and the slope of the regression curve to evaluate the adherence of SF graphs to a heavy-tailed degree distribution, following the methodology outlined in [16].

To support the scale-free characterization, we applied three Extreme Value Theory estimators—Hill, Moment, and Kernel—to the degree sequences of representative brain graphs. These provide statistically robust tail estimates, improving on naïve slope-fitting methods. Figure S5 displays the empirical log-log CCDFs (black dots) overlaid with the fitted curves from each method across five subjects. Across subjects, Hill and Kernel estimators yield γ values in the range of approximately 2.5 to 3, consistent with classic scale-free regimes. In contrast, the Moment estimator returns larger values, roughly 3.5 to 5.5, likely due to its known upward bias in finite samples with noisy tails. These discrepancies highlight the importance of using multiple estimators to assess the robustness of tail exponent estimates.

To further interpret these exponents, we apply Voitalov's taxonomy based on the tail index $\xi = 1/(\gamma - 1)$. The estimated ξ values from the Hill and Kernel methods fall around 0.5 to 1, placing most graphs in either power-law with divergent second moment or power-law. The Moment estimator's higher γ values imply $\xi < 0.25$, suggesting a few graphs fall into hardly power-law [63]. These results support the presence of heavy-tailed degree distributions across EEG-based brain graphs.

2.4. Datasets

This study employed two datasets: (i) motor-imagery task data from the BCI Competition III Dataset IVa [64, 65], and (ii) resting-state EEG data sourced from PhysioNet [66-68]. In the remainder of this paper, we refer to these two datasets as Dataset-1 and Dataset-2, respectively.

Dataset-1: Dataset IVa from BCI competition III consists of EEG signals from five healthy subjects, acquired using 118 electrodes at a sampling rate of 1000 Hz. A 2D projection of the electrode layout is shown in Fig. 4 (Left). Subjects were presented with visual cues for a duration of 3.5 seconds, during which they performed MI tasks involving their right hand or right foot. Random interruptions in the display of target cues, lasting between 1.75 and 2.25 seconds, served as rest periods for the subjects [64, 65]. We band-pass filtered the time-course of each electrode using a 5th-order Butterworth filter with a passband of 8-30 Hz to retain sensorimotor rhythms: mu (8-13 Hz) and beta (13-30 Hz), and subsequently down-sampled the signals to a frequency of 100 Hz. We then extracted, for each subject, 280 trials related to two classes of MI tasks (right hand and right foot), 140 trials per class. Each trial lasted 3.5 seconds (from 0.5 to 4 seconds after visual cue presentation), consistent with prior related work [69, 70].

Dataset-2: The resting-state EEG data from PhysioNet comprises EEG signals from 109 subjects, captured by 64 electrodes via the BCI2000 system [66, 67] at a sampling rate of 160 Hz. The location of 64 EEG electrodes in brain fingerprinting from Dataset-2 is shown in Fig. S6(a) in the supplementary material. Here, we used a subset of this dataset consisting of one-minute resting-state runs per subject (in the "eyes open" mode). For Dataset-2, we temporally bandpass filtered the resting-state EEG signals into six canonical frequency bands: delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), low beta (13-20 Hz), high beta (20-30 Hz), and gamma (30-50 Hz).

2.5. Application 1: MI Decoding

To perform MI decoding, we used features derived either directly from the EEG signals or by using the inferred SW, SF, SWSF, and Kron-reduced EEG graphs. Firstly, we extracted a range of classical features from the EEG signals. These features fall into eight broad categories: statistical, Hjorth, variation, energy, entropy, band power, time series, and those based on common spatial patterns (CSP). We also extracted two GSP-based features named $DTV(\mathbf{f}_i, \mathbf{f}_j)$ and $\bar{\mathbf{e}}_{\mathcal{F}}$. Given the large set of extracted features, we aimed to reduce their number to prevent redundancies, decrease computational cost, and improve performance generalization. Various methods have been proposed for feature selection within the context of MI decoding [71-74]. To this end, we used a meta-heuristic method known as differential evolution (DE) [75], a technique that has demonstrated superior performance in several related studies [60, 76]. We used the best subset of 10 features determined by this technique. These features were then used for classification using a support vector machine classifier with the radial basis function. We assessed the robustness of

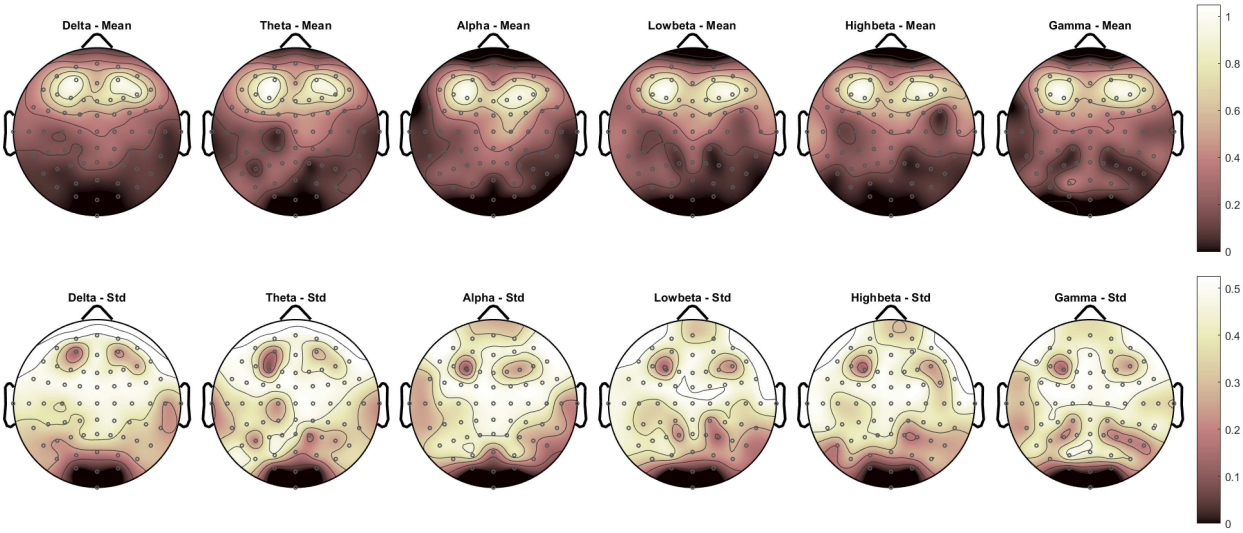

Fig. 3: Mean (top) and standard deviation (bottom) topographic plots of the 21 selected vertices (V_{seed}) across 109 subjects in the brain fingerprinting dataset, shown for all six frequency band.

classification results using 10-fold cross-validation, with 10 rounds of repetition.

2.6. Application 2: Fingerprinting

For Dataset-2, we preprocessed the resting-state EEG signals recorded during the "Eyes Open" mode by applying temporal bandpass filters to separate them into six standard frequency bands: delta (0.5-4 Hz), theta (4-8 Hz), alpha (8-13 Hz), low beta (13-20 Hz), high beta (20-30 Hz) and gamma (30-50 Hz). Data in each band was split into two parts, treating the initial 30 seconds as the test set and the second 30 seconds as the retest set. We then generated PLV, SWSF, and $\text{SWSF}_{\text{kron}}$ brain graphs for both the test and retest sets for each of the 109 subjects. The feature sets were defined as the vectorized form of the upper triangular section of the adjacency matrices corresponding to these graphs.

Maximizing human functional connectome fingerprinting is based on the principle that two connectivity profiles derived from the same subject should exhibit greater similarity than two connectivity profiles derived from two different subjects. We assessed the fingerprinting power of our proposed graphs by deriving identifiability measures introduced in [22], as previously used in the context of EEG [61] and MEG [77] data. Specifically, given S subjects, we denote the test and retest feature vectors of subject j as $\mathbf{y}_j^{(\text{test})}$ and $\mathbf{y}_j^{(\text{retest})}$ respectively, where $j = 1, \ldots, S$. We then construct the $S \times S$ identifiability matrix whose (i, k) entry is the PLV value computed between $\mathbf{y}_i^{(\text{test})}$ and $\mathbf{y}_k^{(\text{retest})}$. This matrix is not necessarily symmetric [77, 78].

Using the identifiability matrix, a scalar quantity named I_{self} can be defined as the average of the main diagonal elements. These elements represent the PLV value between the test and retest feature sets of identical subjects, indicating the average self-similarity across subjects. Similarly, another scalar value named I_{others} can be defined as the average of the off-diagonal elements. These elements represent the PLV value between the test and retest feature sets of different subjects, quantifying the average cross-subject similarity based on the given feature set. Additionally, a third scalar quantity named $I_{\text{diff}} = (I_{\text{self}} - I_{\text{others}}) \times 100$, can be defined. This value provides a robust group-level estimate of identifiability at the individual connectome level, where a higher value indicates greater individual fingerprinting [22, 77].

3. Results and Discussion

This section presents an evaluation of the mentioned brain graph designs, MI decoding results derived from SWSF and $\text{SWSF}_{\text{Kron}}$ brain graphs, along with brain fingerprinting results based on SWSF and $\text{SWSF}_{\text{Kron}}$ brain graphs. Specifically, we analyze the influence of the accompaniment of GSP-based and classical features, the results of the Kron reduction efficacy in the proposed approach, and the accuracy of the SVM-RBF using various brain graphs to assess the separability of two-class MI tasks. Additionally, we discuss the fingerprinting capabilities of SWSF and $\text{SWSF}_{\text{Kron}}$ over PLV-based graph and the influence of EEG duration on brain fingerprinting.

3.1. Graph-based Evaluation of Architectures

The adjacency matrix of the thresholded PLV, SW, SF, and SWSF graphs of a representative subject are shown in Fig. 4. In the SW structure, electrodes are predominantly interconnected with their neighboring electrodes in both the clockwise and anti-clockwise directions. Consequently, the adjacency matrix of the SW brain graph exhibits a dominant banded structure around the main diagonal, reflecting the preserved local connectivity among spatially adjacent electrodes. In the SF graph, based on the "preferential attachment", the highest edge weights within the adjacency matrix are observed around the initial electrodes. Given the

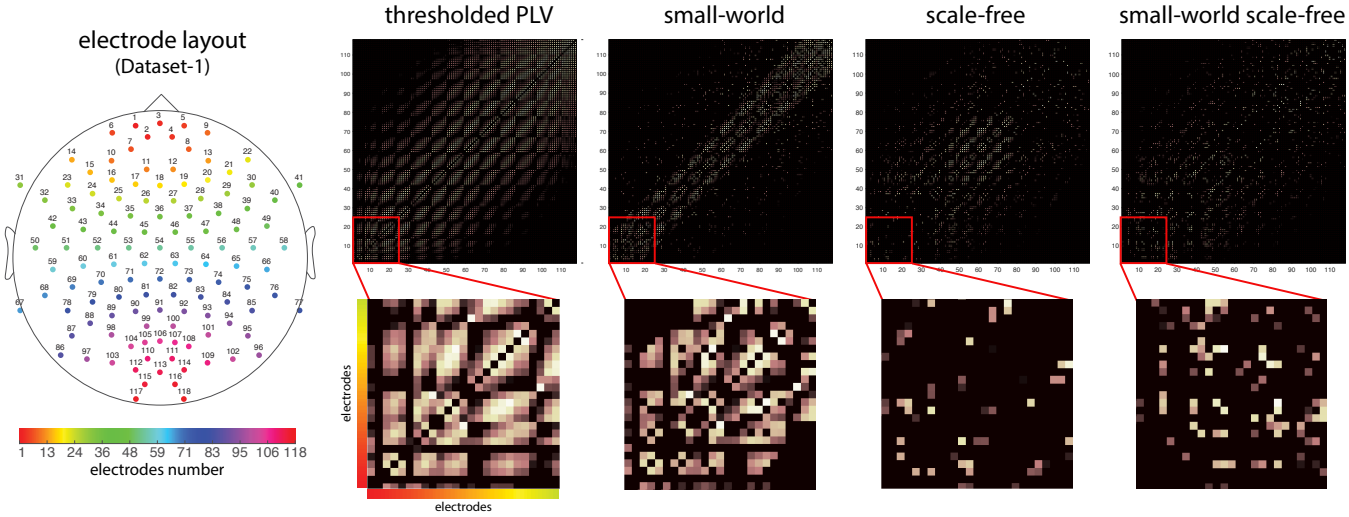

Fig. 4: Adjacency matrices for the thresholded PLV, SW, SF, and SWSF graphs of subject aa; cortical location of electrodes are shown on the left; the correspondence between electrode positions and anatomical regions is shown in Fig. S6(b) in the supplementary material. Adjacency matrices for the PLV, SW, SF, and SWSF graphs of the other subjects of Dataset-1 are shown in Fig. S7 in the Supplementary Material, respectively.

physiology of the motor cortex channels (used for MI decoding), we have placed these initial electrodes in the central brain regions. In the SWSF graph, high-significance nodes connect to adjacent ones, forming a SW structure visible as diagonal-adjacent stripes in the adjacency matrix. Increasing connections among these nodes introduces SF-like "preferential attachment," evident in the matrix's central rows and columns during MI decoding.

Figure 5 illustrates the degree distribution of SWSF brain graphs, which demonstrates heavy-tailed characteristics, as opposed to Gaussian distribution traits. To examine the SWSF brain graphs more precisely, we computed both the SW and SF measurements for all subjects; see Table 1. The mean C and mean L of the SWSF graphs are 0.234 and 3.449, respectively, whereas the mean I and P values are 1.041 and 0.640, respectively, which reflect the small-worldness of these graphs. The same measures but for the SW graphs are presented in Table S2 in the Supplementary Material.

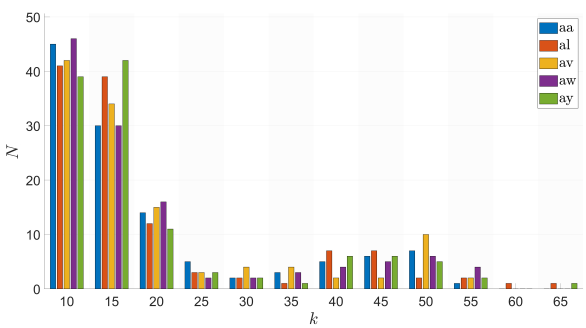

Fig. 5: Degree distributions of the SWSF graphs for all five subjects in the MI decoding application.

Furthermore, we found that the average exponent of the fitted power-law-like distribution for subjects in the SWSF design is approximately 2.41, with an R-squared value close to 0.85. Therefore, in comparison with the results obtained for the SF measures, it becomes evident that the application of the SW scheme leads to a reduction in both the exponent parameter and R-squared, which reflects the scale-free nature of these networks. The average number of hubs (N_h) for the SWSF graph across subjects is 22, nearly identical to SF graphs.

Table 1: Graph metrics for the constructed SWSF brain graphs in the MI decoding application. For all graphs and subjects: $N = 118$.

Metric	aa	al	av	aw	ay	Average
C	0.242	0.226	0.222	0.238	0.243	0.234
L	3.422	3.460	3.447	3.431	3.485	3.449
$C/C^{(rnd)}$	1.502	1.466	1.471	1.487	1.519	1.489
$L/L^{(rnd)}$	1.417	1.437	1.430	1.425	1.442	1.430
I	1.060	1.020	1.029	1.044	1.053	1.041
ΔC	0.506	0.485	0.492	0.513	0.477	0.495
ΔL	0.121	0.092	0.116	0.108	0.147	0.117
P	0.632	0.651	0.643	0.629	0.647	0.640
γ	2.51	2.45	2.40	2.36	2.34	2.41
R^2	0.84	0.85	0.85	0.84	0.86	0.85
N_h	23	22	22	22	21	22

The SW_{Kron}, SF_{Kron}, and $\text{SWSF}_{\text{Kron}}$ graphs for subjects 'aa,' 'al,' 'av,' 'aw,' and 'ay' each comprise 24 selected nodes and 276 corresponding edges, see Fig. 6. The three Kron-reduced graphs (SW_{Kron}, SF_{Kron}, and $\text{SWSF}_{\text{Kron}}$) originate from the same PLV-based graph, which consists of 118 nodes. Subsequently, these graphs are derived from the SW, SF, and SWSF graphs, each consisting of 118 nodes and characterized by distinct metrics and topologies.

3.2. Application 1: MI Decoding via SWSF Brain Graphs

MI decoding analysis was performed on Dataset-1. From the available set of classical and GSP features, the following top ten features from the different classes were selected

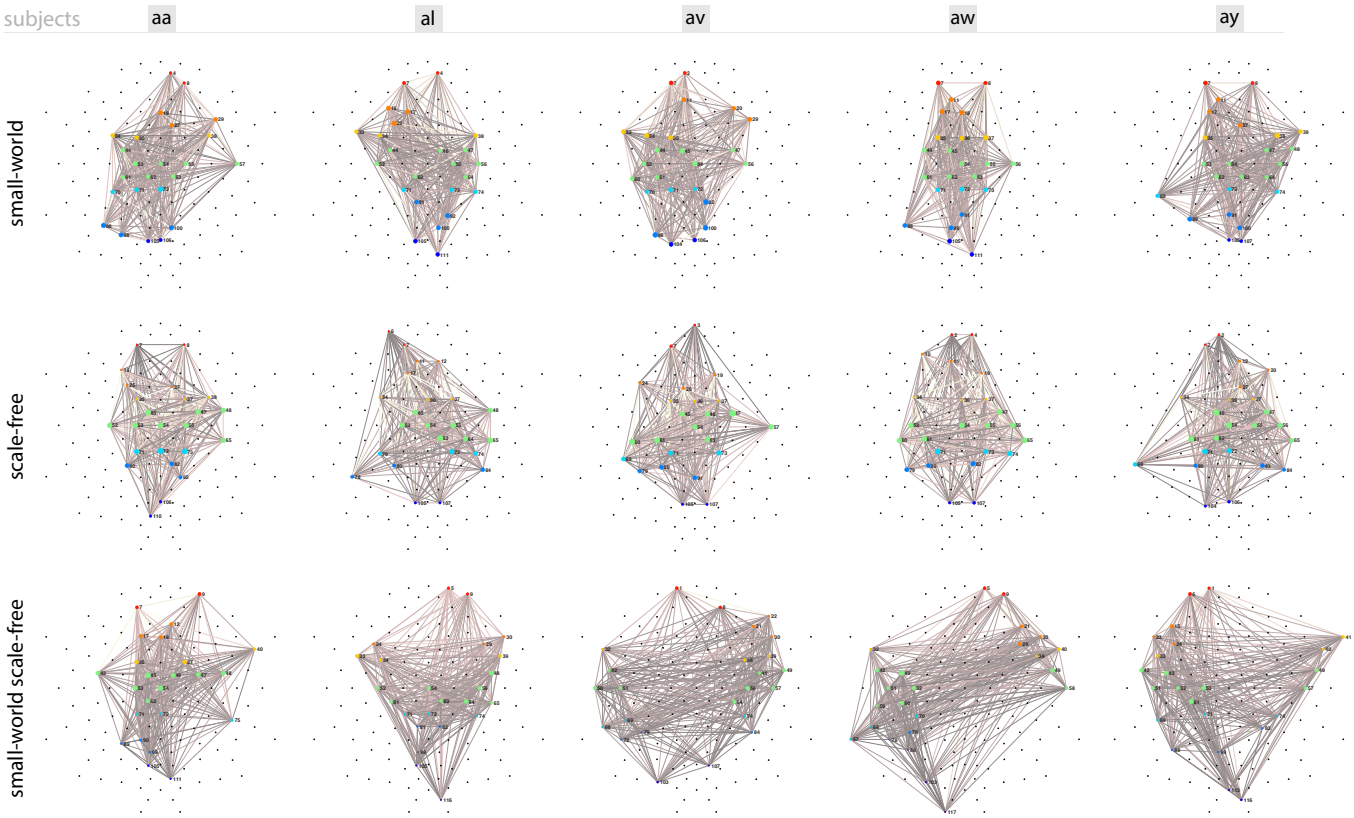

Fig. 6: Kron-reduced SW, SF, and SWSF graphs across all five subjects in the MI decoding application.

through differential evolution (cf. Section 2.5), which consist of, GSP features: $\bar{\mathbf{e}}_{\mathcal{F}}$ $DTV(\mathbf{f}_i, \mathbf{f}_j)$; temporal features: second difference, standard deviation, log root sum of sequential variations, mean Teager energy, log energy entropy; CSP features: CSP (1st filter); frequency features: μ band power; time series features: AR (4th coefficient). The GSP and classical features are described in detail in Section 2.1 and Table S3, respectively. The same 10 features were selected across all graph designs. Based on DE feature selection, the best fitness function was obtained when GSP-based features were combined with classic MI features. Combining $\bar{\mathbf{e}}_{\mathcal{F}}$ and $DTV(\mathbf{f}_i, \mathbf{f}_j)$ with classical features, including multiple time-domain features, the CSP feature, μ band power, and autoregressive (AR) coefficient, led to optimal results. We used a statistical two-sample t-test with a significance level of 0.01, to examine the level of distinguishability of the features. The t-test results showed that the variance ratio for the first CSP filter has the highest separability. Compared to the CSP feature, the GSP-based, time, and frequency features had lower absolute t-values and higher p-values.

Table 2 compares MI decoding classification accuracy results using five different methods: CLA, PLV, SW, SF, and SWSF. The SW, SF, and SWSF graphs outperform the PLV graph and CLA; the CLA method only uses classical features whereas PLV, SW, SF, and SWSF use both classical and GSP features as they are graph-based. The performance gradually increases when transitioning from the CLA model to the SWSF model, with SWSF showing superior performance over all the other methods.

The results of MI decoding using the Kron-reduced graphs are reported in Table 3. Kron reduction enhanced the classification accuracy across all graphs. The $\text{SWSF}_{\text{Kron}}$ model outperforms SW_{Kron}, SF_{Kron}, and PLV_{Kron} models in all subjects. Moreover, a comparison of Tables 2 and 3 shows that the average performance of all four models improved when combined with Kron reduction, demonstrating the utility of this method. Compared to the original schemes (PLV, SW, SF, and SWSF), the average MI decoding accuracy in the Kron-reduced schemes (PLV_{Kron}, SW_{Kron}, SF_{Kron}, and $\text{SWSF}_{\text{Kron}}$) increased by 1.65%, 2.79%, 1.79%, and 3.00%, respectively. As confirmed by the results in Table 3, similar to those in Table 2, SWSF brain graphs yield superior results in MI decoding compared to the SW and SF brain graphs.

Table 2: MI decoding accuracy (mean±std,%) of the proposed method for five different models of brain graph ($N = 118, M = 1180$). For CLA, the top 10 classical features were used whereas for PLV, SW, Sf, and SWSF the top 10 GSP-based and classical features were used .

Subject	CLA	PLV	SW	SF	SWSF
aa	78.9±3.0	79.6±3.2	82.1±2.3	80.4±2.7	**84.3±1.7**
al	92.1±3.2	90.0±3.1	95.0±2.0	92.9±2.2	**96.1±1.6**
av	72.9±2.5	73.2±4.4	78.9±2.6	73.9±2.8	**79.3±2.2**
aw	82.1±2.8	83.9±3.9	88.6±2.5	85.4±2.0	**90.7±1.9**
ay	83.2±2.5	82.1±4.6	88.6±2.6	82.9±2.7	**91.8±1.7**
mean±std	81.8±2.8	81.8±3.8	86.6±2.4	83.1±2.5	**88.4±1.8**

Table 3: MI decoding accuracy (mean±std,%) of the proposed method for four different models of brain graph with Kron reduction.

Subject	$\mathbf{PLV}_{\text{Kron}}$	$\mathbf{SW}_{\text{Kron}}$	$\mathbf{SF}_{\text{Kron}}$	$\mathbf{SWSF}_{\text{Kron}}$
aa	80.7±5.9	85.4±2.1	81.8±2.0	**88.2±1.5**
al	92.5±3.1	97.1±2.3	93.6±1.8	**98.6±1.9**
av	75.0±3.1	80.4±2.0	76.1±2.1	**82.9±2.0**
aw	85.7±2.4	91.8±1.8	86.4±1.9	**94.6±1.5**
ay	83.2±5.0	92.5±2.2	86.4±1.7	**92.9±1.5**
mean±std	83.4±3.9	89.4±2.1	84.9±1.9	**91.4±1.7**

3.3. Application 2: Fingerprinting via SWSF Brain Graphs

Brain fingerprinting analysis was performed on Dataset-2. In the following, we compare the performance of PLV, SWSF, and $\text{SWSF}_{\text{kron}}$ brain graphs in brain fingerprinting under varying rewiring probabilities and EEG time lengths.

Figure 7 shows fingerprinting results as a function of the rewiring probability for SWSF, and $\text{SWSF}_{\text{kron}}$ graphs, across six frequency bands; the parameter ρ does not apply to PLV graphs. Optimal subject identifiability was found for $\rho = 0.6$ for the Delta and Theta bands, and $\rho = 0.3$ for the Alpha, Low-Beta, High-Beta, and Gamma bands. In both SWSF and $\text{SWSF}_{\text{kron}}$ graphs, the Alpha and Beta bands, which are associated with both motor imagery and fingerprinting, exhibited the best performance for $\rho = 0.3$. In contrast, for the Delta and Theta bands — which we study only in the context of fingerprinting — the best identifiability was obtained for $\rho = 0.6$. Lastly, $\text{SWSF}_{\text{kron}}$ generally outperforms SWSF across all six frequency bands, except for the Gamma band, where both methods demonstrated comparable performance.

Table 4 presents a comparison of the best performances of SWSF and $\text{SWSF}_{\text{kron}}$ in brain fingerprinting across six frequency bands for the optimal ρ. As shown in Fig. 7, these results are benchmarked against PLV for evaluation. The results demonstrate a clear superior performance of the inferred graphs (SWSF and $\text{SWSF}_{\text{kron}}$) in the I_{diff} measure, compared to the conventional PLV-based graphs, across all frequency bands.

Table 4: I_{diff} (%) for fingerprinting on Dataset-2.

Frequency-Band	PLV	SWSF	$\text{SWSF}_{\text{Kron}}$
Delta	18.56	22.35	**25.28**
Theta	20.47	29.77	**40.99**
Alpha	15.74	26.66	**30.47**
Low-Beta	14.58	28.02	**39.53**
High-Beta	16.75	35.30	**43.96**
Gamma	13.44	26.45	**28.29**

To investigate the impact of EEG time length on fingerprinting, we performed identification using varying trial lengths in Dataset-2. To enhance the stability and robustness of the investigation, the bootstrap method was employed to select random data subsets, with 30 repetitions. In Dataset-2, the length of EEG epochs used to derive the graphs was varied from 2 to 30 seconds (2, 3, 4, 5, 6, 7, 8, 10, 12, 15, 20, 30) for both the test (session 1) and retest (session 2) sets, due to limited trial availability. The power of identifiability is depicted as a function of the epoch length, ranging from 2 to 30 seconds, for all six frequency bands. Differential identifiability was calculated as the average of 30 bootstrap iterations for each trial instance. In each iteration, n trials (n seconds) were randomly selected from the EEG data of each subject across all six frequency bands, and the PLV, SWSF, and $\text{SWSF}_{\text{kron}}$ brain graphs were subsequently derived.

Figure 8 shows fingerprinting performance versus resting-state data length. Overall, performance improves with longer recordings across all models, except in the Gamma and Theta bands. In the Gamma band, all models peak at 3 seconds, then show minor fluctuations up to 30 seconds. In the Theta band, the SWSFkron graph peaks at 3 seconds and slightly declines thereafter. As EEG length increases, Idiff rises and its standard deviation decreases across all bands, enhancing identifiability and fingerprinting effectiveness.

Across all settings, SWSF and $\text{SWSF}_{\text{kron}}$ models consistently outperform PLV graphs, except for the Delta band where the performances are nearly identical. The most significant performance difference between SWSF and $\text{SWSF}_{\text{kron}}$, in comparison to PLV graphs, is observed in the Low Beta and High Beta bands. A pattern that remains consistent across varying epoch lengths.

4. Conclusion

We proposed a novel method for constructing brain graphs from EEG data, generating subject-specific networks that reflect small-world and scale-free properties based on the phase-locking value between electrode time courses. To reduce complexity while retaining essential information, we applied Kron reduction. By combining graph-derived features with classical ones, we demonstrated improved performance in motor imagery decoding and brain fingerprinting across two independent datasets. These results underscore the potential of our graph design for broader EEG analysis and applications in cognitive neuroscience.

CRediT authorship contribution statement

Mohammad Davood Khalili: Conceptualization, Methodology, Software, Data curation, Formal analysis, Investigation, Visualization, Writing – original draft, Writing – review & editing. **Vahid Abootalebi:** Conceptualization, Methodology, Supervision, Writing – review & editing. **Hamid Saeedi-Sourck:** Supervision, Writing – review & editing. **Andrea Santoro:** Methodology, Writing – review & editing. **Harry H. Behjat:** Conceptualization, Methodology, Software, Investigation, Visualization, Supervision, Writing – review & editing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have

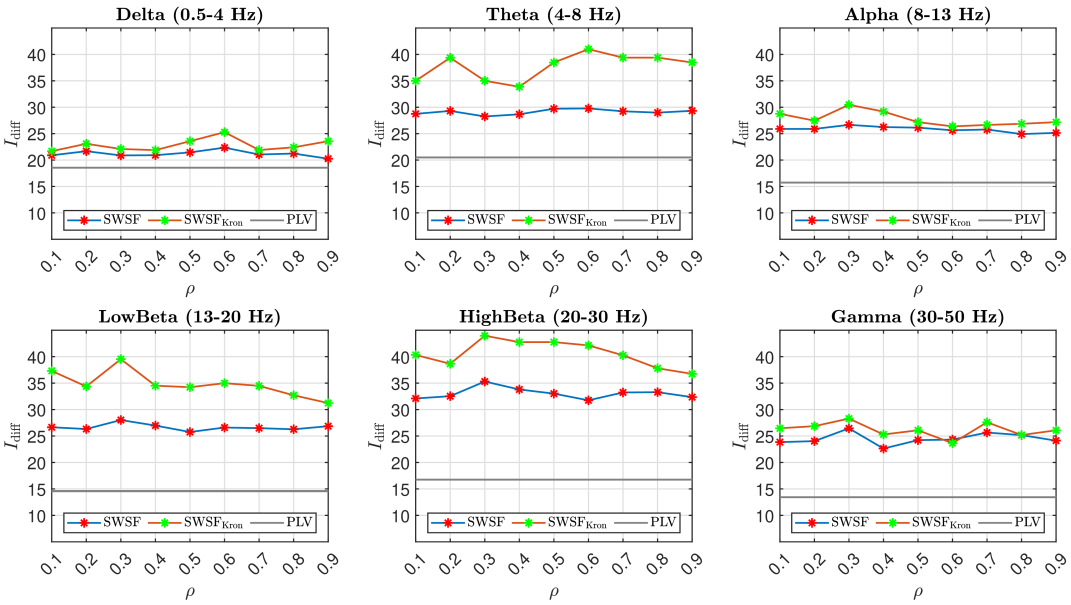

Fig. 7: Differential identifiability power of PLV, SWSF, and SWSF$_{\text{kron}}$ graphs as a function of the rewiring probability (ρ) across six frequency bands on Dataset-2.

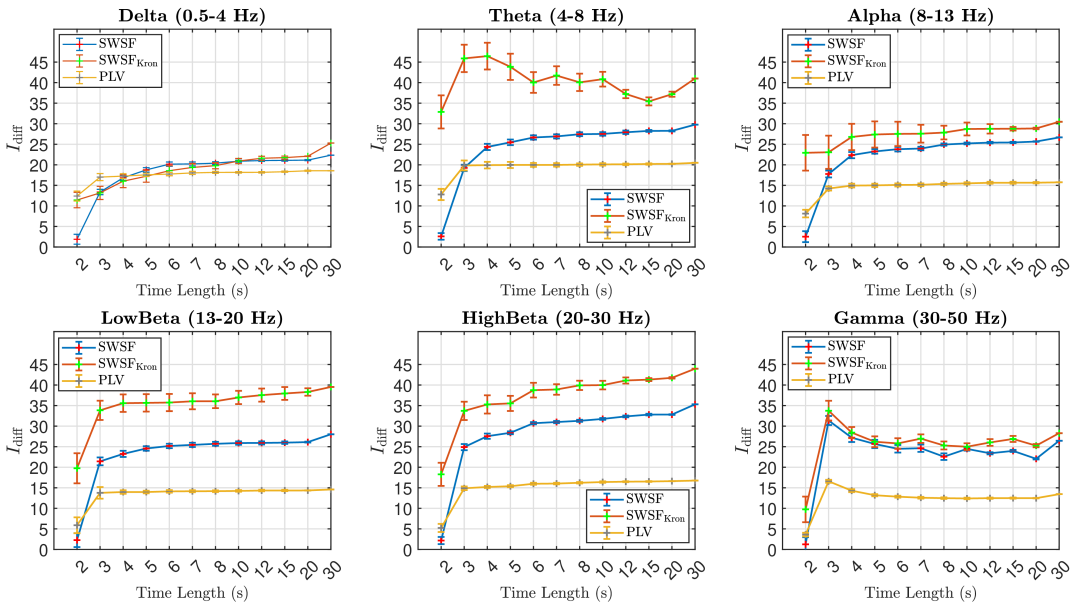

Fig. 8: Differential identifiability power of PLV, SWSF, and SWSF$_{\text{kron}}$ graphs as a function of the length of epochs used to derive the graphs for six frequency bands on Dataset-2. The error bars show the standard deviation across 30 bootstrap iterations.

appeared to influence the work reported in this paper.

Ethics Statement

This research study was conducted retrospectively using human subject data made available in open access. Ethical approval for processing the data was not required as confirmed by the license attached with the open access data. The data analyzes that we performed are in compliance with relevant local laws and institutional guidelines. No funding was received for conducting this study.

Data and Code Availability

The datasets used in this paper are publicly available; Dataset-1: `www.bbci.de/competition/iii/`; Dataset-2: `https://physionet.org/content/eegmmidb/1.0.0/`. We customized four MATLAB toolboxes to implement various stages of our proposed approach. Specifically, EEGLAB [79] for EEG preprocessing and scalp projection, Brainstorm [80] for brain projection, GSPBOX [81] for generating the graph figures, and the Brain Connectivity Toolbox (BCT) [82] for calculating graph metrics.

Acknowledgments

Dataset-1 used in this work was provided as part of the BCI Competition III, organized by the Berlin Brain-Computer Interface team (`https://www.bbci.de/`). Dataset-2 was made available by the developers of the BCI2000 instrumentation system for brain-computer interface research (`https://www.bci2000.org/downloads/doc/paper.pdf`).

References

[1] G. Buzsáki, C. A. Anastassiou, and C. Koch, "The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes," *Nature reviews neuroscience*, vol. 13, no. 6, pp. 407-420, 2012.

[2] C. M. Michel, and M. M. Murray, "Towards the utilization of EEG as a brain imaging tool," *Neuroimage*, vol. 61, no. 2, pp. 371-385, 2012.

[3] O. Sporns, and J. D. Zwi, "The small world of the cerebral cortex," *Neuroinformatics*, vol. 2, no. 2, pp. 145-162, 2004.

[4] O. Sporns, "Networks of the Brain," MIT Press, 2016.

[5] D. S. Bassett, and O. Sporns, "Network neuroscience," *Nature neuroscience*, vol. 20, no. 3, pp. 353-364, 2017.

[6] S. Sardellitti, S. Barbarossa, and P. Di Lorenzo, "Graph topology inference based on sparsifying transform learning," *IEEE Transactions On Signal Processing*, vol. 67, no. 7, pp. 1712-1727, 2019.

[7] L. Stanković, and E. Sejdić, "Vertex-frequency analysis of graph signals," ed. Berlin, Germany: Springer International Publishing, 2019.

[8] T. Adamovich, I. Zakharov, A. Tabueva, and S. Malykh, "The thresholding problem and variability in the EEG graph network parameters," *Scientific Reports*, vol. 12, no. 1, pp. 18659, 2022.

[9] J. Jin *et al.*, "A novel classification framework using the graph representations of electroencephalogram for motor imagery based brain-computer interface," *IEEE Transactions On Neural Systems And Rehabilitation Engineering*, vol. 30, pp. 20-29, 2021.

[10] S. Pain, M. Sarma, and D. Samanta, "Graph signal processing and graph learning approaches to Schizophrenia pattern identification in brain Electroencephalogram," *Biomedical Signal Processing And Control*, vol. 100, pp. 106954, 2025.

[11] R. F. Betzel, and D. S. Bassett, "Multi-scale brain networks," *Neuroimage*, vol. 160, pp. 73-83, 2017.

[12] A. Avena-Koenigsberger, B. Misic, and O. Sporns, "Communication dynamics in complex brain networks," *Nature reviews neuroscience*, vol. 19, no. 1, pp. 17-33, 2018.

[13] D. S. Bassett, and E. T. Bullmore, "Small-world brain networks," *The Neuroscientist*, vol. 12, no. 6, pp. 512-523, 2006.

[14] D. S. Bassett, and E. T. Bullmore, "Small-world brain networks revisited," *The Neuroscientist*, vol. 23, no. 5, pp. 499-516, 2017.

[15] M. P. van den Heuvel, C. J. Stam, M. Boersma, and H. H. Pol, "Small-world and scale-free organization of voxel-based resting-state functional connectivity in the human brain," *Neuroimage*, vol. 43, no. 3, pp. 528-539, 2008.

[16] C. Forlim, S. Haghiri, S. Düzel, and S. Kühn, "Efficient small-world and scale-free functional brain networks at rest using k-nearest neighbors thresholding," *BioRxiv*, pp. 628453, 2019. https://doi.org/10.1101/628453

[17] B. C. M. Van Wijk, C. J. Stam, and A. Daffertshofer, "Comparing brain networks of different size and connectivity density using graph theory," *PloS One*, vol. 5, no. 10, pp. e13701, 2010.

[18] G. Pfurtscheller, and C. Neuper, "Motor imagery and direct brain-computer communication," *Proceedings Of The IEEE*, vol. 89, pp. 1123-1134, 2001.

[19] C. Neuper, R. Scherer, S. Wriessnegger, and G. Pfurtscheller, "Motor imagery and action observation: modulation of sensorimotor brain rhythms during mental control of a brain–computer interface," *Clinical Neurophysiology*, vol. 120, no. 2, pp. 239-247, 2009.

[20] H. Yuan, and B. He, "Brain–computer interfaces using sensorimotor rhythms: current state and future perspectives," *IEEE Transactions On Biomedical Engineering*, vol. 61, no. 5, pp. 1425-1435, 2014.

[21] E. S. Finn *et al.*, "Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity," *Nature neuroscience*, vol. 18, no. 11, pp. 1664–1671, 2015.

[22] E. Amico, and J. Goñi, "The quest for identifiability in human functional connectomes," *Scientific Reports*, vol. 8, no. 1, pp. 8254, 2018.

[23] M. Demuru, and M. Fraschini, "EEG fingerprinting: Subject-specific signature based on the aperiodic component of power spectrum," *Computers In Biology And Medicine*, vol. 120, pp. 103748, 2020.

[24] I. Hameed, D. M. Khan, S. M. Ahmed, S. S. Aftab, and H. Fazal, "Enhancing motor imagery EEG signal decoding through machine learning: A systematic review of recent progress," *Computers in Biology and Medicine*, vol. 185, p. 109534, 2025.

[25] V. Abootalebi, M. Moradi, and M. Khalilzadeh, "A new approach for EEG feature extraction in P300-based lie detection," *Computer Methods And Programs In Biomedicine*, vol. 94, no. 1, pp. 48-57, 2009.

[26] F. Riaz, A. Hassan, S. Rehman, I. Niazi, and K. Dremstrup, "EMD-based temporal and spectral features for the classification of EEG signals using supervised learning," *IEEE Transactions On Neural Systems And Rehabilitation Engineering*, vol. 24, no. 1, pp. 28-35, 2015.

[27] B. Gosala, P.D. Kapgate, P. Jain, R.N. Chaurasia, and M. Gupta, "Wavelet transforms for feature engineering in EEG data processing: An application on Schizophrenia," *Biomedical Signal Processing and Control*, vol. 85, pp. 104811, 2023.

[28] H. Ramoser, J. Muller-Gerking, and G. Pfurtscheller, "Optimal spatial filtering of single trial EEG during imagined hand movement," *IEEE Transactions On Rehabilitation Engineering*, vol. 8, no. 4, pp. 441-446, 2000.

[29] F. Lotte, and C. Guan, "Regularizing common spatial patterns to improve BCI designs: unified theory and new algorithms," *IEEE Transactions On Biomedical Engineering*, vol. 58, no. 2, pp. 355-362, 2010.

[30] A. Ortega, P. Frossard, J. Kovačević, J. M. Moura and P. Vandergheynst, "Graph signal processing: Overview, challenges, and applications," in *Proceedings of the IEEE*, vol. 106, no. 5, pp. 808-828, 2018.

[31] I. Gutman, "The energy of a graph: old and new results," *Algebraic Combinatorics And Applications*, pp. 196-211, 2001.

[32] I. Gutman, and B. Zhou, "Laplacian energy of a graph," *Linear Algebra And Its Applications*, vol. 414, no. 1, pp. 29-37, 2006.

[33] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE Signal Processing Magazine*, vol. 30, no. 3, pp. 83-98, 2013.

[34] H. Behjat, U. Richter, D. Van De Ville, and L. Sörnmo, "Signal-adapted tight frames on graphs," *IEEE Transactions on Signal Processing*, vol. 64, no. 22, pp. 6017–6029, 2016.

[35] A. Sandryhaila and J. M. Moura, "Discrete signal processing on graphs: Frequency analysis," *IEEE Transactions On Signal Processing*, vol. 62, no. 12, pp. 3042-3054, 2014.

[36] H. B. Ahmed, D. Dare, and A. O. Boudraa, "Graph signals classification using total variation and graph energy informations," *2017 IEEE Global Conference On Signal And Information Processing (GlobalSIP)*, pp. 667-671, 2017.

[37] P. Mathur, and V. K. Chakka, "Graph signal processing based cross-subject mental task classification using multi-channel EEG signals," *IEEE Sensors Journal*, vol. 22, no. 8, pp. 7971-7978, 2022.

[38] E. T. Bullmore, and O. Sporns, "Complex brain networks: graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 3, pp. 186-198, 2009.

[39] P. Erdős, and A. Rényi, "On the evolution of random graphs," *Publ. Math. Inst. Hung. Acad. Sci*, vol. 5, no. 1, pp. 17-60, 1960.

[40] C. J. Stam, and J. C. Reijneveld, "Graph theoretical analysis of complex networks in the brain," *Nonlinear Biomedical Physics*, vol. 1, no. 1, pp. 1-19, 2007.

[41] V. Latora, V. Nicosia, and G. Russo, "Complex networks: principles, methods and applications," Cambridge University Press, 2017.

[42] M. E. J. Newman, "Networks," Oxford University Press, 2018.

[43] A.-L. Barabási, and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509-512, 1999.

[44] A.-L. Barabási, "Linked: The new science of networks," American Association of Physics Teachers, 2003.

[45] D. J. Watts, and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440-442, 1998.

[46] M. D. Humphries, and K. Gurney, "Network 'small-world-ness': a quantitative method for determining canonical network equivalence," *PloS One*, vol. 3, no. 4, pp. e0002051, 2008.

https://doi.org/10.1371/journal.pone.0002051.

[47] S. F. Muldoon, E. W. Bridgeford, and D. S. Bassett, "Small-world propensity and weighted brain networks," *Scientific Reports*, vol. 6, no. 1, pp. 1-13, 2016.

[48] E. Rodriguez, N. George, J-P. Lachaux, J. Martinerie, B. Renault, and F. J. Varela, "Perception's shadow: long-distance synchronization of human brain activity," *Nature*, vol. 397, no. 6718, pp. 430-433, 1999.

[49] E. Gysels, and P. Celka, "Phase synchronization for the recognition of mental tasks in a brain-computer interface," *IEEE Transactions on neural systems and rehabilitation engineering*, vol. 12, no. 4, pp. 406-415, 2004.

[50] Z. Wang, Y. Tong, and X. Heng, "Phase-locking value based graph convolutional neural networks for emotion recognition," *IEEE Access*, vol. 7, no. 4, pp. 93711-93722, 2019.

[51] A. Gong, J. Liu, L. Lu, G. Wu, C. Jiang, and Y. Fu, "Characteristic differences between the brain networks of high-level shooting athletes and non-athletes calculated using the phase-locking value algorithm," *Biomedical Signal Processing And Control*, vol. 51, pp. 128-137, 2019.

[52] L. Marple, "Computing the discrete-time analytic signal via FFT," *IEEE Transactions on signal processing*, vol. 47, no. 9, pp. 2600-2603, 1999.

[53] S. Micheloyannis, E. Pachou, C. J. Stam, M. Vourkas, S. Erimaki, and V. Tsirka, "Using graph theoretical analysis of multi-channel EEG to evaluate the neural efficiency hypothesis," *Neuroscience Letters*, vol. 402, no. 3, pp. 273-277, 2006.

[54] T. Behrouzi, and D. Hatzinakos, "Graph variational auto-encoder for deriving EEG-based graph embedding," *Pattern Recognition*, vol. 121, pp. 108202, 2022.

[55] F. Váša, and B. Mišić, "Null models in network neuroscience," *Nature Reviews Neuroscience*, vol. 23, no. 8, pp. 493-504, 2022.

[56] V. M. Eguiluz, D. R. Chialvo, G. A. Cecchi, M. Baliki, and A. V. Apkarian, "Scale-free brain functional networks," *Physical Review Letters*, vol. 94, no. 1, pp. 018102, 2005.

[57] A. Fornito, A. Zalesky, and E. T. Bullmore, "Fundamentals of brain network analysis," Academic Press, 2016.

[58] D. I. Shuman, M. J. Faraji, and P. Vandergheynst, "A multiscale pyramid transform for graph signals," *IEEE Transactions on Signal Processing*, vol. 64, no. 8, pp. 2119-2134, 2015. https://doi.org/10.1109/TSP.2015.2512529.

[59] A. Loukas, "Graph Reduction with Spectral and Cut Guarantees," *Journal of Machine Learning Research*, vol. 20, no. 116, pp. 1-42, 2019.

[60] M. D. Khalili, V. Abootalebi, and H. Saeedi-Sourck, "A Dimensionality Reduction Approach for Motor Imagery BCI using Functional Clustering, Graph Signal Processing, and Differential Evolution," bioRxiv, 2023. https://www.biorxiv.org/content/10.1101/2023.08.28.555094v1. https://doi.org/10.1101/2023.08.28.555094.

[61] M. Miri, V. Abootalebi, E. Amico, H. Saeedi-Sourck, D. Van De Ville, and H. Behjat, "Brain fingerprinting using EEG graph inference," *2023 31st European Signal Processing Conference (EUSIPCO)*, IEEE, pp. 1025-1029, 2023. https://doi.org/10.23919/EUSIPCO58844.2023.10289864

[62] A. Griffa, E. Amico, R. Liégeois, D. Van De Ville, and M. Preti, "Brain structure-function coupling provides signatures for task decoding and individual fingerprinting," *NeuroImage*, vol. 250, pp. 118970, 2022.

[63] I. Voitalov, P. Van Der Hoorn, R. Van Der Hofstad, and D. Krioukov, "Scale-free networks well done," *Physical Review Research*, vol. 1, no. 3, pp. 033034, 2019.

[64] K. R. Müller and B. Blankertz, "Berlin Institute of Technology, Charité-University Medicine Berlin," Available: https://www.bbci.de/competition/iii/desc_IVa.html, Accessed: 03.03.2023 [Online].

[65] B. Blankertz *et al.*, "The BCI competition III: validating alternative approaches to actual BCI problems," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 14, no. 2, pp. 153-159, June 2006.
https://doi.org/10.1109/TNSRE.2006.875642.

[66] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000: a general-purpose brain-computer interface (BCI) system," *IEEE Transactions on biomedical engineering*, vol. 51, no. 6, pp. 1034–1043, 2004.

[67] G. Schalk, D. J. McFarland, T. Hinterberger, N. Birbaumer, and J. R. Wolpaw, "BCI2000," Available: https://physionet.org/content/eegmmidb/1.0.0/, Accessed: 03.03.2023 [Online].

[68] A. L. Goldberger *et al.*, "PhysioBank, PhysioToolkit, and PhysioNet: components of a new research resource for complex physiologic signals," *circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

[69] M. D. Khalili, V. Abootalebi, and H. Saeedi-Sourck, "Dimensionality Reduction for Motor Imagery BCI System Using Kron Reduction, Graph Fourier Transform and Differential Evolution," *Iranian Journal of Biomedical Engineering*, vol. 16, no. 1, pp. 75-94, 2022. ([Published in Persian]).

[70] M. Miri, V. Abootalebi, H. Saeedi-Sourck, D. Van De Ville, and H. Behjat, "Spectral Representation of EEG Data using Learned Graphs with Application to Motor Imagery Decoding," *Biomedical Signal Processing and Control*, vol. 87, pp. 105537, 2024.

[71] F. Lotte *et al.*, "A review of classification algorithms for EEG-based brain–computer interfaces: a 10-year update," *Journal Of Neural Engineering*, vol. 15, no. 3, pp. 031005, 2018. https://doi.org/10.1088/1741-2552/aab2f2.

[72] N. Padfield, J. Zabalza, H. Zhao, V. Masero, and J. Ren, "EEG-based brain-computer interfaces using motor-imagery: Techniques and challenges," *Sensors*, vol. 19, no. 6, pp. 1423, March 2019. https://doi.org/10.3390/s19061423.

[73] Z. Khademi, F. Ebrahimi, and H. M. Kordy, "A transfer learning-based CNN and LSTM hybrid deep learning model to classify motor imagery EEG signals," *Computers in Biology and Medicine*, vol. 143, p. 105288, 2022.

[74] M. M. Abdelaty, M. A. Rushdi, M. E. Rasmy, and M. H. Annaby, "Graph vertex and spectral features for EEG-based motor imagery classification," *Computers in Biology and Medicine*, vol. 189, p. 109944, 2025.

[75] S. Das, and P. N. Suganthan, "Differential evolution: A survey of the state-of-the-art" *IEEE Transactions On Evolutionary Computation*, vol. 15, no. 1, pp. 4-31, 2010.

[76] M. Z. Baig, N. Aslam, H. P. Shum, and L. Zhang, "Differential evolution algorithm as a tool for optimal feature subset selection in motor imagery EEG," *Expert Systems with Applications*, vol. 90, pp. 184-195, 2017. https://doi.org/10.1016/j.eswa.2017.07.033.

[77] E. Sareen, S. Zahar, D. Van De Ville, A. Gupta, A. Griffa, and E. Amico, "Exploring MEG brain fingerprints: Evaluation, pitfalls, and interpretations," *NeuroImage*, vol. 240, pp. 118331, 2021.

[78] P. Sorrentino *et al.*, "Clinical connectome fingerprints of cognitive decline," *NeuroImage*, vol. 238, pp. 118253, 2021.

[79] A. Delorme, and M. Scott, "EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis," *Journal of neuroscience methods*, vol. 134, no. 1, pp. 9-21, 2004.

[80] F. Tadel, S. Baillet, J. C. Mosher, D. Pantazis, and R. M. Leahy, "Brainstorm: a user-friendly application for MEG/EEG analysis," *Computational intelligence and neuroscience*, vol. 2011, no. 8, pp. 1-13, 2011.

[81] N. Perraudin *et al.*, "GSPBOX: A toolbox for signal processing on graphs," *arXiv* preprint arXiv:1408.5781, 2014.

[82] M. Rubinov, and O. Sporns, "Complex network measures of brain connectivity: uses and interpretations," *Neuroimage*, vol. 52, no. 3, pp. 1059-1069, 2010.

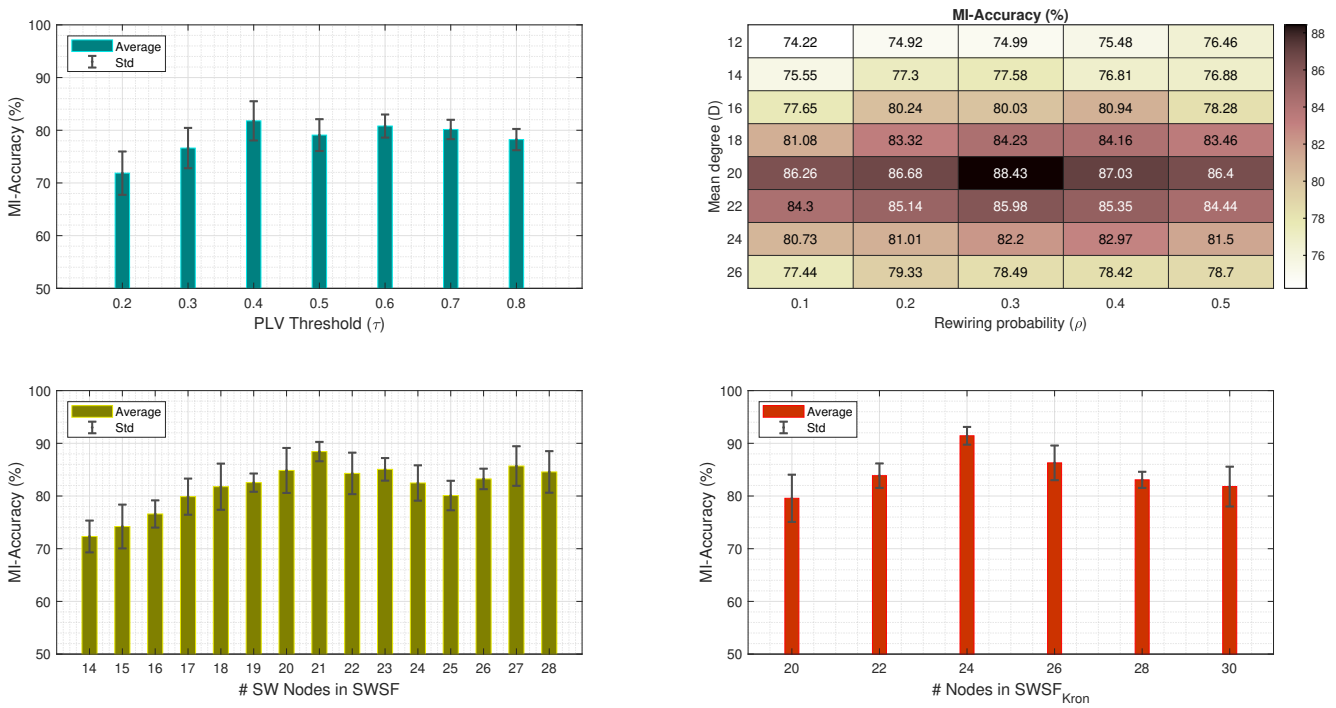

Fig. S1: Cross-validation strategy for parameter selection. To ensure generalization across tasks, a systematic 10-fold cross-validation (CV) was applied to both motor imagery (MI) decoding (Dataset-1) and brain fingerprinting (Dataset-2). Five parameters were optimized: PLV threshold (τ), rewiring probability (ρ), average degree (D), initial seed vertices (V_{seed}), and selected vertices for Kron reduction. For MI decoding, τ was tuned independently, while ρ and D were optimized jointly via grid search. V_{seed} and Kron-reduction nodes were selected using a hybrid strategy combining functional clustering with degree centrality, followed by empirical evaluation to identify subsets maximizing validation accuracy. (Top-left) shows performance across τ values; (Top-right) shows joint tuning of ρ and $\bar{D}$. (Bottom-left) explores varying the number of V_{seed} nodes in the SWSF graph; (Bottom-right) shows accuracy trends for different node selections in the Kron-reduced graphs. For brain fingerprinting, ρ and D were re-optimized using CV, while τ, V_{seed}, and Kron nodes were transferred from the MI task. Biological plausibility constraints ($I > 1$, $P > 0.6$) were enforced to ensure small-world structure.

Table S1: Classification accuracy in MI decoding using small-world brain graphs derived by using different values of the free parameter ρ.

Subjects	**SW(ρ = 0.1)**	**SW(ρ = 0.2)**	**SW(ρ = 0.3)**	**SW (ρ = 0.4)**	**SW(ρ = 0.5)**
aa	74.29±1.89	78.57±2.64	82.14±2.43	80.36±2.26	75.00±2.05
al	83.93±1.94	89.64±1.44	95.00±2.17	91.79±2.54	84.64±1.89
av	75.71±2.17	77.86±2.21	78.93±2.29	76.43±2.34	75.36±2.41
aw	78.21±2.18	82.14±2.57	88.57±2.15	84.29±2.27	81.79±1.84
ay	76.43±2.26	81.43±2.08	88.57±1.99	83.92±2.39	77.86±2.28
Mean±std	77.71±2.09	81.93±2.19	86.64±2.21	83.36±2.36	78.93±2.09

Table S2: Graph metric values used to validate the small-world characteristics of the constructed small-world brain graphs. For all graphs and subjects: $N = 118$, $M = 1180$, $D = 20$, $\bar{E} = 0.17$, and $\rho = 0.3$.

Metric	**aa**	**al**	**av**	**aw**	**ay**	**Average**
C	0.361	0.384	0.369	0.372	0.377	0.373
L	2.088	2.145	2.072	2.168	2.123	2.119
I	2.047	1.963	2.028	2.008	1.969	2.003
P	0.785	0.769	0.781	0.776	0.764	0.775

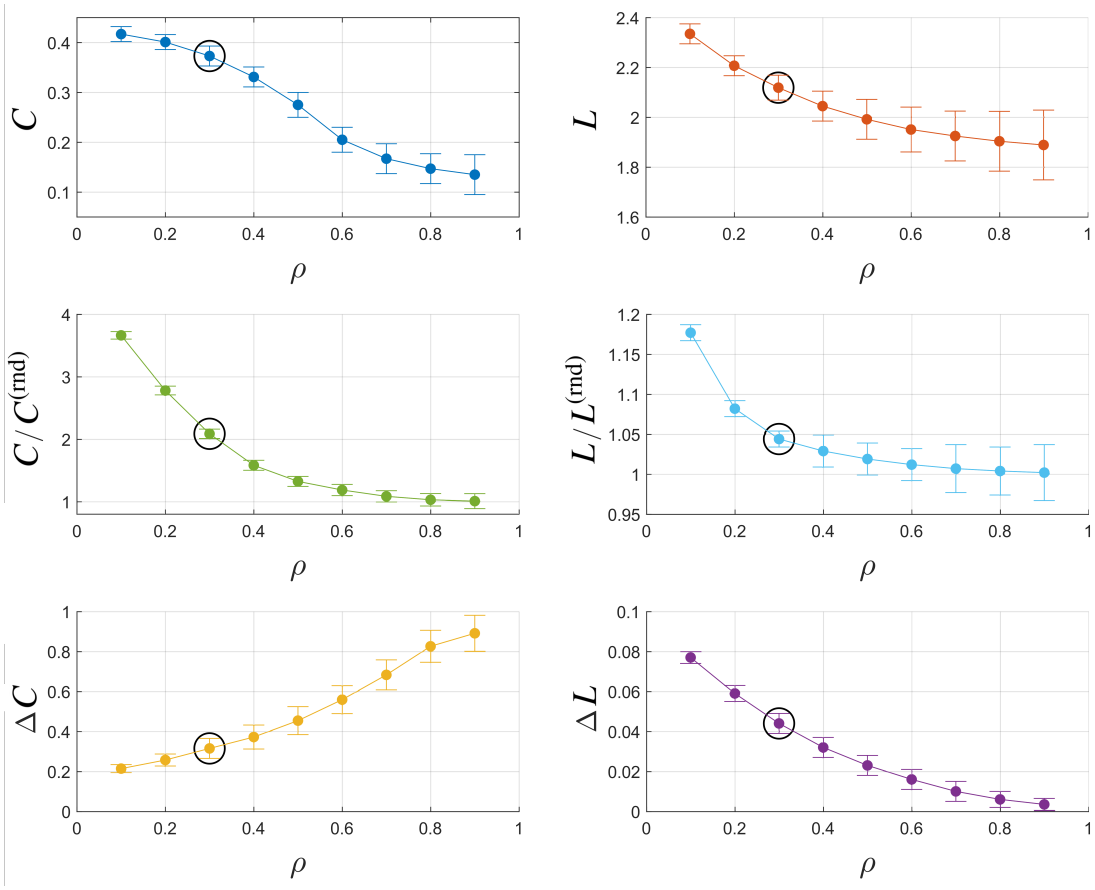

Fig. S2: Effect of ρ on C, L, $C/C^{(\mathrm{rnd})}$, $L/L^{(\mathrm{rnd})}$, ΔC, and ΔL for SW brain graphs over all subjects in MI decoding.

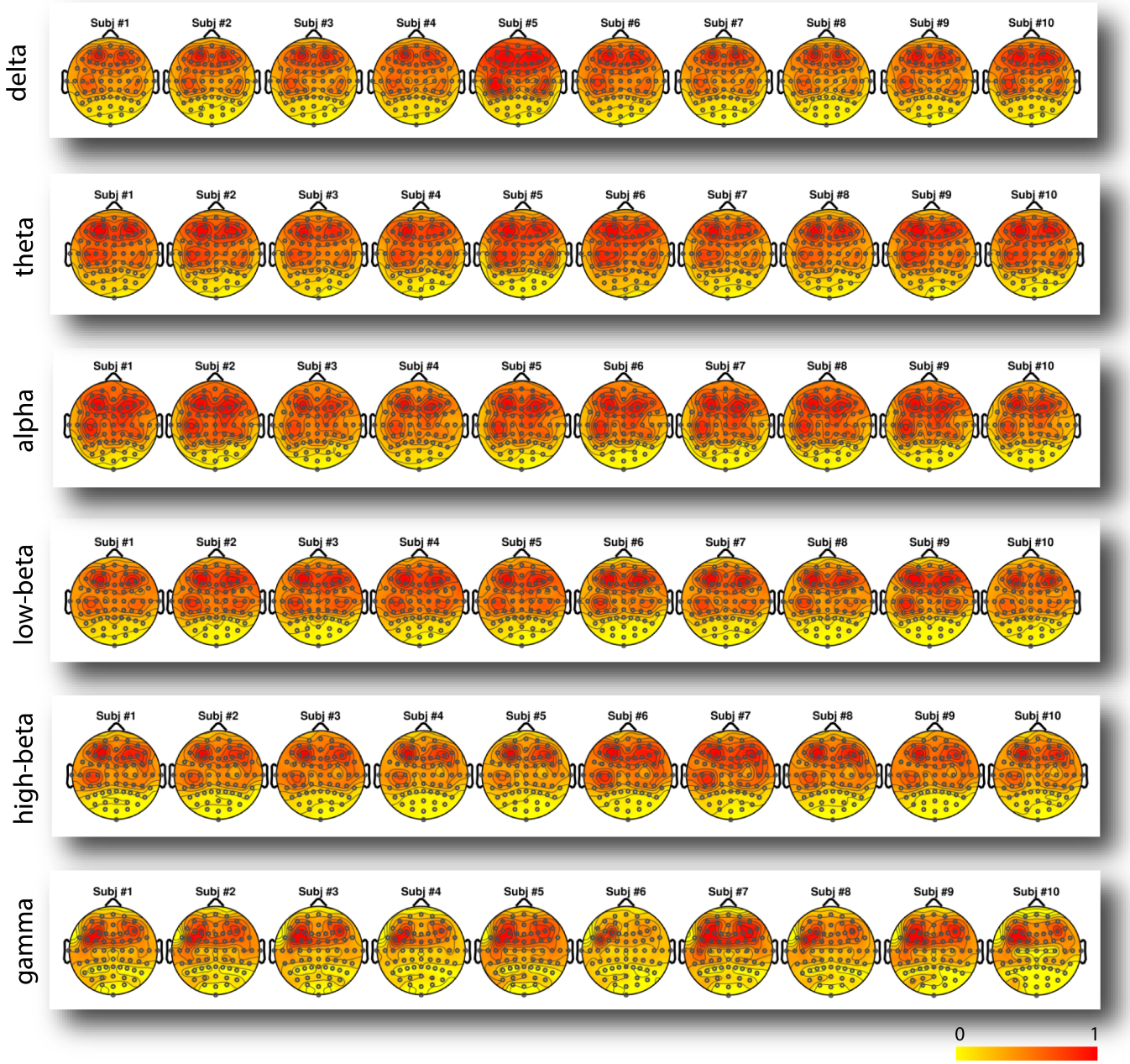

Fig. S3: Topographic plots showing electrode significance for brain fingerprinting across the first 10 subjects of Dataset-2, displayed by frequency band.

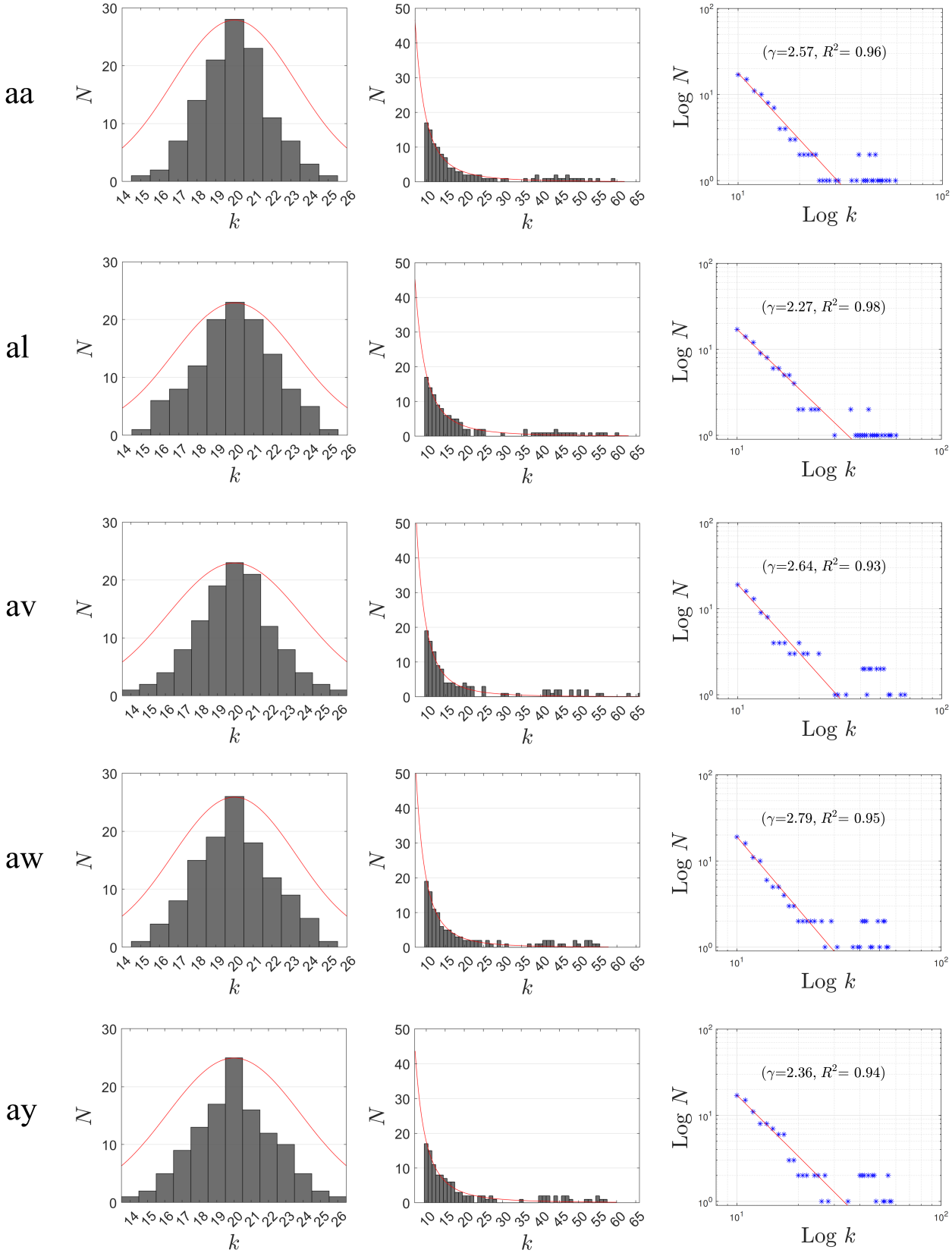

Fig. S4. Left: Histogram of small-world degree distribution with their corresponding fitted Normal distribution. Center: Histogram of scale-free degree distribution (center) with fitted power-law curves, exhibiting exponents between 2 and 3. Right: log-log plots highlighting scale-free properties, with corresponding γ values and R^2 scores shown. The number of hubs (N_h) is 21 for aa, al, av, aw and 20 for ay, based on 118 electrodes, 1180 edges, and $D = 20$. The average power-law exponent across subjects is 2.53 with $R^2 \approx 0.95$. Hubs are defined as nodes with degrees at least 1.5 times the mean, averaging 21 hubs (18% of nodes) per subject.

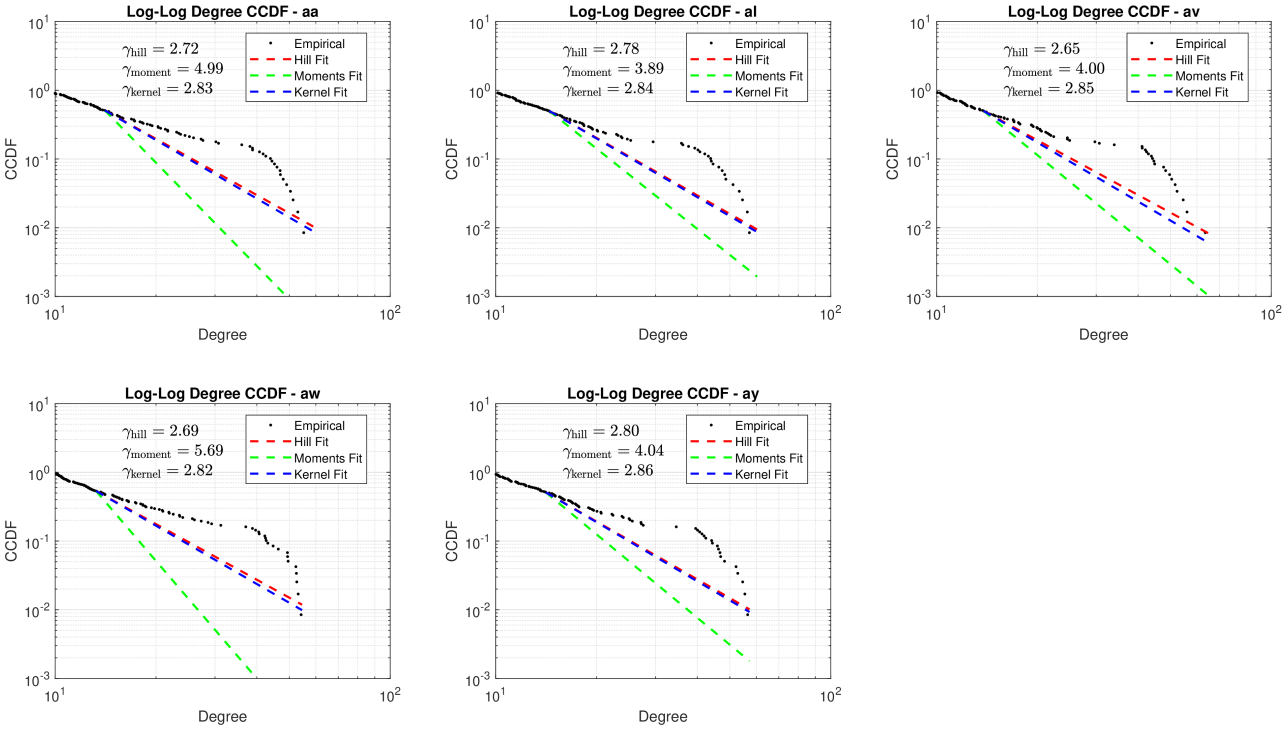

Fig. S5: Log-Log CCDFs of degree distributions for five subjects in MI dataset, overlaid with tail exponent estimates from three EVT-based methods: Hill (red), Moment (green), and Kernel (blue). Estimated γ values reflect heavy-tailed, scale-free–like structure in brain graphs.

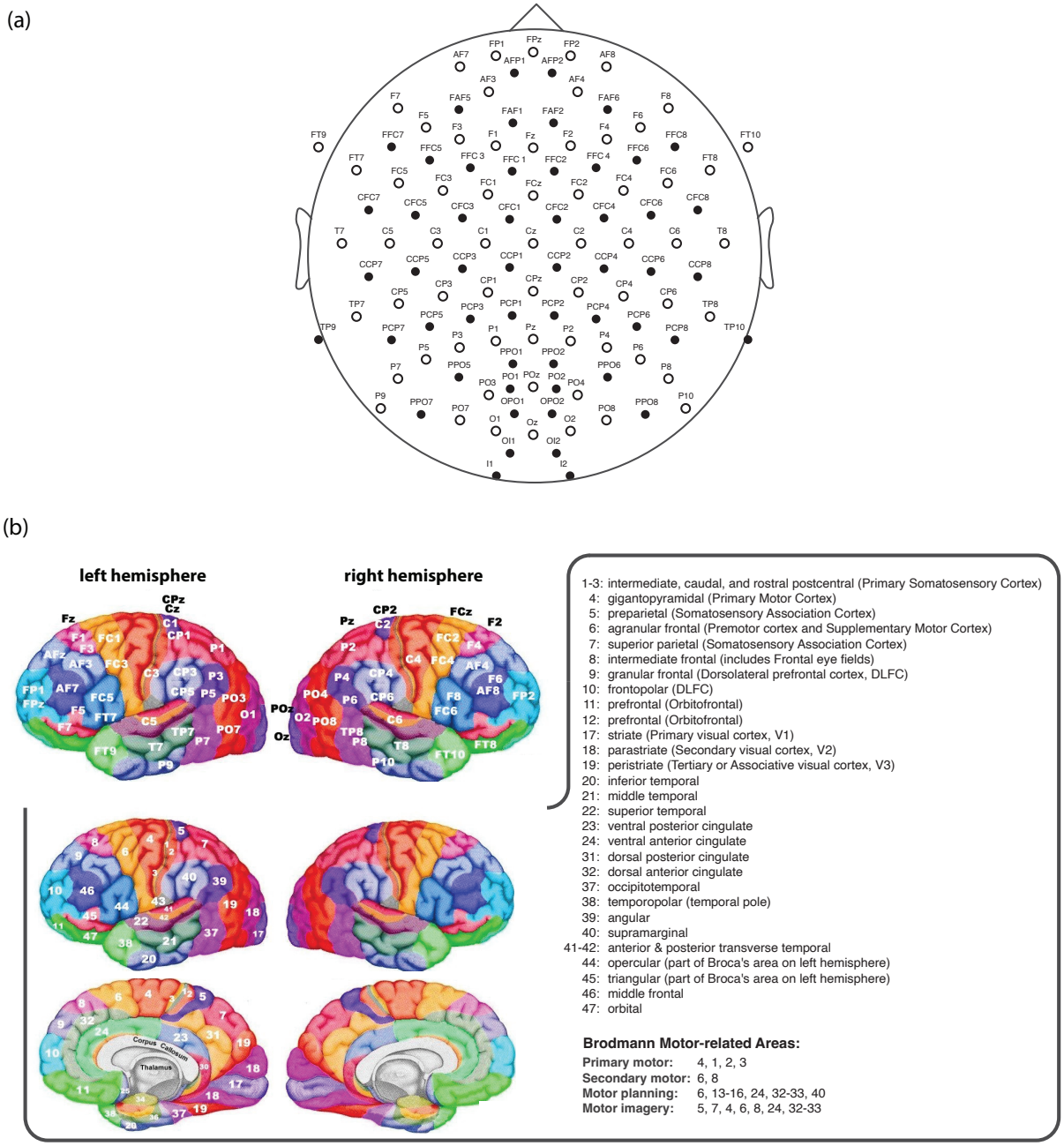

Fig. S6: (a) Arrangement of 118 EEG electrodes. Labels, corresponding to cortical locations: frontal (F), temporal (T), central (C), parietal (P), occipital (O), Inion (I). The letter 'z' indicates the midline and letter combination refer to intermediate cortical regions. Odd and even numbers indicate the left and right hemisphere, respectively. (b) The position of a subset (64) of the electrodes (marked as white in (a)) are shown on a rendering of the cerebral cortex (top left) relative to Brodmann area centers (bottom left). The surface maps were adopted from `http://www.brainm.com/software/pubs/dg/BA_10-20_ROI_Talairach/nearesteeg.html`.

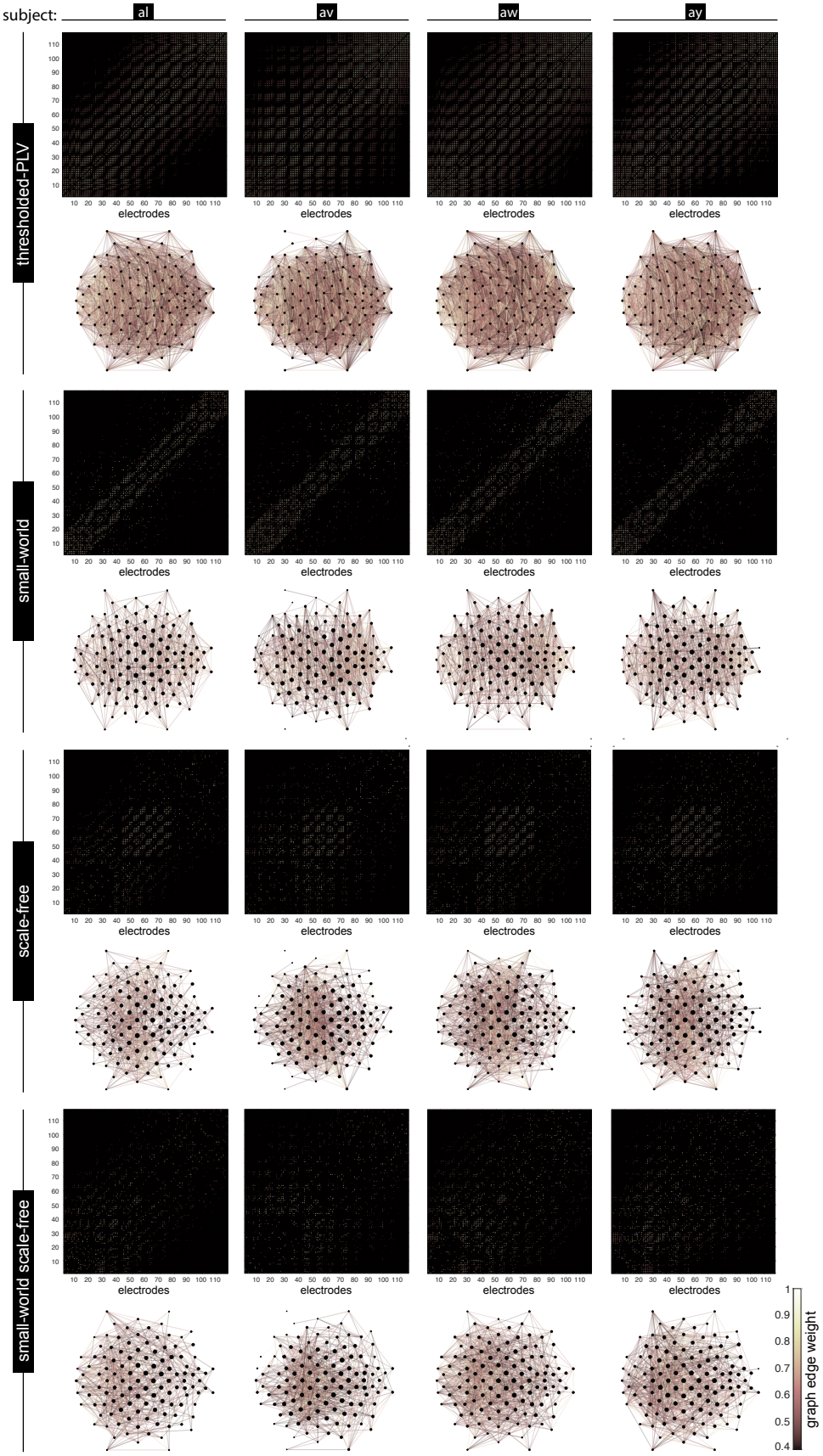

Fig. S7: Brain graphs derived from PLV (thresholded), SW, SF, and SWSF models, along with their corresponding weighted adjacency matrices shown for subjects al, av, aw, and ay.

Table S3: Description of classical metric considered in the bag of features used for MI decoding. A subset of these classical features in combination with GSP features were selected (cf. Section 3.2) through differential evolution (cf. Section 2.5) for MI decoding. ($\mathbf{x}$ = denotes the timeseries associated to an electrode.)

Group	Name	Calculation
Statistical features	Maximum	$Max(\mathbf{x}) = max(\mathbf{x}_t)$
	Minimum	$Min(\mathbf{x}) = min(\mathbf{x}_t)$
	First difference	$D_1(\mathbf{x}) = \frac{1}{T-1} \sum_{t=1}^{T-1} \mid \mathbf{x}_{t+1} - \mathbf{x}_t \mid$
	Second difference	$D_2(\mathbf{x}) = \frac{1}{T-2} \sum_{t=1}^{T-2} \mid \mathbf{x}_{t+2} - \mathbf{x}_t \mid$
	Arithmetic mean	$\mu_{\mathbf{x}} = \frac{1}{T} \sum_{t=1}^{T} \mathbf{x}_t$
	Standard deviation	$\sigma_{\mathbf{x}} = \sqrt{\frac{1}{T-1} \sum_{t=1}^{T} (\mathbf{x}_t - \mu_{\mathbf{x}})^2}$
	Skewness	$Sk(\mathbf{x}) = \frac{1}{(T-1)\sigma_{\mathbf{x}}^3} \sum_{t=1}^{T} (\mathbf{x}_t - \mu_{\mathbf{x}})^3$
	Kurtosis	$Kt(\mathbf{x}) = \frac{1}{(T-1)\sigma_{\mathbf{x}}^4} \sum_{t=1}^{T} (\mathbf{x}_t - \mu_{\mathbf{x}})^4$
Hjorth features	Hjorth activity	$HA_{\mathbf{x}} = \sigma_{\mathbf{x}}^2$
	Hjorth mobility	$HM_{\mathbf{x}} = \frac{\sigma_{\mathbf{x}'}}{\sigma_{\mathbf{x}}};\ \mathbf{x}' = \frac{d}{dx}$
	Hjorth complexity	$HC_{\mathbf{x}} = \frac{\sigma_{\mathbf{x}''}/\sigma_{\mathbf{x}'}}{\sigma_{\mathbf{x}'}/\sigma_{\mathbf{x}}};\ \mathbf{x}'' = \frac{d^2}{dx^2}$
Variation features	Log root sum of sequential variations	$LRSSV(\mathbf{x}) = log_{10} \sqrt{\sum_{t=1}^{T-1} (\mathbf{x}_t - \mathbf{x}_{t-1})^2}$
Energy features	Mean energy	$ME(\mathbf{x}) = \frac{1}{T} \sum_{t=1}^{T} \mathbf{x}_t^2$
	Mean Teager energy	$MTE(\mathbf{x}) = \frac{1}{T} \sum_{t=3}^{T} \left(\mathbf{x}_{t-1}^2 - \mathbf{x}_t\, \mathbf{x}_{t-2}\right)$
Entropy features	Shannon entropy	$ShEn = -\sum_{i=1}^{n} P_i\, log_2(P_i)$; $[P_i$: Probability$]$
	Rényi entropy	$ReEn^{(\xi)} = \frac{1}{1-\xi} \left(log_2 \sum_{i=1}^{n} (P_i^{\xi})\right)$; $[\xi \neq 1,$ (Here: $\xi = 2)]$
	Tsallis entropy	$TsEn^{(\xi)} = \frac{1}{\xi-1}(1 - \sum_{i=1}^{n} P_i^{\xi})$; $[\xi \geq 0,\ \xi \neq 1,$ (Here: $\xi = 2)]$
	Log energy entropy	$LeEn = \sum_{t=1}^{T} log(\mathbf{x}_t^2)$
Band power features	Mu band power	$BP_{\mathbf{x},\mu} = \sum_{i=1}^{N} PSD_{\mathbf{x},\mu}(i)$; ($PSD$ = Power Spectral Density)
	Beta band power	$BP_{\mathbf{x},\beta} = \sum_{i=1}^{N} PSD_{\mathbf{x},\beta}(i)$; ($PSD$ = Power Spectral Density)
	Beta/Mu power ratio	$BPR_{\mathbf{x},\beta/\mu} = \frac{BP_{\mathbf{x},\beta}}{BP_{\mathbf{x},\mu}}$
Time series features	Coefficients of AR model (order 4)	$AR(4) : \mathbf{x}_t = \sum_{j=1}^{4} \phi_j \mathbf{x}_{t-j} + \epsilon_t$; $[\phi_j : j^{th}$ Coefficient, $\epsilon_t \sim N(0, \sigma^2)]$
CSP-based features	Log variance ratio of the CSP filtered signal	$CSP_{logv} = log\left[\frac{var(W_k X)}{\sum_{i=1}^{4} var(W_i X)}\right]$; ($W_i$ = CSP filters)