PAPER • OPEN ACCESS

From single layer to multilayer networks in mild cognitive impairment and Alzheimer's disease

View the article online for updates and enhancements.

Journal of Physics: Complexity

PAPER

From single layer to multilayer networks in mild cognitive impairment and Alzheimer's disease

Ignacio Echegoyen[1,2,3,*] [ID], David López-Sanz[4,5], Fernando Maestú[4,6,7] and Javier M Buldú[1,2,8]

1 Complex Systems Group & GISC, Universidad Rey Juan Carlos, Madrid, 28933, Móstoles, Madrid, Spain
2 Laboratory of Biological Networks, Center for Biomedical Technology, Universidad Politécnica de Madrid, 28223 Madrid, Spain
3 Department of Psychology, Universidad Pontificia Comillas, 28049 Madrid, Spain
4 Laboratory of Cognitive and Computational Neuroscience, Centre for Biomedical Technology, Universidad Politécnica de Madrid (UPM), 28223 Madrid, Spain
5 Department of Psychobiology and Methodology in Behavioral Sciences, Universidad Complutense de Madrid, 28223 Pozuelo de Alarcón, Madrid, Spain
6 Department of Experimental Psychology, Universidad Complutense de Madrid, 28223 Pozuelo de Alarcón, Madrid, Spain
7 Research Networking Center in Bioengineering, Biomaterials and Nanomedicine, 28029 Zaragoza, Spain
8 Institute of Unmanned System and Center for OPTical IMagery Analysis and Learning (OPTIMAL), Northwestern Polytechnical University, Xi'an 710072, People's Republic of China
* Author to whom any correspondence should be addressed.

E-mail: i.echegoyen@comillas.edu

Abstract

We investigate the alterations of functional networks of patients suffering from mild cognitive impairment and Alzheimer's disease (AD) when compared to healthy individuals. Departing from the magnetoencephalographic recordings of these three groups, we construct and analyse the corresponding single layer functional networks at different frequency bands, both at the sensors and the regions of interest (ROI) levels. Different network parameters show statistically significant differences, with global efficiency being the one having the most pronounced differences between groups. Next, we extend the analyses to the frequency-band multilayer networks (MN) of the same dataset. Using the mutual information as a metric to evaluate the coordination between brain regions, we construct the $\alpha\beta$ MN and analyse their algebraic connectivity at baseline λ_{2-BSL} (i.e., the second smallest eigenvalue of the corresponding Laplacian matrices). We report statistically significant differences at the sensor level, despite the fact that these differences are not clearly observed when networks are obtained at the ROIs level (i.e., after a source reconstruction procedure). Next, we modify the weights of the inter-links of the multilayer network to identify the value of the algebraic connectivity λ_{2-T} leading to a transition where layers can be considered to be fully merged. However, differences between the values of λ_{2-T} of the three groups are not statistically significant. Finally, we developed nested multinomial logistic regression models (MNR models), with the aim of predicting group labels with the parameters extracted from the MN (λ_{2-BSL} and λ_{2-T}). Using these models, we are able to quantify how age influences the risk of suffering AD and how the algebraic connectivity of frequency-based multilayer functional networks could be used as a biomarker of AD in clinical contexts.

1. Introduction

Network Neuroscience [1] consists of the analysis of structural and functional networks coming from different brain recording modalities. The rationale behind this approach is based on the fact that the brain is a complex, nonlinear, dynamical system where neural populations are constantly coupling and decoupling over different spatial and temporal scales with the aim of transmitting and processing information. Importantly, brain functions emerge from these interactions, from basic biological regulation to highly complex cognitive abilities. Network neuroscience is a suitable tool to study these processes, where nodes can be single neurons

(at the microscale), cortical columns (mesoscale) or large neuronal populations (macroscale). Nowadays, we can track these interactions experimentally at different time scales, i.e. microseconds using electro- or magneto-encephalography (EEG and MEG, respectively) or seconds using, for example, functional magnetic resonance imaging (fMRI). This perspective has revolutionised the field of neuroscience, leading to great advances in the understanding of the healthy and diseased brain [2, 3]. The analyses of how functional brain networks organise to transmit and process information have shown an emerging hierarchical topology, with massively connected areas (hubs) and communities of highly connected nodes [4]. Under this new perspective, the picture drawn on different psychiatric disorders has changed dramatically, revealing important changes in the topological organisation of functional networks [3], and helping to develop new biomarkers based on network parameters, useful to classify patients and predict their evolution [5, 6].

Although the analysis of functional networks is still a hot topic and provides answers to many unresolved questions, most of the studies collapse all the information contained at different scales into one single network, having profound consequences on the conclusions one can draw [7]. In this sense, the multilayer formalism is more realistic to describe the transmission of information in the brain, whose general framework was described some years ago [8–10], along with the generalisation of some quantities from single layer networks to multilayer networks (MN) [11].

Functional brain networks are constantly changing over time. Tracking the evolution of the topology's dynamics across multiple scales during the execution of a task or following development and ageing is an interesting scientific quest, impossible to tackle with single layer networks, which are temporal snapshots of brain activity. To overcome this issue, some authors have proposed time-evolving multiplex networks, where layers represent whole-brain connectivity at different time steps [12, 13]. Recent generalisations of the modularity parameter Q to MN allow for graph partitions across layers [14], enabling the study of modules stability through different time scales (applied in [12, 13] during a learning task and training). However, how to quantify the weights of the past dynamics in a multilayer brain network is still an open question.

On the other hand, we must take into account that cross-frequency interactions are of major importance for information integration [15–18]. Fast oscillations allow precise timing of events in small neuronal populations, while slow ones recruit larger populations in long-range communication [18]. These interactions have been observed empirically in two distinct modes: phase-amplitude-coupling (PAC) and amplitude-envelope correlations (AEC) [15]. It is still a matter of debate if both occur at the same time and how they coordinate during different cognitive operations. In any case, it has been hypothesised that such mechanisms are behind information binding to form coherent cognitive percepts (for example, integrating different features of the same visual stimuli), movement coordination and memory consolidation. To study how different frequency bands interact, some authors propose the application of the multilayer framework, where each layer accounts for a different frequency band. Brookes *et al* [19] used this approach to analyse a visuomotor task recorded with MEG, studying differences between patients suffering from schizophrenia and healthy individuals. They measured the connectivity within layers with the envelope correlation and found that it correlates with the severity of the persistent symptoms. Yu *et al* [20], analysed MEG registers from subjects with Alzheimer's disease (AD) and controls, finding several 'multilayer hubs' affected by the disease that were not detected in the individual layers. These hubs (left hippocampus, posterior areas from the default mode network and occipital regions) showed different connectivity patterns and their vulnerability correlated with disturbances in the cognitive function and abnormal cerebrospinal fluid amyloid-β. Similarly, De Domenico *et al* [21] identified hubs in a frequency band multiplex network built from fMRI, improving notably the detection of schizophrenic patients. These findings reflect that multilayer hubs might not be the same as single layers hubs, drawing a much more complex picture of the functional role of highly connected nodes.

A novel question posed in frequency bands multilayer/multiplex networks concerns the level of interdependence between layers, which determines how information flows from any node to the rest of the network across layers. It can be studied in terms of the second smallest eigenvalue of the Laplacian matrix, λ_2, associated to the Fiedler vector. The λ_2 has been typically studied in networks to explore diffusive processes [22], synchronisation and robustness [23] (directly related to graph partitioning problems). It has been proven analytically that the behaviour of λ_2 in multiplex networks is a function of the number of identical, diagonal interlinks [24], enabling supperdiffusion (diffusion faster than that of any of the layers considered individually), at least for some range of the diffusive constant D [22]. This behaviour seems to be ubiquitous in MN, regardless of the topology of each individual network, at least for classical network science models (i.e., random regular, Barabási–Albert, Watts–Strogatz and lattices), as demonstrated numerically in [25]. In all cases, for two-layer multiplex networks with identical interlayer links, λ_2 follows two distinct regimes. While layers are sufficiently independent and act as distinct entities, the value of λ_2 is independent of the structure of the two layers. As more interlayer links are added, the network undergoes a sharp structural transition [24], and λ_2 stabilizes at $(1/2)(\mathbf{L}_A + \mathbf{L}_B)$, half the eigenvalue corresponding to the superposition of both layers

$(\lambda_S/2)$ [22]. In this regime, the network acts as a whole [24]. This finding enables new strategies to build networks maximising λ_2, as applied in [26], and can be useful in clinical contexts. With this regard, Tewarie *et al* analysed the λ_2 from resting state MEG recordings (healthy subjects) and compared the empirical results to those obtained from a biophysical model [27]. Interestingly, they conclude that the healthy brain works at the transition point between the aforementioned regimes, and hypothesise that such a state could promote layers integration and segregation processes. A more recent work has extended this notion to the case of MN with heterogeneous interlayer links, focussing on the impact of missing links on the value of the λ_2, and observing both regimes in synthetic signals and in real, resting state MEG registers [28].

In this paper, we analysed the single layer and the MN of patients suffering from mild cognitive impairment (MCI) and AD, and compared their structure with the equivalent networks obtained from healthy individuals. Using MEG recordings, we quantified, for every subject, the coordination between every pair of brain regions at the sensors and regions of interest (ROIs) level, but also at each frequency band and between frequency bands. To that end, we chose mutual information (MI) [29] as the connectivity metric. The reason behind this choice is the fact that, at the cost of no clear interpretation in physiological terms, MI is agnostic to the underlying statistical structure of the data. Hence, we can apply the same measure to quantify intra-layer and inter-layer links, ensuring both types of links are in the same range, and without any assumption regarding the mode of cross-frequency interaction (PAC/AEC). First, we obtained different network parameters (average shortest-path, weighted clustering coefficient, local and global efficiency and network outreach) at each individual frequency band (δ, θ, α and β), each of them leading to an independent single layer network. Next, we constructed the frequency-based MN by connecting the α and β layers through inter-layer links assessing the level of coordination between the two frequency bands. We investigated the role of λ_2 in the $\alpha\beta$ MN as a possible biomarker in AD and MCI. Following references [24, 28], we hypothesised that, as the condition advances, we might be able to observe differences in the interaction between frequency bands, especially in α and β bands, since (i) they are the ones accumulating more power in the spectrum, (ii) they have been shown to display significant differences in the statistical complexity and entropy [30], and (iii) they have been reported to be the frequency bands with the highest impairment [31].

2. Methods

2.1. Recruitment, data acquisition and preprocessing

We analyse resting-state MEG recording from 111 elder adults, recruited from the Center for Prevention of Cognitive Impairment, the Seniors Center of Chamartín District and the Hospital Clínico San Carlos Neurology Department in Madrid (Spain). The Hospital Clínico San Carlos Ethics Committee approved the study. Accordingly to their cognitive status, 48 participants were classified as healthy controls (control group (CG)), 46 met the criteria for amnestic MCI, and the remaining 17 received the diagnostic of dementia due to AD. The complete protocol, including diagnostic and excluding criteria, can be found in [30]. Tables 1–3 in the appendix (section 4) summarise demographic information and mini-mental state examination (MMSE) scores for each participant.

We recorded 4 min of resting state brain's magnetic activity through a 306 channel vectorview MEG system from each subject. Due to the tSSS filtering procedure, signals recorded by magnetometers and gradiometers are highly correlated and redundant [32]; thus, only data coming from the 102 magnetometers were employed for further analyses. Ocular, muscular and jump artefacts were removed automatically and visually confirmed later by an MEG expert. Additionally, ocular and electrocardiographic contributions to the MEG signal were discarded via independent component analyses. Clean data were segmented into 4 s epochs, and the resulting dataset comprised 29 epochs (samples) from each subject. Clean epochs for each subject were band-pass filtered between 1 and 45 Hz so that low frequency and power line artefact contributions were discarded. A realistic model using a single-shell was employed to generate the forward model of a 1 cm spacing grid with 2459 sources inside the brain. Magnetic activity at the brain level was reconstructed in each of the nodes by using linearly constrained minimum variance beamformer [33]. Once time series for all of the 2459 sources were calculated, a representative time-course for each cortical region of a reduced version of the Harvard–Oxford atlas [34] was chosen by selecting the centroid source of each region. We retained 62 ROIs at each sample for every subject. More details on the procedure can be found in [30].

Lastly, for all subjects in both modalities (sensors and ROIs), we conducted a frequency decomposition of all signals to get an estimate of each time series into four classical frequency bands: $\delta : 1$–4 Hz; $\theta : 4$–8 Hz; $\alpha : 8$–13 Hz and $\beta : 13$–31 Hz. In order to estimate the signal for each frequency band, we simply computed the Fourier transform of the signal and reconstructed it via inverse Fourier transform in the range of frequencies specified above. γ band was disregarded for all subjects due to lack of power in this specific frequency range (31–50 Hz).

2.2. Network construction

For each subject and sample in both modalities, we constructed two types of functional networks. On the one hand, four single-layer networks, one for each frequency band, and on the other, one $\alpha\beta$ multilayer network. To unify our analyses we estimated the functional connectivity in every network in terms of MI, either it be single layer (evaluated at each frequency band separately) or multilayer (evaluated in α and β bands and between them). MI works comparing the amount of shared information between any two variables. In our case, these two variables are time series (in sensors or ROIs space), either from the same layer (α or β), or from different layers (one from α and one from β). Being X and Y any two time series, the MI between them is defined as:

$$\text{MI}(X;Y) = \sum_{x \in X, y \in Y} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}, \tag{1}$$

where $p(x,y)$ stands for the joint probability of $X = x$ and $Y = y$, and $p(x)$, $p(y)$ their individual probabilities. To estimate probabilities we chose the Freedman–Diaconis rule [35], which performs reasonably good without assuming normality on the data [36].

Note that the multilayer network will be composed of two layers, α and β, and that we allow interlayer links (obtained through the MI) connecting any two nodes from any two layers. In this sense, our networks follow a complete multilayer architecture, not a multiplex one [28].

To test the statistical significance of the weight of every link, we applied an edge-level nonparametric scheme based on circular shifting, a method of the family of permutation testing. To that end, we generated new surrogate data under the null hypothesis of no temporal correlation [37] for each pair of time series considered [38]. We kept the original value of MI obtained, and then we followed an iterative procedure to test if it was statistically significant or not: (i) we kept one of the time series untouched, while in the other, (ii) we shifted a collection of consecutive points to its end, and (iii) we got the MI between both time series (the original and the shifted one). Repeating this process 200 times gave us an ad hoc null distribution of no temporal correlation for that particular pair of time series, against which we test our original estimation of MI. We set every comparison with a significance level of $\alpha' = 0.01$ and kept the original estimation of MI if and only if its probability of appearance in the null distribution was lesser than α'. Otherwise, we set it to zero. This method is preferred over other normalisation schemes due to its ability to control for multiple comparisons at any desired level of significance α', making no assumptions about the underlying distribution [37]. Also, circular shifting preserves the inner autocorrelation structure of the time series [38]. The length of each shifted segment l_s is adjusted for each frequency band across all subjects and samples, as the frequency component will determine the prototypical cycle of autocorrelation. In time points (tp), we chose $\delta = (1,4)$ Hz, $l_s = 40$ tp, $\theta = (4,8)$ Hz, $l_s = 20$ tp, $\alpha = (8,13)$ Hz, $l_s = 10$ tp and $\beta = (13,31)$ Hz, $l_s = 5$ tp.

In contrast to other synchronization measures, MI has no upper bound, potentially biasing any further comparison relying on the strength of the network. To avoid this issue, we normalised all the weights in every network by linearly rescaling it between 0 and 1:

$$p_{ij} = \frac{\text{MI}_{ij} - \min[\text{MI}_{ij}]}{\max[\text{MI}_{ij}] - \min[\text{MI}_{ij}]}, \tag{2}$$

where $\max[\text{MI}_{ij}](\min[\text{MI}_{ij}])$ indicates the largest (smallest) weight (i.e., MI) in the network. Note that this step is carried out for every network before any further calculation. Hence, it will differ depending on the network under consideration. In $\alpha\beta$ MN, we consider all links (inter and intralayer), while in single layer networks, the procedure is conducted over intralayer links (as layers are isolated).

2.3. Multilayer analyses

To study the dependency between layers, we first obtained the Laplacian matrix **L** from each $\alpha\beta$ multilayer network, whose elements are:

$$l_{ij} = \begin{cases} s_i & \text{if } i = j \\ -1 & \text{if } i \neq j \text{ and there is an edge from vertex } i \text{ to } j \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

where $s_i = \sum_{i \neq j} p_{ij}$ is the strength of a node i. Then, solving the eigenvector equation:

$$\mathbf{L}v = \lambda v \tag{4}$$

results in N eigenvalues and their associated eigenvectors, where N is the number of nodes in the network (62 ROIs $\times$ 2 frequency bands $\rightarrow N_{\text{ROI}} = 124$; 102 sensors $\times$ 2 frequency bands $\rightarrow N_{\text{sensor}} = 204$).

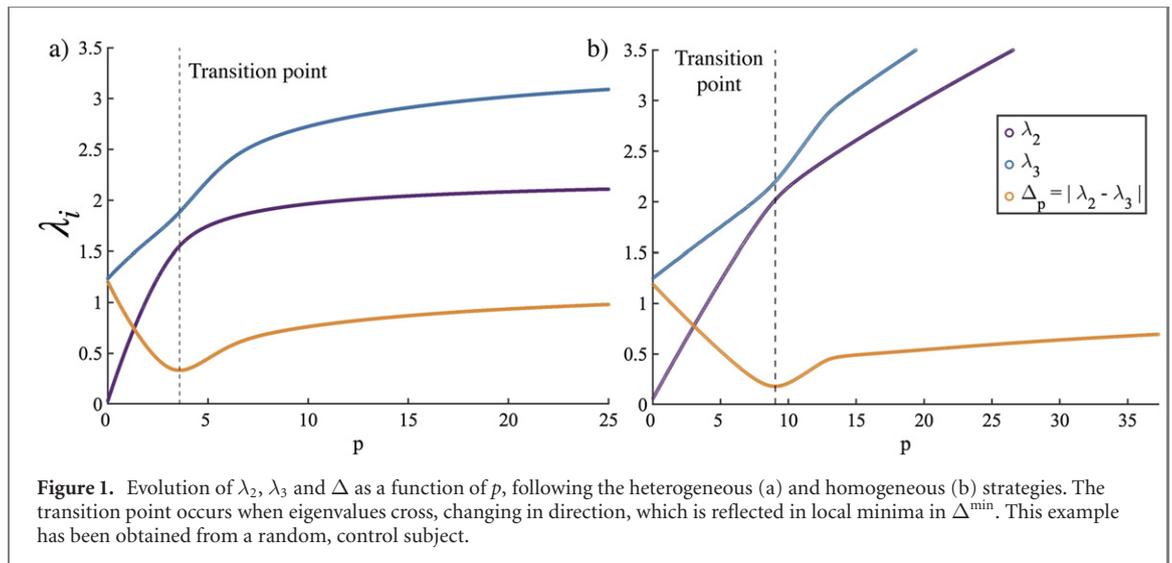**Figure 1.** Evolution of λ_2, λ_3 and Δ as a function of p, following the heterogeneous (a) and homogeneous (b) strategies. The transition point occurs when eigenvalues cross, changing in direction, which is reflected in local minima in $\Delta^{\min}$. This example has been obtained from a random, control subject.

The second smallest eigenvalue λ_2 (i.e., algebraic connectivity) is associated with the structure of the network, and in the multilayer/multiplex case, it undergoes a phase transition when layers become structurally dependent [22, 24, 28]. Given the fact that the evolution of the disease involves massive synapse loss and brain areas disconnection, we hypothesise that, functionally, this should be reflected in the integration between frequency bands, particularly between α and β. Particularly, we expect layers to be more disconnected than what we expect in a normal state (CG). Therefore, we will test (a) if the integration between bands, as quantified by the algebraic connectivity, is altered in MCI and AD at baseline (λ_{2-BSL}) and (b) how far each group is from the transition point, where λ_2 undergoes the transition typically observed in multilayer/multiplex networks when layers are structurally merged (λ_{2-T}).

The former test is achieved simply by comparing the values of λ_2 obtained from each network between groups, which we will call λ_{2-BSL} from now on. Bear in mind that this nomenclature (baseline) is simply used to differentiate the values of λ_2 when obtained empirically from every network in every subject. The latter, however, requires simulating how the network would evolve towards the transition and the calculation of λ_2 at that point, which we will call λ_{2-T}. We followed two different strategies to obtain λ_{2-T}, resembling the approach presented in [28]. In each $\alpha\beta$ network, we (a) multiplied every interlayer link by a constant (p; *heterogeneous strategy*), or (b) we first added the mean strength of the interlayer quadrant of the network and then multiplied every interlayer link by a constant (p; *homogeneous strategy*). We explored values of p from 0.1 to 70, obtaining the Laplacian matrix and its associated eigenvalues at each step. To detect the transition point, we obtained the absolute difference of λ_2 and λ_3 as a function of $p(\Delta(p) = |\lambda_2 - \lambda_3|)$. Local minima in $\Delta(p)$ indicate crossing points between eigenvalues; we were interested in the first minimum, where λ_2 and λ_3 cross, reflecting the structural transition in the multilayer. For illustrative purposes, figure 1 shows one example of λ_2, λ_3 and $\Delta(p)$ as a function of p, for the heterogeneous (a) and homogeneous (b) strategies, obtained from a random control subject.

Note that eigenvalues (λ_2 and λ_3) increase faster in the homogeneous strategy, as we are adding the average interlayer strength to all interlayer links before multiplying by the constant. As with λ_{2-BSL}, we will compare the values obtained from both the heterogeneous (λ_{2-T1}) and homogeneous (λ_{2-T2}) strategies between groups. Given that groups are not balanced, we had to implement an iterative strategy that uses all the information available in the trials/samples, ruling out any effect of inter-sample variability, and any group difference in such variability (*iterative two-way* ANOVA).

Lastly, we fitted a multinomial logistic regression model to the data, trying to predict group labels. This procedure allowed us to quantify a conversion rate to AD, taking into account λ_{2-BSL}, λ_{2-T}, age and sex. To fit the model, we previously averaged values of λ_{2-BSL} and λ_{2-T} across samples. A schematic representation of the pipeline to conduct our multilayer analyses can be found in the appendix (section 4, figure 7).

2.4. Single layer analyses

For each layer network (coming from sensors or ROIs) we obtained its weighted clustering coefficient [39], local and global efficiency [40], shortest path and outreach [5]. In unweighted networks, the clustering coefficient, also known as *transitivity*, measures the ratio of neighbours of a node that are also neighbours between them, or equivalently, the number of triangles in the graph. Its average, hence, quantifies the cliquishness of the network [41]. In weighted networks, the *weighted* clustering coefficient measures the local cohesiveness

taking into account the intensity of the interactions present on the local triplets [39]:

$$c^w(i) = \frac{1}{s_i(k_i - 1)} \sum_{j,h} \frac{(w_{ij} + w_{ih})}{2} a_{ij} a_{ih} a_{jh}. \tag{5}$$

The normalisation factor $(s_i(k_i - 1))$ ensures that $c^w(i)$ is bounded between 0 and 1, accounting for the weight of each link times the maximum number of possible triplets.

To measure the local and global efficiency, we followed to the formulation proposed by [40], considering the ability of connecting nodes in a low number of steps at each scale (local and global):

$$E_g = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}}, \tag{6}$$

where d_{ij} is the shortest topological distance between nodes i and j, computed using the Dijkstra's algorithm [42]. In this sense, the global efficiency is the harmonic mean of the inverse of the characteristic path length. Local efficiency, on the other hand, is obtained as:

$$E_l = \frac{1}{N} \sum_{i \in G} E(\mathbf{G_i}), \tag{7}$$

where $\mathbf{G_i}$ is the subgraph formed by the neighbours of i without i. This property is local in the sense that it is obtained at each node as the inverse of the average shortest path of the neighbours of i if that node were not present in that subgraph. In biophysical terms, the efficiency of the network reflects a trade-off between the cost of the paths connecting different nodes and its performance in carrying out a given task.

Lastly, the outreach parameter, formulated originally in [5], can be obtained as:

$$O(i) = \sum_{j \in V(i)} w_{ij} r_{ij}, \tag{8}$$

where r_{ij} is the physical distance between nodes and $V(i)$ is the subgraph of nearest neighbours of i. To obtain r_{ij}, we compute the Euclidean distance of the three spatial coordinates, which are given by the position of each sensor over the scalp or by the position of each ROI's centroid. The average outreach over all nodes reflects the dominance of long-range (high outreach) or short-range (low outreach) connections, and has been found to be altered in MCI and Alzheimer patients [5, 43].

2.5. Group comparisons: methodological considerations

Groups are not balanced: control and MCI groups are more populated ($N = 48$ and 46, respectively) than AD ($N = 17$). Also, we count with 29 trials/samples from each individual. Irrespective of the level of the type of network (multilayer or single layer), any group comparison entails the methodological drawback of different sample sizes. Under this circumstance, we can either average the parameter across samples, or treat trials from each subject as different observations (realisations) of the parameter from each group. In the former, every test would imply comparing three vectors of sizes 48, 46 and 17 and necessarily drawing upon non-parametric statistics, with more relaxed assumptions on the data (normality is not required, and violating the assumption of homoscedasticity is not as dramatic as in parametric tests), at the cost of less sensitivity to real group differences. In the latter, vectors to compare would be of length $48 \times 29 = 1392$, $46 \times 29 = 1334$ and $17 \times 29 = 493$, which allow us to make use of parametric statistics. Nonetheless, we face a possible problem of intra-subjects variability. Hence, we claim that to test for groups differences, the statistical approach should (i) take into account all sources of variability: between samples, in each individual; between subjects, in each group; and between groups, (ii) track the (possible) interactions between variability sources, (iii) be as sensitive as possible to group differences while (iv) using as much available data as possible.

Considering all together, each comparison was conducted with an iterative two-way ANOVA. The analysis of variance is parametric, thus more sensitive to real differences in the data in comparison to nonparametric approaches, and it is able to take into account the variability from samples, as well as the interaction between sample variability and group belonging. In each realisation, we subsampled 17 individuals from the control and MCI groups at random, thus ensuring the design is balanced. For each subject, we kept the value of the parameter under consideration in every sample (that is, 29 values for each subject) and performed a two-way ANOVA. We set GROUP and SAMPLE as factors for the analysis (analogous to a repeated measures ANOVA), where the main effects will be differences between groups and differences between samples. The interaction (GROUP $\times$ SAMPLE) would indicate that the variability in samples is group dependant and not homogeneous.

P-values indicate significant differences, but do not inform about the practical significance of results, something far more important in empirical studies [44]. At each step, then, we also estimated the size of each

effect, as measured by the partial eta squared (η_p^2), a common measure in mixed effects designs that relate to the amount of variance explained by the factor in consideration. We repeated the process (iterative two-way ANOVA, also calculating the effect size for each comparison) 100 times, with a different subsample in each iteration, ensuring the maximal efficiency in statistical terms, as we were using all available information from every subject. We were especially interested in observing group differences (main effect of group), and expected that the main effect due to samples, as well as the interaction between group and samples remained non-significant. In this regard, we considered that an effect was significant if and only if 90 out of 100 iterations were significant, evaluating in each case its associated *p*-value after correcting for multiple comparisons with FDR. In such cases, the final effect size was the average over the significant iterations.

3. Results

Every group comparison has been conducted with the iterative two-way ANOVA explained above. For every test, we provide *p*-values, always meeting the criteria established before (90 out 100 iterations reaching that *p*-value after FDR correction). We focussed on the results for the main effect of group, as our goal is to demonstrate differences that depend on the stage of the disease. Note that, in any of our analyses, the main effect of sample and the interaction between sample and group were significant, indicating that we can safely attribute differences to group belonging and not to variability in the samples. Additionally, we qualified each result with its corresponding size effect (η_p^2). Recall that the effect size quantifies the amount of variance in the data explained by the effect under consideration (group membership, samples or interaction between them). To grasp the implications of a particular value, it would be preferable to relate our results to previous literature [45], but given the lack of studies in which effect sizes are reported, we will follow the general recommendations originally provided by Cohen [46] to define small $(\eta_p^2 = 0.01)$, medium $(\eta_p^2 = 0.06)$ and large $(\eta_p^2 = 0.14)$ effects.
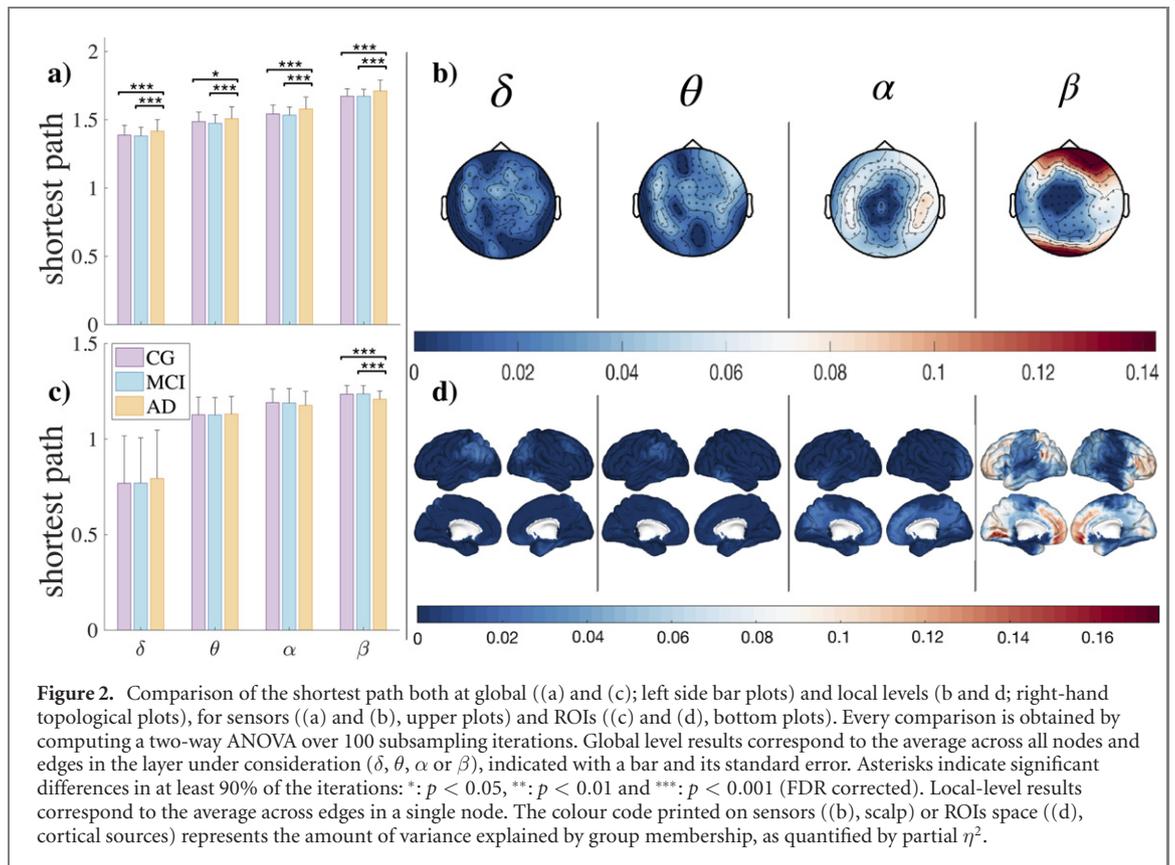
3.1. Single layer analyses

We carried out a single layer analysis of each of the four main frequency bands of the human brain, obtaining an independent network for each frequency. Networks were obtained in two modalities, at the (i) sensor level and (ii) the ROIs level (using the Harvard–Oxford atlas). Next, we calculated five network parameters for each modality and group, namely the weighted clustering coefficient, the global efficiency, the local efficiency, the outreach, and the shortest path.

Figure 2 shows, for every frequency layer (δ, θ, α and β), the comparison between one of these parameters, the shortest path, for the MCI, AD and control groups at two scales. At the global level, parameters are averaged over all nodes (or edges if it is an edge-level metric), yielding one single value for each sample and layer. Bars (with corresponding standard error) indicate the average in that layer for each group, with asterisks indicating the *p*-value of the comparison after FDR correction only if that *p*-value is met in 90 out of 100 iterations ((a) and (c) in the figure for sensors and ROIs respectively). At the local (node) level, results are illustrated over scalp plots (sensors, (b)) or cortical surface plots (ROIs, (d)), with the average value of η^2 colour-coded. In that case, we set to 0 the value of η^2 if that comparison is not significant, or take the average over all significant iterations if it is ($p < 0.05$ after FDR correction in at least 90% iterations).

In this example, we illustrate the results found in shortest path, showing that in sensors there are significant global differences between controls and AD, and between MCI and AD (figure 2(a)) in δ ($p < 0.001$ in both cases, $\eta_p^2 = 0.0475$); θ ($p < 0.05$ and $p < 0.001$, $\eta_p^2 = 0.0422$); α ($p < 0.001$ in both cases, $\eta_p^2 = 0.0754$) and β ($p < 0.001$, both cases $\eta_p^2 = 0.0838$). In this regard, there is an effect of group in all bands, with sizes being small in δ and θ and medium in α and β, which is coherent with previous reports indicating that these last two bands are the most affected. The effect of samples and the interaction with group were not significant in any comparison. Note how, in any case, the effects are smoothed when taken globally: when we make comparisons at the node level, local alterations show greater effect sizes (figure 2(b)), specially in temporo-parietal areas (right hemisphere, α, $\eta_p^2 \approx 0.1$), and temporo-parietal, frontal and occipital areas, again with a right hemisphere preference in β band ($\eta_p^2 \approx 0.14$).

In ROIs, global alterations are less pronounced, as we only found differences in β band between controls and AD ($p < 0.001$) and between MCI and AD ($p < 0.001$), with a medium effect size ($\eta_p^2 = 0.0866$). Local differences are coherent with sensors: we find small effect sizes in δ and θ ($\eta_p^2 \approx 0.02$) corresponding to alterations in the parietal cortex (somatosensory association cortex), supramarginal gyrus, and entorhinal cortex. In the α band, the significantly affected areas ($\eta_p^2 \approx 0.04$) are: the orbitofrontal prefrontal cortex, and specially cuneus, precuneus, superior, and medial orbitofrontal cortex, as well as the cingulate cortex. Nonetheless, the largest effects, as in sensors, are found in the β band, including all previously mentioned areas, but including broader regions (prefrontal cortex) and lingual gyrus ($\eta_p^2 \approx 0.15$). Interestingly, in α and β, the alteration is bilateral, not confined to one hemisphere.

Figure 2. Comparison of the shortest path both at global ((a) and (c); left side bar plots) and local levels (b and d; right-hand topological plots), for sensors ((a) and (b), upper plots) and ROIs ((c) and (d), bottom plots). Every comparison is obtained by computing a two-way ANOVA over 100 subsampling iterations. Global level results correspond to the average across all nodes and edges in the layer under consideration (δ, θ, α or β), indicated with a bar and its standard error. Asterisks indicate significant differences in at least 90% of the iterations: *: $p < 0.05$, **: $p < 0.01$ and ***: $p < 0.001$ (FDR corrected). Local-level results correspond to the average across edges in a single node. The colour code printed on sensors ((b), scalp) or ROIs space ((d), cortical sources) represents the amount of variance explained by group membership, as quantified by partial η^2.

Next, we calculated the weighted clustering (equation (5); figures 3(a) and (b)), the local efficiency (equation (7); panels (c) and (d)), the outreach (equation (8); panels (e) and (f)) and the global efficiency (equation (6); panels (g) and (h)). See section 2.3 for more details on the equations to obtain every measure. Interestingly, the weighted clustering coefficient on sensors was significantly different between CG–AD ($p < 0.001$) and MCI–AD ($p = 1e − 06$) in the δ band. We also found differences in θ but only between CG–AD ($p < 0.001$, $\eta_p^2 = 0.029$). When the same analysis is translated to the ROIs level differences are reported at higher frequency bands: the θ layer showed differences between CG and AD ($p < 0.05$, $\eta_p^2 = 0.023$), as well as the α layer (CG–AD, $p < 0.001$, $\eta_p^2 = 0.036$). Finally, the β layer is the one showing more differences both at CG–AD ($p = 1e − 06$) groups and MCI–AD ($p = 1e − 05$, $\eta_p^2 = 0.093$). We found a similar trend in the local efficiency (only δ being significant in sensors, and β in ROIs). The outreach parameter showed differences between the analysis at the sensor and the ROIs level as well. Again, δ and θ layers are the ones showing statistical significant differences between CG–AD and MCI–AD, while at the ROIs level differences appear at layers α and β, i.e. move to higher frequencies and only between MCI–AD (α) and CG–AD (β). Only global efficiency was consistent in both modalities, with all bands showing differences between controls and AD group, and between MCI and AD group. Note that global efficiency is the only measure that works at a global scale by construction, while the rest has been calculated at the node level and averaged over the whole network.

3.2. Multilayer analyses

Next, we investigated the differences between groups and modalities when $\alpha\beta$ MN are constructed (see section methods) and analysed. First, we checked whether network density and strength (i.e., the sum of the links' weights) differ between groups. As shown in figures 4(a) and (b), we could not find any difference between groups in density, regardless of the modality. The effect of the samples and the interaction between group and samples were not significant either. Figures 4(c) and (d) compares the strength between groups. In this case, any tested effect was statistically significant for sensors (main effects of group and sample, and interaction between them). On the contrary, we found a significant difference between MCI and AD groups ($p < 0.05$), with a small effect size ($\eta_p^2 = 0.046$) in ROIs. This could be due to the normalisation procedure (circular shifting) or due to real differences in strength, derived from the disease. Take into account that, although the effect is small (around 4.6% of the variance in strength is attributable to the clinical condition), this difference could bias further analyses.
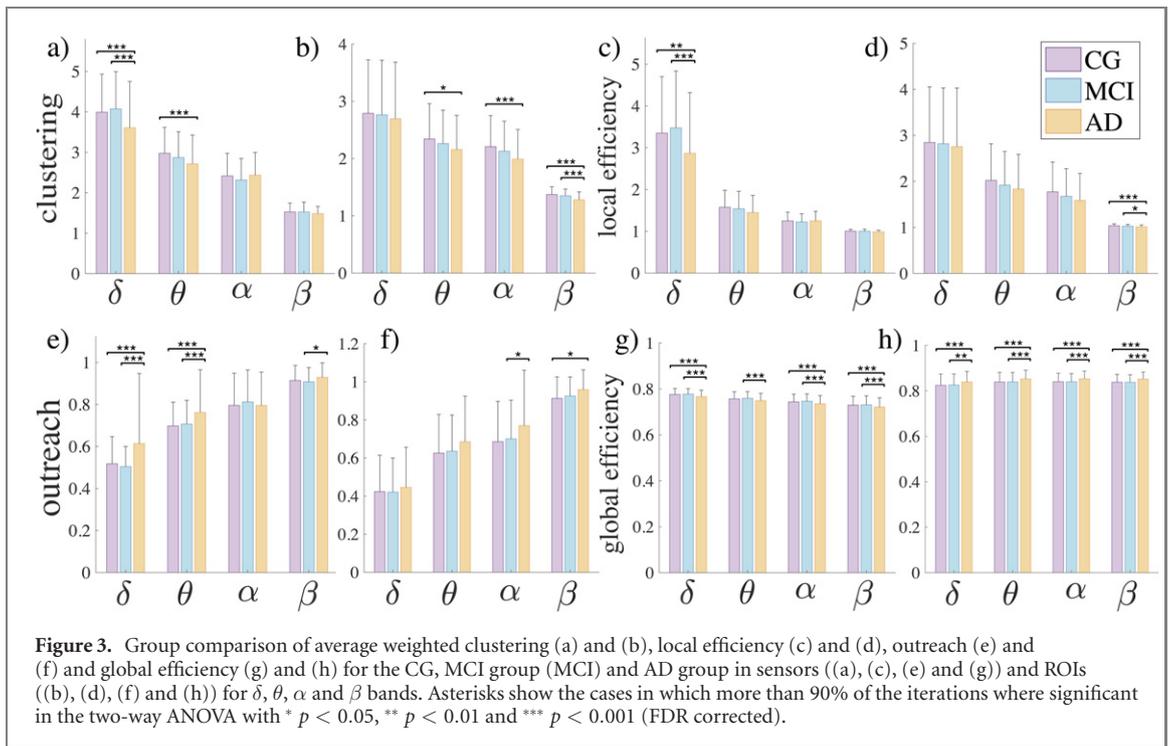
Figure 3. Group comparison of average weighted clustering (a) and (b), local efficiency (c) and (d), outreach (e) and (f) and global efficiency (g) and (h) for the CG, MCI group (MCI) and AD group in sensors ((a), (c), (e) and (g)) and ROIs ((b), (d), (f) and (h)) for δ, θ, α and β bands. Asterisks show the cases in which more than 90% of the iterations where significant in the two-way ANOVA with * $p < 0.05$, ** $p < 0.01$ and *** $p < 0.001$ (FDR corrected).
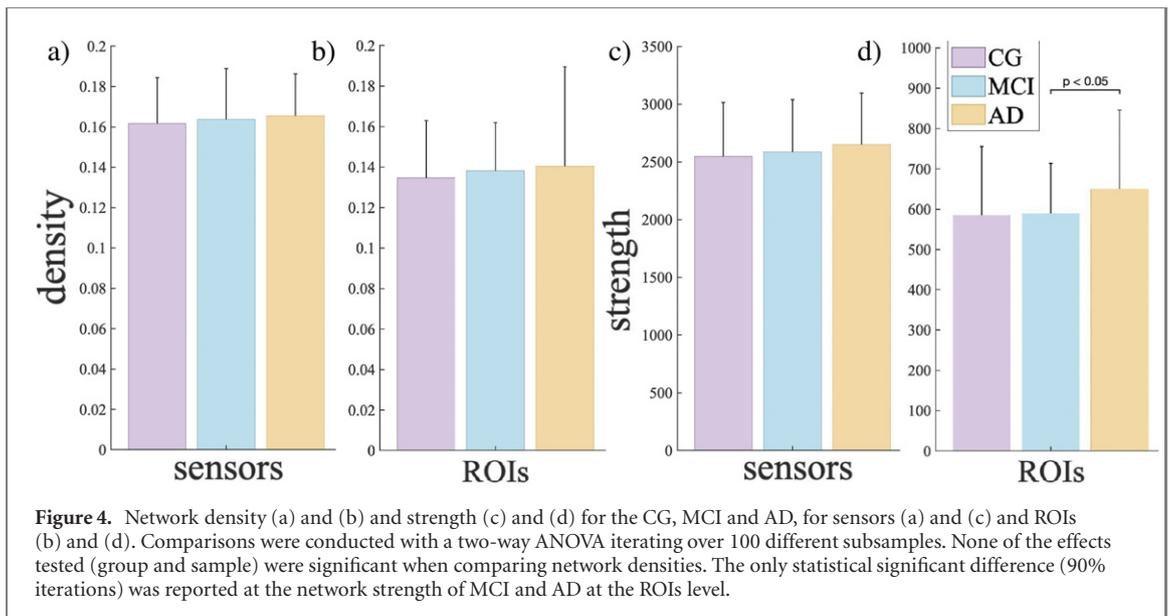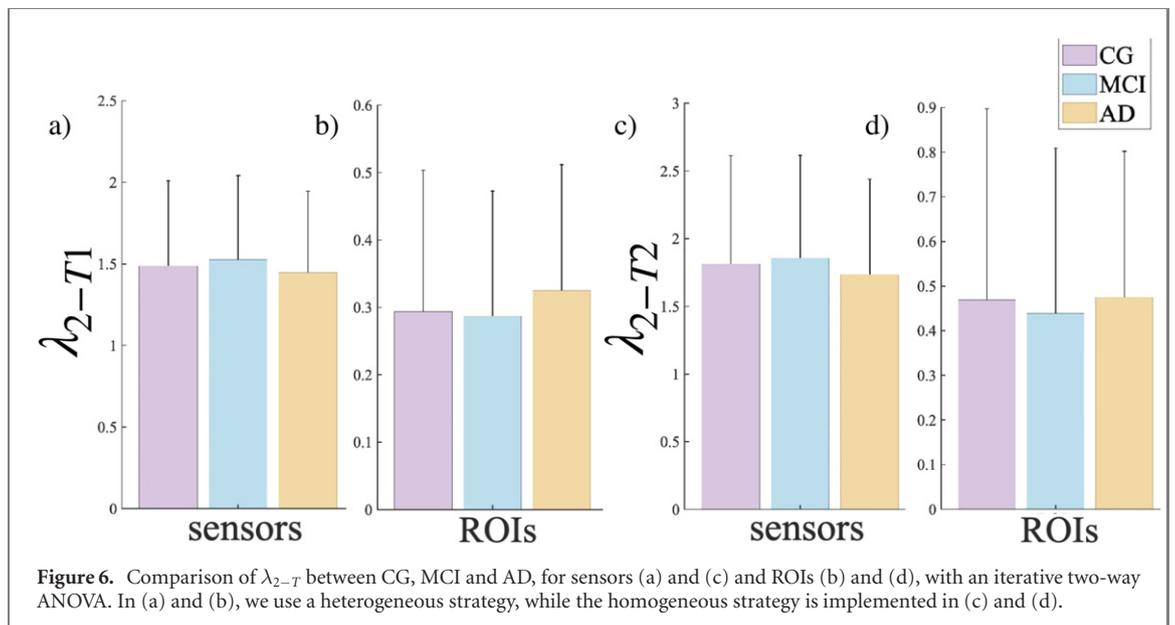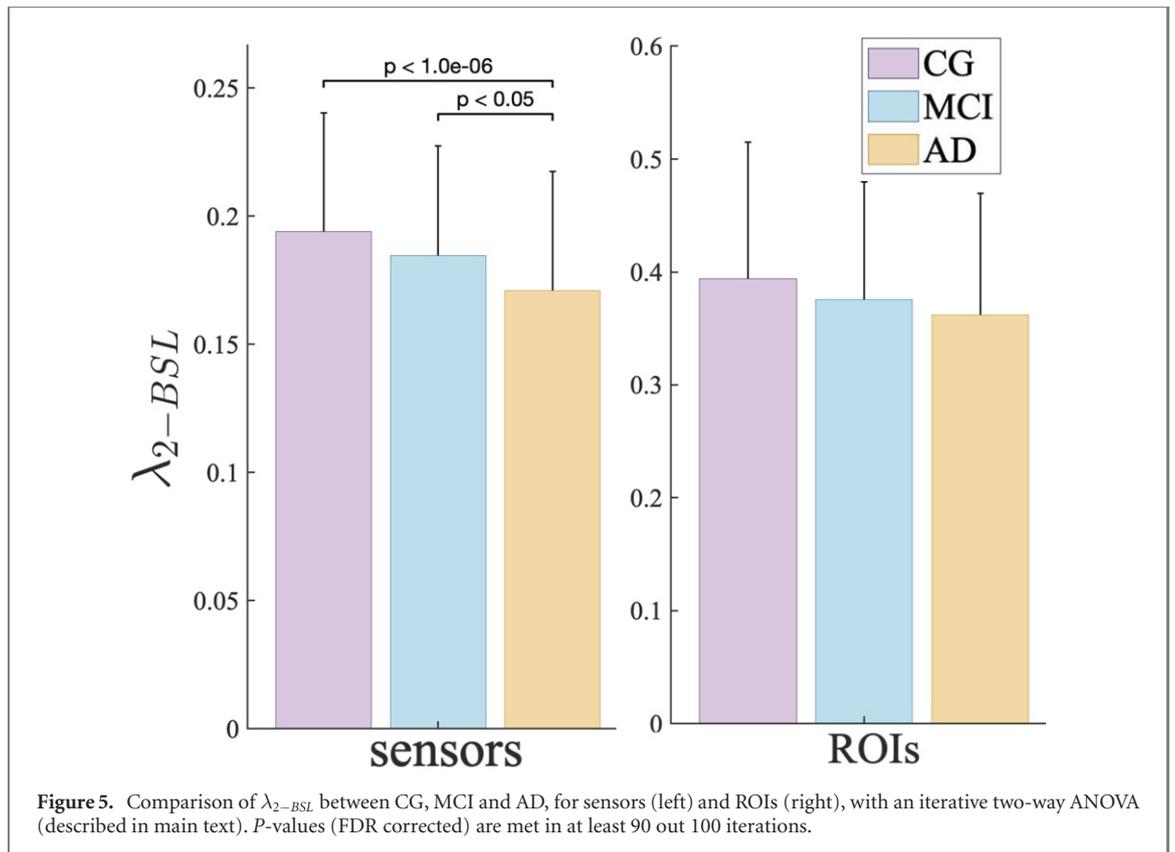

Figure 4. Network density (a) and (b) and strength (c) and (d) for the CG, MCI and AD, for sensors (a) and (c) and ROIs (b) and (d). Comparisons were conducted with a two-way ANOVA iterating over 100 different subsamples. None of the effects tested (group and sample) were significant when comparing network densities. The only statistical significant difference (90% iterations) was reported at the network strength of MCI and AD at the ROIs level.

Next, we compared the values of λ_2 at baseline and at the transition point, the latter obtained from the homogeneous and heterogeneous strategies (see methods), using the iterative two-way ANOVA. It is worth noting that, as a consequence of the normalisation scheme applied (circular shifting), it is possible that in some cases a network is disconnected, leading to an algebraic connectivity (λ_2) equal to zero. None of the networks built from sensors met this condition, whereas in ROIs, this was the case of the 14.5% of networks in the CG, 15.97% in MCI and 9.13% in AD. All disconnected networks were discarded for further analyses.

We found significant differences between controls and MCI ($p < 1e - 06$) and between MCI and AD ($p < 0.05$) in λ_2 at baseline (figure 5), with a small size effect ($\eta_p^2 = 0.0465$). The effect of samples and the interaction between group and sample were not significant. Although the group effect could explain around 4.6% of the variance in λ_2, it is indicating that as the disease progresses, α and β bands become more independent, given that at each group, the average decreases. This might be reflecting a loss in the ability of layers to interact between them, which might be paramount for the brain to work properly. Interestingly, this tendency is also observed in ROIs, although in this case, differences are not statistically significant. Take into account

Figure 5. Comparison of λ_{2-BSL} between CG, MCI and AD, for sensors (left) and ROIs (right), with an iterative two-way ANOVA (described in main text). *P*-values (FDR corrected) are met in at least 90 out 100 iterations.

Figure 6. Comparison of λ_{2-T} between CG, MCI and AD, for sensors (a) and (c) and ROIs (b) and (d), with an iterative two-way ANOVA. In (a) and (b), we use a heterogeneous strategy, while the homogeneous strategy is implemented in (c) and (d).

that this could be a side effect of the missing values reported before (as we are excluding some data from the analyses due to network disconnection).

Next, we investigate what the value of λ_2 would be when reaching the transition point. With this aim, we increase the weight of the inter-layer links, as explained in the method section, using (i) a homogeneous strategy and (ii) a heterogeneous strategy. Both approximations to reach the transition point are consistent in each modality (sensors/ROIs), albeit results are opposite (figure 6): in sensors (panels (a) and(c)), for both approaches, the AD group had a smaller λ_2 when reached the transition point, followed by CG, with MCI having the largest. On the contrary, in ROIs (panels (b) and (d)), AD had the highest values of λ_2 when compared to the other two groups, and MCI the smallest. However, none of these differences was statistically significant.

Finally, we tried to fit different, nested multinomial logistic regression models (MNR models), with the aim of predicting group labels with the parameters extracted from the MN (λ_{2-BSL} and λ_{2-T}). Importantly,

we first averaged the parameters across samples with the consequent information and predictive power loss. Nonetheless, this step is necessary to conduct the logistic regression, as keeping all samples would incur in overfitting. In the context of linear regressions, keeping all samples would be analogous to duplicate subjects, increasing the percentage of cases correctly classified artificially.

MNR models are a special case of generalised linear models (GLM), with the particularity that, instead of predicting probabilities, they predict the logarithm of the odds of a certain category (relative risk), using the logit function as the canonical link function [47]. This makes them especially suitable to classify subjects and have been extensively used in machine learning approaches. In our case, the advantage of applying an MNR model is that it enables us to give a probability of membership to each group depending on the values of each parameter. Hence, we could predict a conversion ratio if any of those parameters change beyond a given value. The goal was to keep the most parsimonious model, i.e. keeping the smallest number of terms while explaining as much variance as possible. Thus, we proceeded iteratively, adding one extra term (parameter) at each step. In all cases, being π_k the theoretical probability associated to each category in the response variable (CG/MCI/AD), and taking the CG as the reference category, K, we can define $K - 1$ non-redundant equations of the form:

$$\ln \frac{\pi_k}{\pi_K} = \beta_{0k} + \beta_{1k} X_1 + \beta_{2k} X_2 + \cdots + \beta_{jk} X_j + \beta_{pk} X_p, \tag{9}$$

where K is the reference category (CG) and k any of the other two (MCI/AD). Each regression coefficient had one subindex ($j = 1, 2, \ldots, p$) for each of the p independent variables (parameters added to the model) and another one ($k = 1, 2, \ldots, K - 1$) to identify each equation. As we only had three categories, every model considered was composed of two equations, each one describing the relative risk of one subject being MCI(AD) with respect to CG, and containing a coefficient for the intercept and a separate slope on each predictor (parameter). To decide which model is preferred over the others, we observe how adding or removing a parameter reduces the deviance from the fit, calculated as:

$$\text{dev} = 2 \sum_i^n \sum_j^n (y_{ij}) \log \left(\frac{y_{ij}}{\hat{\pi}_{ij} m_i} \right) = \sum_i^n rd_i, \tag{10}$$

where y_{ij} is the ith observation in the jth category (group), $\hat{\pi}_{ij}$ its categorical, estimated probability and m_i its corresponding sample size. The deviance parameter is obtained in GLM as the difference in the fit between the full and reduced models [48]. The full model uses a separate parameter for each sample, hence having as many parameters as subjects (it is also called 'complex model'). The reduced model includes only a set of parameters and, although its fit is always worse in comparison to the full model, it does not require as many parameters. In the case of logistic (either binary or multinomial) regression models, the deviance parameter is estimated comparing the maximum likelihood of the full model with the likelihood of the reduced one, and asymptotically follows a χ^2 distribution [49]. As mentioned above, at each step, we added one term (variable) and compared it with the previous model to test if adding one extra term reduces the deviance of the fit. To compare between models, we simply subtracted their deviances and checked the cumulative probability of the result under its χ^2 distribution with degrees of freedom equal to the difference of each model's degrees of freedom. A significant ($1 - P_{(\text{dev1}-\text{dev2})} < \alpha$) indicates that the coefficient of the new term is significantly different from zero, and we can safely include it in the model, reducing its deviance.

We first defined a simple model ($mdl(1)$) containing only demographic variables (age and sex) to predict group labels. More details regarding their deviances and values for the coefficients can be found in the appendix, section 4. We depart from this model, adding more terms for each parameter under consideration. From all the other models we have fitted, the best scores are obtained when we only add one extra term for λ_2 at baseline, in both modalities. This is expected, as λ_{2-BSL} was the only parameter showing significant group-dependent differences (statistically significant in sensors and with a clear tendency in ROIs).

The model for sensors ($mdl(2S)$) is:

$$\ln \frac{\pi_{\text{MCI}}}{\pi_{\text{CG}}} = -1.3938 - 11.9491 \times \lambda_{2-BSL} + 0.0554 \times \text{age} - 0.2331 \times \text{sex} \tag{11}$$

$$\ln \frac{\pi_{\text{AD}}}{\pi_{\text{CG}}} = -15.7062 - 29.3990 \times \lambda_{2-BSL} + 0.288\,6 \times \text{age} - 0.917\,0 \times \text{sex} \tag{12}$$

with $\text{dev}_{mdl(2S)} = 197.2716$. For the first equation, the p-values associated to each term were (from left to right), $p = 0.7089$, $p = 0.1271$, $p = 0.2385$, $p = 0.5827$. For the second, $p = 0.0161$, $p = 0.0247$, $p = 0.0003$, $p = 0.1856$. In this case, every coefficient in the first equation (relative risk of MCI over CG) was not significant; while the intercept, and the effects associated to λ_{2-BSL} and age must be taken into account when comparing AD and CG. If we take $\text{dev}_{1vs2S} = \text{dev}_{mdl(1)} - \text{dev}_{mdl(2S)} \rightarrow 203.8440 - 197.2716 = 6.5724$, and given that $dev_{1vs2S} \sim \chi^2(dfe_1 - dfe_{2S} = 2)$, we can state that the new model adds significant information, as

$1 - \chi^2_{6.5724}(2) = 0.0374$. This result indicates that, in principle, λ_{2-BSL} could be a useful biomarker in clinical contexts.

The model for ROIs ($mdl(2R)$) is:

$$\ln \frac{\pi_{\text{MCI}}}{\pi_{\text{CG}}} = -1.590\,7 - 6.497\,3 \times \lambda_{2-BSL} + 0.061\,9 \times \text{age} - 0.2496 \times \text{sex} \tag{13}$$

$$\ln \frac{\pi_{\text{AD}}}{\pi_{\text{CG}}} = -19.2093 - 11.1362 \times \lambda_{2-BSL} + 0.3171 \times \text{age} - 0.7697 \times \text{sex} \tag{14}$$

with $\text{dev}_{mdl(2R)} = 197.9304$, and associated p-values (first equation) $p = 0.6556$, $p = 0.0708$, $p = 0.1858$, $p = 0.5584$ and (second equation) $p = 0.0015$, $p = 0.0357$, $p = 0.0001$, $p = 0.2469$. Results were consistent with the model for sensors data, with first equation's coefficients being not statistically significant, and intercept, λ_2 and age effects being relevant. Comparing the deviance with the simplest model gave us $1 - \chi^2_{5.9136}(2) = 0.0520$, indicating that adding the new term does not reduces the deviance of the fit significantly. Nonetheless, the intercept and the coefficient of λ_2 were significant in the second equation (AD vs CG), which again is expected given that these groups should be the most different.

4. Conclusions

In this paper, we first investigated the differences between single layer functional networks of patients suffering from MCI and AD. Next, we extended our analysis to frequency-band MN and presented an application of the algebraic connectivity (λ_2) as a biomarker of AD in a clinical context.

We provided a complete analysis of single layer networks, focussing on the shortest-path parameter but also analysing differences between groups in the weighted clustering coefficient, the global and local efficiency and the outreach. It is worth noting that path length-based metrics are hard to interpret [50, 51], but we included them as an initial step before analyzing the multilayer structure. We found that each parameter had its own distribution over sensor and ROI space, consistent between modalities, that strongly depended on the band and group under consideration. Take into consideration that each rhythm has its own functional role, spatial pattern of synchronization [52] and complexity/entropy distribution [30]. While all network parameters showed some statistically significant differences between controls and AD, the global efficiency was the one showing more differences, both at the sensor and ROIs level. Interestingly, we reported difference between the network parameters when both modalities were compared with a particular (and general) pattern: at the sensor level, differences between groups are more significant at networks related to low-frequency bands, while at the ROIs level, differences are more significant at higher frequencies. Note that every measure used to analyse MEG/EEG signals is affected by signal leakage and source reconstruction [53–55], which may introduce spurious estimates of functional connectivity, biasing the estimation of the network topology. Although there are different algorithms and atlases at disposal to project brain activity, all of them have advantages and pitfalls; thus, some degree of error in the estimation of the network parameters is unavoidable. It is precisely in this regard that we analysed the functional networks obtained at the sensor and ROIs level, with the aim of detecting differences between both projections of the brain activity. Furthermore, we applied the method of circular shifting, ensuring that every weight in the network is significant in comparison to a data-driven null model of connectivity. In this regard, the choice of MI over other FC metrics less sensitive to signal leakage comes from the fact that it enables us to estimate intra and interlayer connectivity in the same scale, and under the same assumptions. Given that we expect leakage to affect the three groups similarly, we consider that results won't be biased in any direction.

Concerning the multilayer analysis, we focussed on algebraic connectivity, a parameter related to the structural and dynamic properties of MN. We constructed the $\alpha\beta$ MN, computed their algebraic connectivity and simulated the value of this parameter at the transition point where the MN suffer a well-known transition where layers are structurally merged. We reported significant differences between controls, MCI and AD in sensors and a strong (but not statistically significant) tendency in ROIs, consistent with the former. Our inability to prove these differences at the ROIs level might come from missing values that correspond to disconnected networks (where $\lambda_2 = 0$), as such cases were discarded from our analyses. Nonetheless, both modalities were consistent, and the underlying differences in λ_2 were informative when trying to classify subjects, as predicted by our logistic regression model presented in section 3. We conclude that in the biological progression of the disease, α and β bands tend to disconnect, losing the optimal level of communication. In this direction, clinical approaches might focus on reconnecting functionally these layers. Nonetheless, more data, preferably from follow-up sessions of the same subjects, could improve our statistical power and give new hints on how this parameter evolves with the disease progression. Our goal was to explore if the interdependency between bands is associated with the disease and if it could be used as a biomarker in AD. Thus, our model included only this term along with demographic variables. In the future, we could connect this finding with other biomarkers

Table 1. Table with demographic information from the CG.
MMSE = mini-mental state examination. Sex: 1 = Male; 2 = female.

ID	Group	Age	Sex	MMSE
1	CG	67	1	30
2	CG	72	1	30
3	CG	78	2	29
4	CG	77	1	28
5	CG	70	2	30
6	CG	70	2	30
7	CG	72	1	27
8	CG	69	1	28
9	CG	66	2	29
10	CG	70	1	28
11	CG	69	1	30
12	CG	71	2	30
13	CG	68	2	27
14	CG	65	2	27
15	CG	66	1	30
16	CG	76	2	30
17	CG	73	2	29
18	CG	77	1	27
19	CG	78	1	30
20	CG	65	2	30
21	CG	73	2	30
22	CG	69	2	30
23	CG	75	1	27
24	CG	74	1	29
25	CG	67	2	30
26	CG	71	1	30
27	CG	73	2	28
28	CG	75	1	27
29	CG	65	2	28
30	CG	73	1	29
31	CG	72	1	27
32	CG	80	1	27
33	CG	73	2	30
34	CG	68	2	28
35	CG	70	2	28
36	CG	68	2	30
37	CG	70	1	30
38	CG	75	2	28
39	CG	80	2	27
40	CG	71	1	30
41	CG	70	2	30
42	CG	75	1	27
43	CG	69	2	30
44	CG	74	2	30
45	CG	65	2	30
46	CG	75	2	28
47	CG	60	2	30
48	CG	70	1	27
	MEAN	71.23	1.56	28.83
	SD	4.31	0.50	1.26

found in the literature, and formulate a model that includes more terms from different sources: single layer network measures that differ between groups, along with MMSE scores and physiological variables, in line with previous literature aiming at quantifying the MCI–AD conversion rate [56–61].

Regarding our simulations on the transition point of the algebraic connectivity (λ_{2-T}), we could not find any remarkable difference, regardless of the strategy followed (homogeneous or heterogeneous). This could reflect an ill-posed formulation of the problem, and alternative strategies might be more successful. Note that the phase transition in λ_2 has been only observed in ideal setups, where all interlayer weights have the same value, and density is fixed for every network.

Finally, it is worth noting that we consider the current study just as a starting point. We believe that more research is needed to have a full description of the problem, with longitudinal data and more patients in each group, to unveil differences to this level of detail. Also, this circumstance points towards a crucial weakness of our study in particular and many others in general: the small sample sizes, especially in the AD group. Despite the fact that two-way ANOVA revealed systematically that the effect of samples (and the interaction with group)

Table 2. Table with demographic information from the MCI group.
MMSE = mini-mental state examination. Sex: 1 = Male; 2 = female.

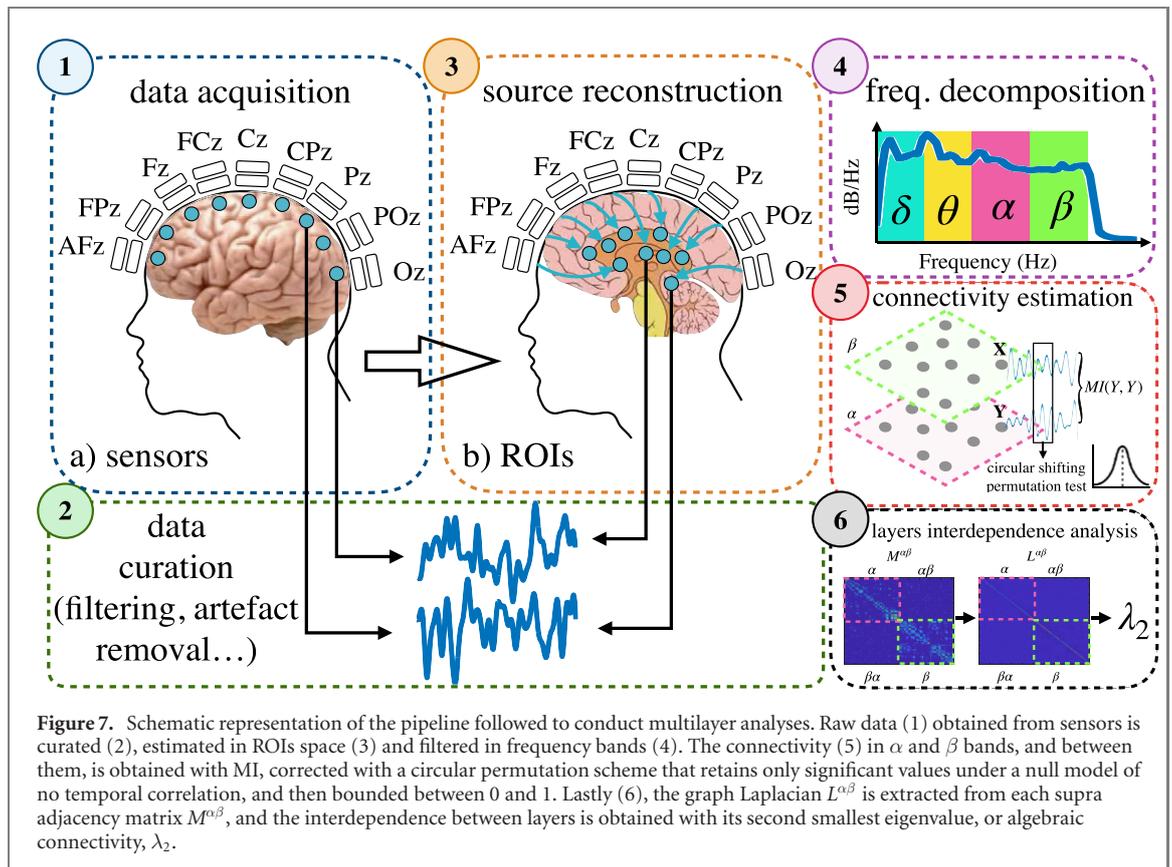| ID	Group	Age	Sex	MMSE
49	MCI	72	2	28
50	MCI	79	2	28
51	MCI	77	1	28
52	MCI	73	2	28
53	MCI	67	1	26
54	MCI	81	1	18
55	MCI	72	2	27
56	MCI	69	1	28
57	MCI	78	2	22
58	MCI	68	2	29
59	MCI	75	2	23
60	MCI	72	2	25
61	MCI	81	2	26
62	MCI	77	2	27
63	MCI	74	1	26
64	MCI	64	2	27
65	MCI	69	2	26
66	MCI	73	1	30
67	MCI	81	2	27
68	MCI	73	2	28
69	MCI	75	2	25
70	MCI	79	2	29
71	MCI	66	2	29
72	MCI	72	1	27
73	MCI	68	2	28
74	MCI	67	2	27
75	MCI	74	2	29
76	MCI	69	2	24
77	MCI	76	1	29
78	MCI	66	1	26
79	MCI	69	1	23
80	MCI	72	1	23
81	MCI	78	1	28
82	MCI	77	2	27
83	MCI	69	1	29
84	MCI	69	2	28
85	MCI	78	2	20
86	MCI	73	1	28
87	MCI	66	1	30
88	MCI	66	1	27
89	MCI	71	1	30
90	MCI	72	1	27
91	MCI	68	1	28
92	MCI	69	1	29
93	MCI	73	1	27
94	MCI	71	1	27
	MEAN	72.35	1.52	26.76
	SD	4.56	0.50	2.53

was not significant, and hence we could safely compare all datum from every subject, we require more data to give robustness to our conclusions. In future studies, it would be interesting to explore differences in the interplay between other frequency bands, building MN comprising other combinations of layers, or possibly including all of them. In our study, we chose α and β because those ranges accumulate more power in the spectrum, show clear differences in complexity and entropy [30] and have been identified as the ones with highest impairment in AD [31].

Data availability statement

The data generated and/or analysed during the current study are not publicly available for legal/ethical reasons but are available from the corresponding author on reasonable request.

Table 3. Table with demographic information from the AD.
MMSE = mini-mental state examination. Sex: 1 = Male; 2 = female.

ID	Group	Age	Sex	MMSE
95	AD	77	2	18
96	AD	74	1	29
97	AD	70	1	28
98	AD	80	1	17
99	AD	77	1	23
100	AD	79	1	25
101	AD	63	1	28
102	AD	80	2	16
103	AD	73	2	18
104	AD	80	2	20
105	AD	83	2	21
106	AD	76	1	25
107	AD	81	2	23
108	AD	82	2	20
109	AD	77	2	17
110	AD	76	1	19
111	AD	87	1	23
	MEAN	77.35	1.47	21.76
	SD	5.46	0.51	4.16

Figure 7. Schematic representation of the pipeline followed to conduct multilayer analyses. Raw data (1) obtained from sensors is curated (2), estimated in ROIs space (3) and filtered in frequency bands (4). The connectivity (5) in α and β bands, and between them, is obtained with MI, corrected with a circular permutation scheme that retains only significant values under a null model of no temporal correlation, and then bounded between 0 and 1. Lastly (6), the graph Laplacian $L^{\alpha\beta}$ is extracted from each supra adjacency matrix $M^{\alpha\beta}$, and the interdependence between layers is obtained with its second smallest eigenvalue, or algebraic connectivity, λ_2.

Appendix A

A.1. Demographic information
See tables 1–3

A.2. Multilayer pipeline analysis
See figure 7

A.3. Null model for λ_2

The simplest model ($mdl(1)$) containing only demographic variables (age and sex) to predict group labels reads:

$$\ln \frac{\pi_{\text{MCI}}}{\pi_{\text{CG}}} = -3.64 + 0.053\,8 \times \text{age} - 0.17 \times \text{sex} \tag{15}$$

$$\ln \frac{\pi_{\text{AD}}}{\pi_{\text{CG}}} = -22.6304 + 0.3032 \times \text{age} - 0.6336 \times \text{sex}. \tag{16}$$

The deviance of the model is $\text{dev}_{mdl(1)} = 203.8440$. The coefficients express both the effects of the predictor variables on the relative risk and the logarithmic odds of being in one category versus the reference category (CG). Intercepts were $-3.64 (p = 0.2760)$ and $-22.6304 (p = 0.0001)$, with coefficients $0.0538 (p = 0.2397)$ and $0.3032 (p = 0.0001)$ for age and $-0.1700 (p = 0.6836)$ and $-0.6336 (p = 0.3265)$ for sex. Note how the intercept in the first equation and the coefficients for sex in both equations were not statistically significant, while the intercept and effect of age was statistically significant only when we compare AD and CG groups (second equation), which makes sense if we take into account that these two groups should differ more. With this model, we could safely state that the relative risk of being AD versus CG increases $\exp(0.3032) = 1.3542$ times for each year increase, given that the rest of the variables are equal. Interpretations regarding sex must be taken with caution, given that the coefficients associated with this effect are not significant. Nonetheless, we will keep this term for practical purposes.

ORCID iDs

Ignacio Echegoyen ⓘ https://orcid.org/0000-0001-8016-7836

References

[1] Sporns O 2011 *Networks of the Brain* (Cambridge, MA: MIT Press)
[2] Bullmore E and Sporns O 2009 Complex brain networks: graph theoretical analysis of structural and functional systems *Nat. Revi. Neurosci.* **10** 186–98
[3] Stam C J 2014 Modern network science of neurological disorders *Nat. Rev. Neurosci.* **15** 683–95
[4] Meunier D, Lambiotte R, Fornito A, Ersche K D and Bullmore E 2009 Hierarchical modularity in human brain functional networks *Fron. Neuroinf.* **3** 37
[5] Buldú J M *et al* 2011 Reorganization of functional networks in mild cognitive impairment *PloS one* **6** e19584
[6] Zanin M, Papo D, Sousa P A, Menasalvas E, Nicchi A, Kubik E and Boccaletti S 2016 Combining complex networks and data mining: why and how *Phys. Rep.* **635** 1–44
[7] Zanin M 2015 Can we neglect the multi-layer structure of functional networks? *Physica* A **430** 184–92
[8] de Domenico M, Solé-Ribalta A, Cozzo E, Kivelä M, Moreno Y, Porter M A, Gómez S and Arenas A 2013 Mathematical formulation of multilayer networks *Phys. Rev.* X **3** 041022
[9] Boccaletti S, Bianconi G, Criado R, Del Genio C I, Gómez-Gardeñes J, Romance M, Sendiña-Nadal I, Wang Z and Zanin M 2014 The structure and dynamics of multilayer networks *Phys. Rep.* **544** 1–122
[10] Kivelä M, Arenas A, Barthelemy M, Gleeson J P, Moreno Y and Porter M A 2014 Multilayer networks *J. Complex Netw.* **2** 203–71
[11] Nicosia V, Bianconi G, Latora V and Barthelemy M 2013 Growing multiplex networks *Phys. Rev. Lett.* **111** 058701
[12] Bassett D S, Wymbs N F, Porter M A, Mucha P J, Carlson J M and Grafton S T 2011 Dynamic reconfiguration of human brain networks during learning *Proc. Natl Acad. Sci.* **108** 7641–6
[13] Bassett D S, Wymbs N F, Rombach M P, Porter M A, Mucha P J and Grafton S T 2013 Task-based core-periphery organization of human brain dynamics *PLoS Comput. Biol.* **9** e1003171
[14] Mucha P J, Richardson T, Macon K, Porter M A and Onnela J-P 2010 Community structure in time-dependent, multiscale, and multiplex networks *Science* **328** 876–8
[15] Helfrich R F, Herrmann C S, Engel A K and Schneider T R 2016 Different coupling modes mediate cortical cross-frequency interactions *Neuroimage* **140** 76–82
[16] Canolty R T and Knight R T 2010 The functional role of cross-frequency coupling *Trends Cogni. Sci.* **14** 506–15
[17] Fries P 2015 Rhythms for cognition: communication through coherence *Neuron* **88** 220–35
[18] Buzsáki G and Draguhn A 2004 Neuronal oscillations in cortical networks *Science* **304** 1926–9
[19] Brookes M J, Tewarie P K, Hunt B A E, Robson S E, Gascoyne L E, Liddle E B, Liddle P F and Morris P G 2016 A multi-layer network approach to meg connectivity analysis *Neuroimage* **132** 425–38
[20] Yu M, Engels M M A, Hillebrand A, Van Straaten E C W, Gouw A A, Teunissen C, van der Flier W M, Scheltens P and Stam C J 2017 Selective impairment of hippocampus and posterior hub areas in Alzheimer's disease: an MEG-based multiplex network study *Brain* **140** 1466–85
[21] de Domenico M, Sasai S and Arenas A 2016 Mapping multiplex hubs in human functional brain networks *Front. Neurosci.* **10** 326
[22] Gomez S, Diaz-Guilera A, Gomez-Gardenes J, Perez-Vicente C J, Moreno Y and Arenas A 2013 Diffusion dynamics on multiplex networks *Phys. Rev. Lett.* **110** 028701
[23] Newman M 2018 *Networks* (Oxford: Oxford University Press)
[24] Radicchi F and Arenas A 2013 Abrupt transition in the structural formation of interconnected networks *Nat. Phys.* **9** 717–20
[25] Martín-Hernández J, Wang H, Van Mieghem P and D'Agostino G 2014 Algebraic connectivity of interdependent networks *Physica* A **404** 92–105
[26] Shakeri H, Albin N, Darabi Sahneh F, Poggi-Corradini P and Scoglio C 2016 Maximizing algebraic connectivity in interconnected networks *Phys. Rev.* E **93** 030301

[27] Tewarie P *et al* 2016 Integrating cross-frequency and within band functional networks in resting-state MEG: a multi-layer network approach *Neuroimage* **142** 324–36

[28] Buldú J M and Porter M A 2018 Frequency-based brain networks: from a multiplex framework to a full multilayer description *Netw. Neurosci.* **2** 418–41

[29] Pereda E, Quiroga R Q and Bhattacharya J 2005 Nonlinear multivariate analysis of neurophysiological signals *Prog. Neurobiol.* **77** 1–37

[30] Echegoyen I, López-Sanz D, Martínez J H, Maestú F and Buldú J M 2020 Permutation entropy and statistical complexity in mild cognitive impairment and Alzheimer's disease: an analysis based on frequency bands *Entropy* **22** 116

[31] Stam C J 2005 Nonlinear dynamical analysis of EEG and MEG: review of an emerging field *Clinical Neurophysiol.* **116** 2266–301

[32] Garcés P, López-Sanz D, Maestú F and Pereda E 2017 Choice of magnetometers and gradiometers after signal space separation *Sensors* **17** 12

[33] Van Veen B D, Van Drongelen W, Yuchtman M and Suzuki A 1997 Localization of brain electrical activity via linearly constrained minimum variance spatial filtering *IEEE Trans. Biomed. Eng.* **44** 867–80

[34] Desikan R S *et al* 2006 An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest *NeuroImage* **31** 968–80

[35] Freedman D and Diaconis P 1981 On the histogram as a density estimator: L2 theory *Z. Wahrscheinlichkeitstheor. Verwandte Geb.* **57** 453–76

[36] Cohen M X 2014 *Analyzing Neural Time Series Data: Theory and Practice* (Cambridge, MA: MIT Press)

[37] Maris E and Oostenveld R 2007 Nonparametric statistical testing of EEG- and MEG-data *J. Neurosci. Methods* **164** 177–90

[38] Gilson M, Tauste Campo A, Chen X, Thiele A and Deco G 2017 Nonparametric test for connectivity detection in multivariate autoregressive networks and application to multiunit activity data *Netw. Neurosci.* **1** 357–80

[39] Barrat A, Barthelemy M, Pastor-Satorras R and Vespignani A 2004 The architecture of complex weighted networks *Proc. Natl Acad. Sci.* **101** 3747–52

[40] Latora V and Marchiori M 2001 Efficient behavior of small-world networks *Phys. Rev. Lett.* **87** 198701

[41] Watts D J and Strogatz S H 1998 Collective dynamics of 'small-world' networks *Nature* **393** 440–2

[42] Dijkstra E W 1959 A note on two problems in connexion with graphs *Numer. Math.* **1** 269–71

[43] Stam C J and van Straaten E C W 2012 The organization of physiological brain networks *Clinical Neurophysiol.* **123** 1067–87

[44] Lakens D 2013 Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and anovas *Front. Psychol.* **4** 863

[45] Thompson B 2007 Effect sizes, confidence intervals, and confidence intervals for effect sizes *Psychol. Schs.* **44** 423–32

[46] Cohen J 2013 *Statistical Power Analysis for the Behavioral Sciences* (New York: Academic)

[47] Dobson A J and Barnett A G 2018 *An Introduction to Generalized Linear Models* (Boca Raton, FL: CRC Press)

[48] McCullagh P and Nelder J A 2019 *Generalized Linear Models* 2nd edn (Boca Raton: Routledge)

[49] Dunteman G H and Moon-Ho R Ho 2006 *An Introduction to Generalized Linear Models* vol 145 (Thousand Oaks, CA: Sage Publications)

[50] Papo D, Zanin M, Martínez J H and Buldú J M 2016 Beware of the small-world neuroscientist! *Front. Hum. Neurosci.* **10** 96

[51] Rubinov M and Sporns O 2010 Complex network measures of brain connectivity: uses and interpretations *Neuroimage* **52** 1059–69

[52] Von Stein A and Sarnthein J 2000 Different frequencies for different scales of cortical integration: from local gamma to long range alpha/theta synchronization *Int. J. Psychophysiol.* **38** 301–13

[53] Gross J *et al* 2013 Good practice for conducting and reporting meg research *Neuroimage* **65** 349–63

[54] Yu M 2020 Benchmarking metrics for inferring functional connectivity from multi-channel EEG and MEG: a simulation study *Chaos* **30** 123124

[55] Colclough G L, Woolrich M W, Tewarie P K, Brookes M J, Quinn A J and Smith S M 2016 How reliable are MEG resting-state connectivity metrics? *Neuroimage* **138** 284–93

[56] Kantarci K *et al* 2005 DWI predicts future progression to Alzheimer disease in amnestic mild cognitive impairment *Neurology* **64** 902–4

[57] Cui Y *et al* 2011 Identification of conversion from mild cognitive impairment to Alzheimer's disease using multivariate predictors *PLoS One* **6** e21896

[58] Westman E, Muehlboeck J-S and Simmons A 2012 Combining MRI and CSF measures for classification of Alzheimer's disease and prediction of mild cognitive impairment conversion *Neuroimage* **62** 229–38

[59] Gaser C, Franke K, Klöppel S, Koutsouleris N and Sauer H (Alzheimer's Disease Neuroimaging Initiative) 2013 BrainAGE in mild cognitive impaired patients: predicting the conversion to Alzheimer's disease *PloS one* **8** e67346

[60] Brueggen K *et al* 2015 Basal forebrain and Hippocampus as predictors of conversion to Alzheimer's disease in patients with mild cognitive impairment—a multicenter DTI and volumetry study *J. Alzheimer's Dis.* **48** 197–204

[61] López M E, Turrero A, Cuesta P, Rodríguez-Rojo I C, Barabash A, Marcos A, Maestú F and Fernández A 2020 A multivariate model of time to conversion from mild cognitive impairment to Alzheimer's disease *GeroScience* **42** 1715–32