Canonical Hidden Markov Model Networks for Studying M/EEG

Chetan Gohil[1], Rukuang Huang[1], Cameron Higgins[1], Mats W.J. van Es[1], Andrew J. Quinn[2], Diego Vidaurre[1,3], and Mark W. Woolrich[1]

[1]Oxford Centre for Human Brain Activity, Oxford Centre for Integrative Neuroimaging, Department of Psychiatry, University of Oxford, Oxford, UK
[2]Centre for Human Brain Health, School of Psychology, University of Birmingham, Birmingham, UK
[3]Center of Functionally Integrative Neuroscience, Department of Clinical Medicine, Aarhus University, Aarhus, Denmark

October 17, 2025

Abstract

Dynamic brain networks identified in magneto/encephalography (M/EEG) recordings provide new insights into human brain activity. One established method uses Hidden Markov Models (HMMs) and has been shown to infer reproducible, fast-switching brain networks in a variety of cognitive and disease conditions. Often these studies are done on small bespoke (*boutique*) datasets ($N < 100$) and analysed in isolation of other M/EEG datasets. Instead of training a new model for each boutique study, which is computationally expensive, we propose the use of a *canonical HMM*. This provides a common reference through which different studies can be described using the same set of networks. We provide HMMs for a range of model orders (4-16 states) in parcellated source space and sensor space. These HMMs were trained on 1849 MEG recordings (N=621, 18-88 years old, 194 hours), capturing population variability in both rest and task data. We illustrate applications of this canonical HMM approach in parcellated source space using boutique MEG and EEG datasets. Applying the canonical HMM in parcel space requires the boutique dataset to be preprocessed and source reconstructed in the same way as the canonical HMM training data. Apply the canonical HMM in the sensor space requires the same sensor layout and preprocessing as the canonical HMM training data. The canonical HMMs have been made publicly available as an open-access resource, providing sets of canonical brain networks that can be used to compare individuals within and across a range of datasets.

Keywords: Electrophysiology; MEG; EEG; Neuronal Oscillations; Hidden Markov Model; Machine Learning; Dynamic Functional Connectivity; Functional Networks

Highlights

- Instead of training a new HMM for every study, which is computationally expensive and can lead to less reliable results, we propose a *canonical HMM* instead.
- The canonical HMM provides a common reference through which different studies can be described using the same set of networks.
- The canonical HMM was learnt from 1849 MEG recordings (194 hours of data; resting state and task) in participants aged 18–88 years.
- We demonstrate application of the canonical HMM to three boutique studies: an Alzheimer's MEG dataset, an eyes open vs closed EEG dataset, and a sensor-level MEG dataset (without the need for source reconstruction).

- We release the canonical HMM (for varying model orders) as an open-access resource for M/EEG research.

1 Introduction

Cortical activity, as measured by electrophysiological recordings, can be described using transient, re-occurring large-scale networks of phase-locking (coherent) activity (Vidaurre et al., 2018b). These networks, which exhibit fast dynamics ($\sim$ 100 ms) and correspond to transient oscillatory activity across different regions (*spatio-spectral patterns*), are typically estimated using a Hidden Markov Model (HMM) applied to source-reconstructed magneto/encephalography (M/EEG) data (Rabiner and Juang, 2003; Bishop and Nasrabadi, 2006; Cho et al., 2024). The networks and their dynamics are learnt directly from the data, usually without the user specifying the spatio-spectral patterns. See (Gohil et al., 2024a) for a description of the HMM approach applied to M/EEG data and Section 2 for further details.

The transient network perspective provided by the HMM has been used to gain an understanding of cognition (Guichet et al., 2025; Rossi et al., 2023; Fauchon et al., 2022; Higgins et al., 2021) and disease (Kohl et al., 2024; Seedat et al., 2023; Sitnikova et al., 2018). To gain specific insights, HMM analysis is often performed on bespoke datasets; for example, by looking at the network-level responses to tasks (Gohil et al., 2024b), or by contrasting network characteristics between healthy and disease groups (Kohl et al., 2024; Seedat et al., 2023; Sitnikova et al., 2018). In this work, we will refer to these as *boutique* datasets. However, the relatively small size of these datasets (typically less than $\sim$100 participants) often reduces the detail and reliability of the estimated large-scale networks and their dynamics (discussed further in Section 4.6).

The HMM has previously been applied successfully to several independent boutique M/EEG datasets (Vidaurre et al., 2018b; Cho et al., 2024; Gohil et al., 2024a; Guichet et al., 2025; Rossi et al., 2023; Fauchon et al., 2022; Higgins et al., 2021; Kohl et al., 2024; Seedat et al., 2023; Sitnikova et al., 2018; Gohil et al., 2024b; van Es et al., 2023; Gohil et al., 2022; Quinn et al., 2018; Vidaurre et al., 2018a, 2016). Broadly, the HMM finds a consistent set of networks across a variety of tasks and at rest. A similar consistency has also been observed in the functional networks found using functional magnetic resonance imaging (fMRI) (Smith et al., 2009). There is also generally good correspondence between networks found in MEG and EEG (Cho et al., 2024) and between healthy and diseased populations (Kohl et al., 2024).

Given the commonality, we propose adopting a common reference (*canonical*) model[1], rather than training a new HMM for each isolated dataset. In this work, we present a canonical HMM pre-trained on a large MEG dataset, which provides a set of networks that describe re-occurring spatio-spectral patterns of activity in electrophysiological (MEG and EEG) data. This approach has several key benefits. It provides a basis set of canonical networks that can be used to compare different studies, including individuals or groups, in a common reference space. It also improves the stability of network inference; and lowers the computational resources required to perform an HMM analysis on new data.

In this work, we have developed a canonical HMM framework that includes a set of pre-trained models with a different number of states (4 to 16) in both parcellated source space and in MEG sensor space. The main constraint in applying these models is that the new datasets have the same spatial layout/parcellation and are subject to similar preprocessing/source reconstruction to that used in the canonical HMM (described in Section 2 and discussed further in Section 4).

To pre-train the canonical HMM, we have used the Cam-CAN dataset (see Section 2.1.1), which contains recordings from a large healthy population (N=621, 1849 rest and task sessions). This is an ideal dataset due to its size (one of the largest currently available) and its wide age

[1]Similar to the foundation model approach used in the field of deep learning (Bommasani, 2021; Wang et al., 2025).

range of participants (18-88 years old). The canonical HMMs, along with example code for applying them to new datasets, have been made available as a public resource, see Section 5 for further details.

2 Materials and Methods

2.1 Datasets

2.1.1 Training MEG Dataset: Cam-CAN (Resting State and Task, 1849 Sessions)

To train the canonical HMM, we use a group of 621 healthy participants aged between 18 and 88 years from the Cam-CAN (Cambridge Centre for Ageing and Neuroscience) dataset (Shafto et al., 2014; Taylor et al., 2017). Each participant has multiple MEG recordings: an eyes closed resting-state session; a sensorimotor task session; and a passive visual/auditory task session. A total of 1849 sessions were studied in this work. The demographics of the participants studied are shown in Figure S1A. Further information regarding the dataset, protocols, tasks and the MEG/MRI acquisition is provided in (Shafto et al., 2014; Taylor et al., 2017).

2.1.2 Example MEG Boutique Dataset: BioFIND (Resting State, 82 Sessions)

In this work, we studied a subset of the BioFIND dataset (Hughes et al., 2019; Vaghari et al., 2022), which aims to identify biological markers for the early detection of Alzheimer's disease. We analysed data from the Cambridge site and excluded those with missing electrooculography recordings or sMRIs. This resulted in 41 MCI participants and 41 healthy controls, each with a 4-13 minute eyes closed resting-state MEG recording. The demographics of the participants studied are shown in Figure S1B. Further information regarding this dataset, the participants, protocols and the MEG/MRI acquisition can be found in (Hughes et al., 2019; Vaghari et al., 2022).

2.1.3 Example EEG Boutique Dataset: LEMON (Resting State, 108 Sessions)

In this work, we studied a subset of the LEMON (Leipzig Study for Mind-Body-Emotion Interactions) dataset (Babayan et al., 2019) excluding any participant in the full dataset that did not have a structural MRI or Polhemus head shape points. This resulted in 108 healthy participants. Medium density (61 channel) EEG data were recorded from each participant in a resting-state session where they alternated between periods of eyes closed and eyes open every 30 seconds. The demographics of the participants studied are shown in Figure S1C. Further information regarding this dataset, protocols and the MEG/MRI acquisition can be found in (Babayan et al., 2019).

2.2 M/EEG Preprocessing, Source Reconstruction and Parcellation

Each dataset was processed using the `osl-ephys` toolbox (van Es et al., 2025; Quinn et al., 2022). For each session, we applied the following steps[2]:

1. **Filtering and downsampling**. First, we bandpass filter the data (forward-backward fifth-order IIR Butterworth) between 0.5 and 125 Hz, then apply notch filters at 50 Hz and 100 Hz to remove mains artefacts. For Cam-CAN an additional spike artefact at 88 Hz was also removed using a notch filter. Following this, the data were downsampled to 250 Hz.

[2]For Cam-CAN (Elekta MEG), we applied these steps to the publicly available MaxFiltered data. For LEMON, we used the publicly available raw sensor recordings.

2. **Automated bad segment/channel detection**. Bad segments and channels can have a detrimental effect on independent component analysis (ICA) denoising (next step). Therefore, we annotated significantly high variance (p-value < 0.1) using an automated generalised-extreme studentised deviate (G-ESD) algorithm (Rosner, 1983). Bad segment and channel detection was applied separately to the different sensor types (magnetometers, gradiometers, EEG).

3. **ICA artefact removal**. After excluding the bad segments and channels, we applied FastICA (Hyvarinen, 1999) to the sensor-level MEG data, decomposing the signal into 64 independent components. For the LEMON dataset, we decomposed the sensor-level EEG into 30 independent components. Based on the correlation of the ICA component time courses with the EOG/ECG electrodes (0.9 threshold), independent components were marked as ocular/cardiac noise and removed from the MEG data.

4. **Bad channel interpolation**. Bad channels in the ICA-cleaned M/EEG data were replaced using a spherical spline interpolation (Perrin et al., 1989). This was the final step in preprocessing the sensor-level MEG data.

5. **EEG re-referencing**. The EEG data were re-referenced by subtracting the average across all sensors excluding those marked as bad. This was only performed for the LEMON dataset. This was the last step in preprocessing the sensor-level EEG data.

6. **Coregistration**. Scalp, inner skull, and outer brain surface were extracted from the sMRI using FSL BET (Smith, 2002; Jenkinson, 2005). The nose was not included in the scalp surface because the sMRI images were defaced. The scalp surface was then used in OSL RHINO (Quinn et al., 2022) to coregister the sMRI to the M/EEG data based on matching the scalp to a set of digitised Polhemus head shape points and fiducials.

7. **Forward modelling**. The *forward model* gives us the expected signal at each sensor (leadfield) for a dipole placed at a particular location inside the brain. The electric/magnetic field from this dipole transverses different tissue types, which have different electromagnetic properties. It is common in MEG source localisation to model the changes in magnetic properties using a single boundary separating the volume inside the brain from outside. In the current work, we used a boundary element model (Henson et al., 2009), using the inner skull surface as the boundary. For EEG, we need to model the electrical conductivity for the different tissue types and used three layers in the forward model: the inner skull, outer skull and outer head skin. We calculated the leadfields for a dipole placed in the x, y and z-direction on an 8 mm isotropic grid inside the volume bound by the inner skull.

8. **Bandpass filtering and bad segment removal**. We applied a bandpass filter to the preprocessed M/EEG data (fifth-order IIR Butterworth) to focus on the frequency range of interest. Depending on the dataset we used a different frequency range: 1-80 Hz for Cam-CAN and BioFIND and 1-45 Hz for LEMON. Following this, we removed bad segments from the data to ensure these did not impact on the quality of the source localisation.

9. **Volumetric beamforming**. We employed a beamformer (**?**Hillebrand and Barnes, 2005) to source localise the MEG data. The beamformer estimates source activity (at a location inside the brain) as a linear combination of sensor activity. In this step, we calculated the weights for the linear combination, referred to as the *beamformer weights*. We used a unit-noise-gain invariant Linearly Constrained Minimum Variance (LCMV) beamformer. This type of beamformer controls for the depth bias by normalising the beamformer weights (Tait et al., 2021).

   To calculate the beamformer weights, we need to provide the leadfields from the forward model, the covariance of noise (signal measured by the sensors unrelated to brain activity)

and the covariance of the data (signal measured by the sensors containing both noise and brain activity). For the noise covariance, we used a diagonal matrix containing the variance of each sensor type. We filled the diagonal elements corresponding to each sensor type with the average (temporal) variance across the sensors of that type. For the data covariance, we used the covariance of the data from step 8 reduced to a rank of 60 for the MEG datasets and 54 for the EEG dataset (this regularisation improves the robustness of the beamformer).

The beamformer calculates the activity in the x, y and z-direction at each grid point used in the leadfield calculation. This is reduced to a single value by projecting the x, y, z activity onto the axis that maximises power of at that location. After beamforming, we are left with the activity at each grid point inside the brain, referred to as a *voxel*.

10. **Parcellation**. Often, we have a very large number of voxels, so we parcellate the data into particular regions of interest (ROIs). In the current work, we used an anatomically defined 52-ROI parcellation. See (Kohl et al., 2023) for a description of the parcellation. Parcel time courses were obtained by applying PCA to the (demeaned) voxel time courses assigned to a parcel and taking the first principal component. PCA is preferred over simply taking the mean across voxels because there is a sign ambiguity in the voxel time courses due to the beamformer.

11. **Orthogonalisation**. The source localisation of M/EEG data suffers from spatial leakage where the uncertainty in estimating source activity leads to nearby locations in source space exhibiting highly correlated activity. These correlations can be misinterpreted as functional connectivity (so called 'ghost interactions' (Colclough et al., 2015)). To minimise the impact of spatial leakage, we used the symmetric orthogonalisation technique proposed in (Colclough et al., 2015), which removes all zero-lag correlations from the parcel data. This can be seen as a conservative approach as we have likely removed some genuine zero-lag functional connectivity from the data.

12. **Sign flipping**. Unfortunately, there is an ambiguity in the sign of each parcel time course (due to the ambiguity of the dipole direction in the source reconstruction and due to the PCA used in the parcellation). This means the sign of each parcel time course can be misaligned across subjects. After we have the orthogonalised parcel time courses for each subject, we calculate the covariance of time-delay embedded parcel data (see Section 2.3.1) for each subject and flatten the upper triangle into a vector. Comparing the correlation of this vector for each pair of subjects, we selected the median subject (highest average correlation) as a template. Finally, we match the sign of each subject's orthogonalised parcel time course to the template with the random search sign-flip algorithm described in (Vidaurre et al., 2016) using the correlation between flattened time-delay embedded covariance matrices as the metric for alignment.

13. **Standardisation**. Finally, we *standardise* the data by subtracting the mean and dividing by the variance for each parcel individually.

Note, in the current work all references to the 'parcel time courses' are to the standardised, sign-flipped, orthogonalised parcel time courses.

2.3 Parcel-Level Canonical HMM

Here, we describe how we developed a parcel-level canonical HMM based on the Cam-CAN dataset.

2.3.1 Data Preparation

Before training the HMM on Cam-CAN, the parcel time courses were prepared by performing TDE, which augments the data with extra channels corresponding to time-lagged versions of the original data. In the current work, we used ±7 lags (±28 ms), resulting in 780 channels for the TDE data. PCA was then applied at the group level to reduce the TDE data down to 120 channels. Finally, we (temporally) standardised (demeaned and divided by the standard deviation) the TDE-PCA data. These were the data used to train the HMM. These steps are summarised in Figure S2.

2.3.2 Model Training

The HMM (Rabiner and Juang, 2003; Bishop and Nasrabadi, 2006) is a generative model typically used to study time series data. At each time point, t, there is an underlying categorical hidden state, $\theta_t \in \{1, ..., K\}$, where K is the number of states. The observed time series, $\boldsymbol{x}_t$, is generated using a multivariate Normal distribution based on the hidden state:

$$\boldsymbol{x}_t | (\theta_t = k) \sim \mathcal{N}(0, \boldsymbol{D}_k), \tag{1}$$

where $\theta_t = k$ is the state at time point t and $\boldsymbol{D}_k$ is a *state covariance*. Note, we force the mean to be zero to focus on modelling dynamics in the covariance, which contains the connectivity information we are interested in. Dynamics are governed by transitions in the hidden state. Note, it is assumed that the probability of a transition at time point t only depends on the state at the previous time point θ_{t-1} (this is the Markovian constraint).[3] Each pairwise state transition probability is contained in the *transition probability matrix*:

$$A_{i,j} = p(\theta_t = j | \theta_{t-1} = i). \tag{2}$$

When we train an HMM on data, we learn the most likely value for the state covariances, $\{\boldsymbol{D}_1, ..., \boldsymbol{D}_K\}$, and transition probability matrix, $\boldsymbol{A}$, to have generated the data. We use stochastic variational Bayes (Bishop and Nasrabadi, 2006) to do this, where we iteratively update our estimates for $\{\boldsymbol{D}_1, ..., \boldsymbol{D}_K, \boldsymbol{A}\}$ based on a random subset of the training data to minimise the variational free energy. We use the Baum-Welch algorithm (Bishop and Nasrabadi, 2006) to calculate the (*posterior*) probability of each state being active at each time point, $q(\theta_t)$, for each session based on our estimates for $\{\boldsymbol{D}_1, ..., \boldsymbol{D}_K, \boldsymbol{A}\}$. The posterior, $q(\theta_t)$, is a (states, time) array for each session. The training of an HMM is summarised in Figure S3.

Hyperparameters. We used the HMM implemented in the `osl-dynamics` (Gohil et al., 2024a) toolbox. In order to train an HMM, we need to specify a few hyperparameters. An important hyperparameter is the number of states. In the current work, we trained canonical HMMs with 4-16 states, which is in line with previous HMM M/EEG studies (Vidaurre et al., 2018b; Cho et al., 2024; Gohil et al., 2024a; Guichet et al., 2025; Rossi et al., 2023; Fauchon et al., 2022; Higgins et al., 2021; Kohl et al., 2024; Seedat et al., 2023; Sitnikova et al., 2018; Gohil et al., 2024b; van Es et al., 2023; Gohil et al., 2022; Quinn et al., 2018; Vidaurre et al., 2018a, 2016). Other hyperparameters include the batch size, sequence length and learning rates. These had little impact on the HMM inference. The values used are summarised in Table S1.

Run-to-run variability. The final estimates for $\{\boldsymbol{D}_1, ..., \boldsymbol{D}_K, \boldsymbol{A}\}$ can be sensitive to the initial values used at the start of training. The typical approach for overcoming this is to train several models from scratch starting from random initialisation and picking the one with the lowest final variational free energy (deemed the best description of the data) for the subsequence

[3] Note, the Markovian constraint does not prevent us from inferring long-range temporal dependencies in real data. Although the HMM generative model only considers the previous time point, the real data may have been generated from a process with a longer memory. This information is contained in the observed data. In calculating the state probabilities (posterior) we can infer long-range dynamics in the states from the data directly despite the short memory of the HMM.

analysis. Historically, this has produced very reproducible results (Gohil et al., 2024a). In our case, we analyse a particularly large dataset. This makes the HMM inference very stable, and we consistently converge on very similar estimates for $\{\boldsymbol{D}_1, ..., \boldsymbol{D}_K, \boldsymbol{A}\}$. Despite this, we took a cautionary approach and selected the best model from a set of five runs.

2.4 Applying the Parcel-Level Canonical HMM to a Boutique Dataset

2.4.1 Data Preparation

Before we can apply the canonical HMM to a new boutique dataset, we must prepare the parcel time courses in a similar way to how the Cam-CAN data was prepared when training the canonical HMM, i.e. using the steps described in Section 2.2. Importantly, we use the same parcellation and a matching sampling frequency (250 Hz). We also strongly recommend using the template session from the Cam-CAN dataset when doing sign flipping on the new dataset, and using the PCA components from Cam-CAN when applying PCA following TDE.

2.4.2 Inference (Estimating State Probabilities)

We use the canonical HMM pre-trained on Cam-CAN to perform inference on a new dataset. This is shown in Figure S4. In doing this, we obtain a state probability time course for each session in the new dataset.

2.4.3 Dual Estimation: Individualised Post-Hoc Analysis

We use the state probability time course and parcel time course for each session to perform post-hoc analysis, see Figure S5. This process provides session-specific estimates for network properties and is referred to as *dual estimation* (Vidaurre et al., 2017).

State power/coherence spectra. Here, we combine the state probability time course with the original parcel data (pre-TDE-PCA) to estimate the spectral properties (PSD and coherence) of each state using a multitaper (Babadi and Brown, 2014). This involves performing the following steps for each session and state:

1. Temporally standardise (demean and divide by the standard deviation) the parcel time courses and multiply by the state time course.

2. Calculate P/CSD for each parcel/pair of parcels using a multitaper (2 s window, 0% overlap, 7 DPSS tapers, 4 Hz time half bandwidth). This is the standard approach/settings used for HMMs trained on TDE-PCA data (Gohil et al., 2024a; Vidaurre et al., 2018b, 2016).

3. Use $C_{ij} = \frac{|CSD_{ij}|}{\sqrt{PSD_i \, PSD_j}}$, where i and j correspond to parcels, to calculate the coherence from the P/CSD.

4. The amplitude of the PSD is biased by the amount of time each state is active due to the multiplication in step 1. We account for this by scaling the PSD by one over the fractional occupancy of the state.

State power maps. The multitaper results in a [sessions, states, parcels, frequencies] array containing the PSDs. Averaging over the frequency dimension results in a [sessions, states, parcels] array, which we refer to as the *state power map* for each session. In the visualisation of the group-averaged power maps, we displayed each state's power map separately and used the average (parcel-specific) power across states as the reference.

State coherence networks. The multitaper results in a [sessions, states, parcels, parcels, frequencies] array for the coherences. Similar to the state power maps, we reduce the frequency

dimension by taking the average over the frequency dimension. This results in a [sessions, states, parcels, parcels] array, which represents the *state coherence network* for each session. We displayed the group-average coherence network for each state individually using the (edge-specific) average across states as the reference and thresholding the top 3% of edges irrespective of sign.

Summary statistics for dynamics. We calculated a *state time course* by taking the most probable state (*maximum a posteriori probability estimate*) at each time point. From the state time course, we calculate metrics that summarise the dynamics within a session. For each session and state, we calculated:

- Fractional occupancy: the fraction of total time spent in a state.
- Mean lifetime (ms): the average duration a state was active.
- Mean interval (s): the average duration between successive state activations.
- Switching rate (Hz): the average number of state activations per second.

2.5 Sensor-Level Canonical HMM

Alongside the parcel-level canonical HMM, we developed a corresponding sensor-level version. Importantly, the sensor-level canonical HMM was designed to capture the same states as the parcel-level canonical HMM. This allows researchers to apply the sensor-level canonical HMM to new sensor-space dataset without the need for source reconstruction, while still being able to map to the same set of canonical networks. Here, we describe how the sensor-level canonical HMM is obtained (i.e. trained) and how to apply it to new sensor-space datasets.

2.5.1 Data Preparation

The sensor-level canonical HMMs were inferred using the MaxFiltered Cam-CAN data (Elekta, $N = 621$, rest and task). Due to the MaxFiltering the rank of the sensor-level data is reduced to 64. Therefore, the first step was to use PCA to reduce to 64 channels. The data was then prepared in the same way as the parcel time courses with TDE-PCA using ± 7 lags followed by a second (post TDE) PCA. The number of components used in the second PCA was chosen to maximise the agreement (classification accuracy) between the sensor and parcel-level inferred state time courses. The optimum number of PCA components was found to be 160.

2.5.2 Model Training

We seek a canonical HMM that captures the same states as those obtained by an HMM trained on the (prepared) parcel data directly. Therefore, to calculate the sensor-level canonical HMM we did not train a new model from scratch. Instead, we dual estimated the state covariances using the prepared *sensor-space* Cam-CAN data and the state probability time courses inferred using the *parcel-level* canonical HMM. In the dual estimation, the states were assumed to have zero mean and we used the transition probability matrix/initial state probabilities from the parcel-level canonical HMM. The result of this was group-level state canonical covariances for each HMM state in prepared sensor-space.

2.5.3 Application to Boutique Datasets

The sensor-level canonical HMM can be applied to boutique datasets in the same way as the parcel-level HMM. The only difference is that the input data must be the prepared sensor-level data and must have the same sensor layout as in the canonical HMM. Currently, the public repository for code (see Section 5) provides sensor-level canonical HMMs for Elekta MEG data.

2.6 Modelling Approaches: Conventional vs Canonical HMM

Two different modelling approaches were used in this work: a conventional approach, where a new HMM is trained from scratch on the boutique dataset; and a canonical approach, where a pre-trained HMM is used. No HMM training is required in the canonical approach. An overview of each approach is shown in Figure 1. Both approaches use the same preparation steps on the new dataset and the same dual estimation step to obtain individualised networks for each session. Note that in all results we show using the conventional approach, the run with the lowest final variational free energy from a set of 10 was used.

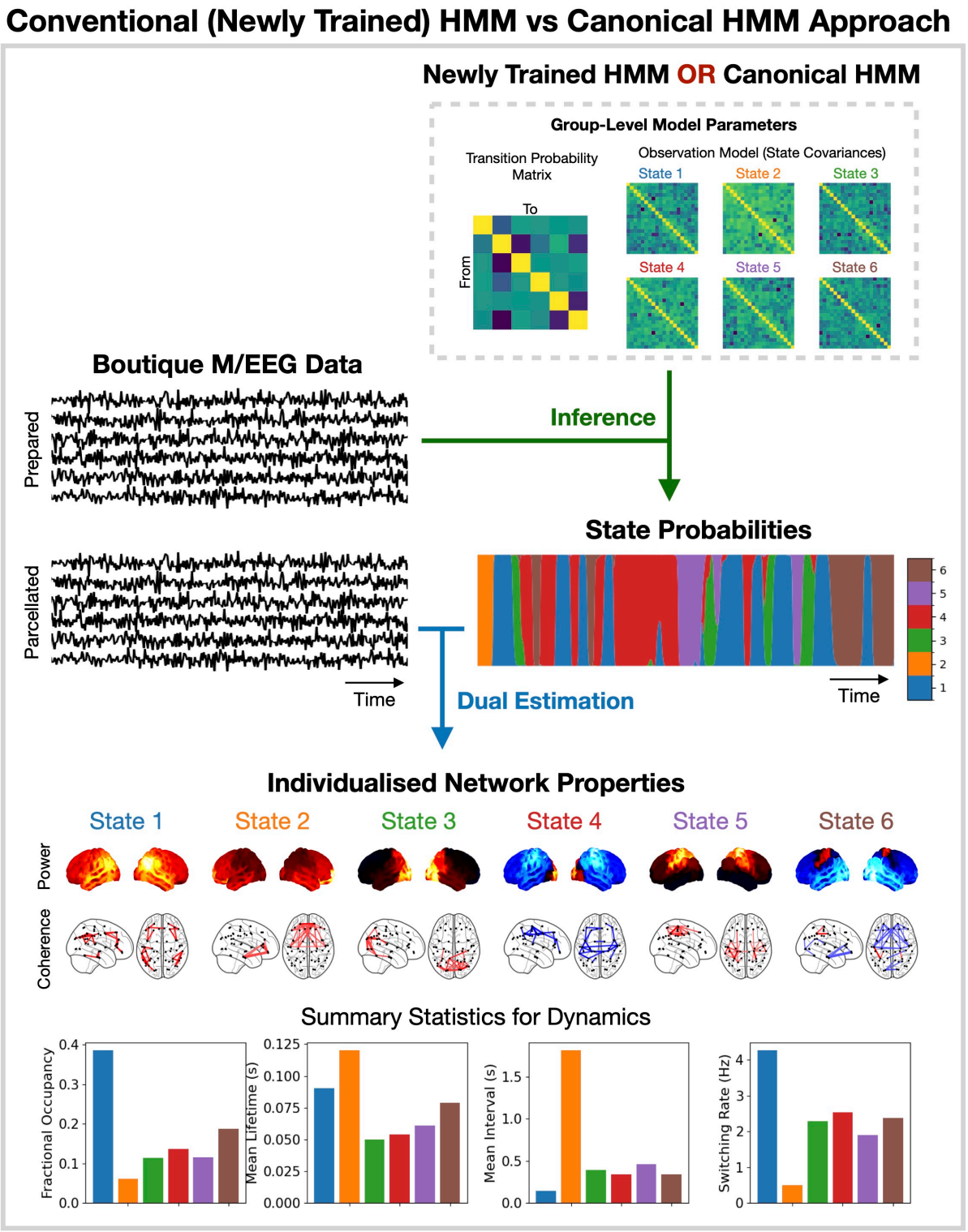

Figure 1: **Overview of HMM approaches**. A conventional HMM approach involves training a new model on the boutique dataset (usually multiple times to ensure results are robust), whereas the canonical HMM approach uses a HMM pre-trained on a large independent dataset. The HMM is used to infer the state probabilities at each time point and dual estimation is used to infer individualised network properties for each state.

3 Results

We present a canonical HMM that describes brain activity using a set of transient networks derived from the Cam-CAN dataset. We describe the spatio-spectral patterns of brain activity associated with each network (in parcel space) and summarise each network's dynamics in Section 3.1. Applying the parcel-level canonical HMM to a boutique MEG dataset in Section 3.2, we study how mild cognitive impairment, an early indicator of Alzheimer's disease, alters brain network activity. In Section 3.3, we illustrate how a parcel-level canonical HMM from MEG can be applied to parcellated EEG data. Finally, we present a sensor-level canonical HMM that can be applied to new sensor-level MEG data in Section 3.4; allowing for a more flexible application of the canonical HMM, where the user does not need to apply the same source localisation used in the parcel-level model.

3.1 Parcel-Level Canonical HMM

We start by presenting the network description provided by the parcel-level canonical HMM for a model order (number of states) of 10. Note that canonical HMMs for a range of model orders (4 to 16) have been made available on the code repository (see Section 5). The 10-state canonical HMM is shown in Figure 2. The networks were inferred on (prepared) parcel data with 52 regions of interest, sampled at 250 Hz, and band-pass filtered in the range 1-80 Hz (see Section 2.2 for the full details). Figure S7 shows the state-specific relative power maps for different narrow-band frequencies rather than the 1-45 Hz range shown in Figure 2A.

The canonical HMM was trained on 1849 MEG recording sessions (rest and task) from the Cam-CAN dataset. Each HMM state corresponds to a network of brain activity with fast transient dynamics ($\sim$ 100 ms). These networks describe recurrent spatio-spectral patterns found in a large population of healthy individuals across a wide age range (N=621, 18-88 years old). Figure 2 shows the canonical HMM networks at the group level (i.e., group-averaged). Individualised networks (for different sessions) are shown in Figure S6.

3.2 Example MEG Boutique Study

Here, we illustrate the use of a 10 state parcel-level canonical HMM in studying a boutique (i.e. bespoke) MEG dataset. We adopted a case-control analysis (comparing the networks for healthy vs. diseased groups) using the BioFIND dataset (Hughes et al., 2019; Vaghari et al., 2022), which aims to identify biomarkers for the early detection of Alzheimer's disease (Scheltens et al., 2021). This study collected MEG data (resting state, eyes closed) from individuals at high risk of developing Alzheimer's disease. These participants showed early signs of cognitive decline, referred to as *Mild Cognitive Impairment* (MCI). See Section 2.1.2 for details regarding the dataset.

The conventional approach for a case-control HMM analysis would be to first train a new HMM on the full BioFIND dataset by concatenating the data from all participants (healthy and MCI) (Gohil et al., 2024a). The HMM is unaware of which participant the data comes from (i.e. it is unsupervised) and finds the best set of network states that describes the entire dataset. This model is referred to as a *group-level HMM*. Following this, subject-level networks are estimated for each participant post-hoc based on the group-level state activations using *dual estimation* (Vidaurre et al., 2017), see Figure S5. By first training the group-level HMM, we ensure the networks are common to all individuals/groups, i.e. state 1 for one individual is the same network as state 1 for another. Figure 3A shows the group-level HMM networks adopting a conventional approach to BioFIND. This shows a set of networks similar to those typically produced from an HMM analysis (Gohil et al., 2024a).

With the canonical HMM approach, we do not need to train a new HMM on the BioFIND dataset. We can instead use the *pre-trained* canonical HMM at the desired model order. We

just need to ensure that BioFIND has been processed into the same parcellated source space as the canonical HMM. We can infer the state probabilities using the canonical HMM and proceed to the dual estimation. Figure 3B shows the group-level HMM networks from applying the canonical HMM to BioFIND. Qualitatively, we see the same networks as the conventional HMM approach (Figure 3A). However, there are a two key benefits of adopting the canonical HMM approach: first, it saves substantial computation by avoiding the need to train a new model; second, it provides a common reference for comparing different studies and datasets.

A quantitative comparison of the agreement of the inferred states using the conventional and canonical HMM is shown in Figure 3C. We see both approaches have a high agreement (0.7-0.9) in the state that is assigned at each time point. Turning to the case-control analysis, we see in Figure 3D, we see that both approaches give the same significant differences in the fractional occupancy of the states. MCI patients increase the time spent in state 2 (the frontal network) and decrease the time spent in state 5 (the sensorimotor network).

3.3 Example EEG Boutique Study

The canonical HMM was inferred on parcellated MEG data (see Section 2). However, source activity within the brain can be estimated using other neuroimaging modalities. EEG offers a cheap non-invasive measure of brain activity with high temporal resolution, albeit at some cost to the spatial resolution and signal-to-noise ratio (Blinowska and Durka, 2006). The same source reconstruction methods as MEG (see Section 2.2) can be applied to medium/high-density (64-256 electrode) EEG recordings (Cho et al., 2024; Makkinayeri et al., 2025). Once the EEG data has been source reconstructed into the same parcel space, we can apply the canonical HMM approach to identify MEG-like networks using EEG. Here, we use the LEMON dataset and compare a conventional HMM approach with the canonical HMM approach.

The conventional HMM approach, i.e. training a new HMM from scratch on the (prepared) parcellated EEG data, is studied in (Cho et al., 2024). Figure 4A shows the spectral properties and dynamics this produces. We see training a new HMM on parcellated EEG data identifies similar networks to parcellated MEG, see (Cho et al., 2024) for a quantitative comparison. However, a few networks that are often identified in MEG are not found: the sensorimotor networks and the temporal networks (states 5, 7, 8 in Figure 2).

When we adopt the canonical HMM approach, we can clear identify the states missing from the conventional analysis (Figure 4B). Figure 4C shows the agreement between the state time courses from each modelling approach. We see states 5 and 7 from the conventional model do not correspond very well with any of the canonical HMM networks, suggesting they represent dataset-specific features or noise. The resting-state recordings in the LEMON dataset alternated between 30 second periods of eyes open and closed conditions. In Figure 4D, we contrast the fractional occupancy of each state in each condition. States 5 and 7 from the conventional model do not discriminate between eyes open and closed conditions further suggesting that they reflect noise.

3.4 Sensor-Level MEG Canonical HMM

Applying a parcel-level canonical HMM to a boutique dataset requires the data to have been processed in the same way as the training data for the canonical model (including source reconstruction and parcellation). Example scripts for how to do this have been provided, see Section 5. A useful alternative approach would be to apply the HMM at the sensor-level, thus avoiding the need for source reconstruction. This facilitates broader applications in scenarios where source reconstruction is not feasible, such as in real-time network detection. It was recently shown in EEG that a reasonable agreement in state classification can be achieved between HMMs inferred at the sensor and parcel level (Cooray et al., 2024).

We developed a canonical HMM that can be applied to sensor-level MEG data, albeit with the requirement that the data have the same sensor layout as the Cam-CAN dataset, i.e. correspond to an Elekta MEG scanner. Details of how the sensor-level canonical HMM was derived are provided in Section 2.5. Notably, the sensor-level canonical HMM was designed to capture the same states as the parcel-level canonical HMM. As with the source-level canonical HMMs, we provide sensor-level canonical HMMs for different model orders (4-16 states).

Figure 5A shows the relative power maps and PSDs for the 10-state sensor-level canonical HMM at the sensor level. Qualitatively, we see a good correspondence between the sensor-level and parcel-level power maps. Note that, by virtue of the sensor-level canonical HMM being obtained via the parcel-level canonical HMM, the states are naturally in the same order.

Figure 5B shows the inferred state probabilities when applying the canonical HMMs at the sensor (top) and parcel (bottom) levels. Quantifying the agreement using a confusion matrix (which counts the number of correct and incorrect classifications between the sensor and source-level state assignment) in Figure 5C, we see the inferred states agree very well. In particular, states 1 and 10 can be identified with a high accuracy at the sensor level: approximately 0.8 vs. a chance accuracy of 0.1.

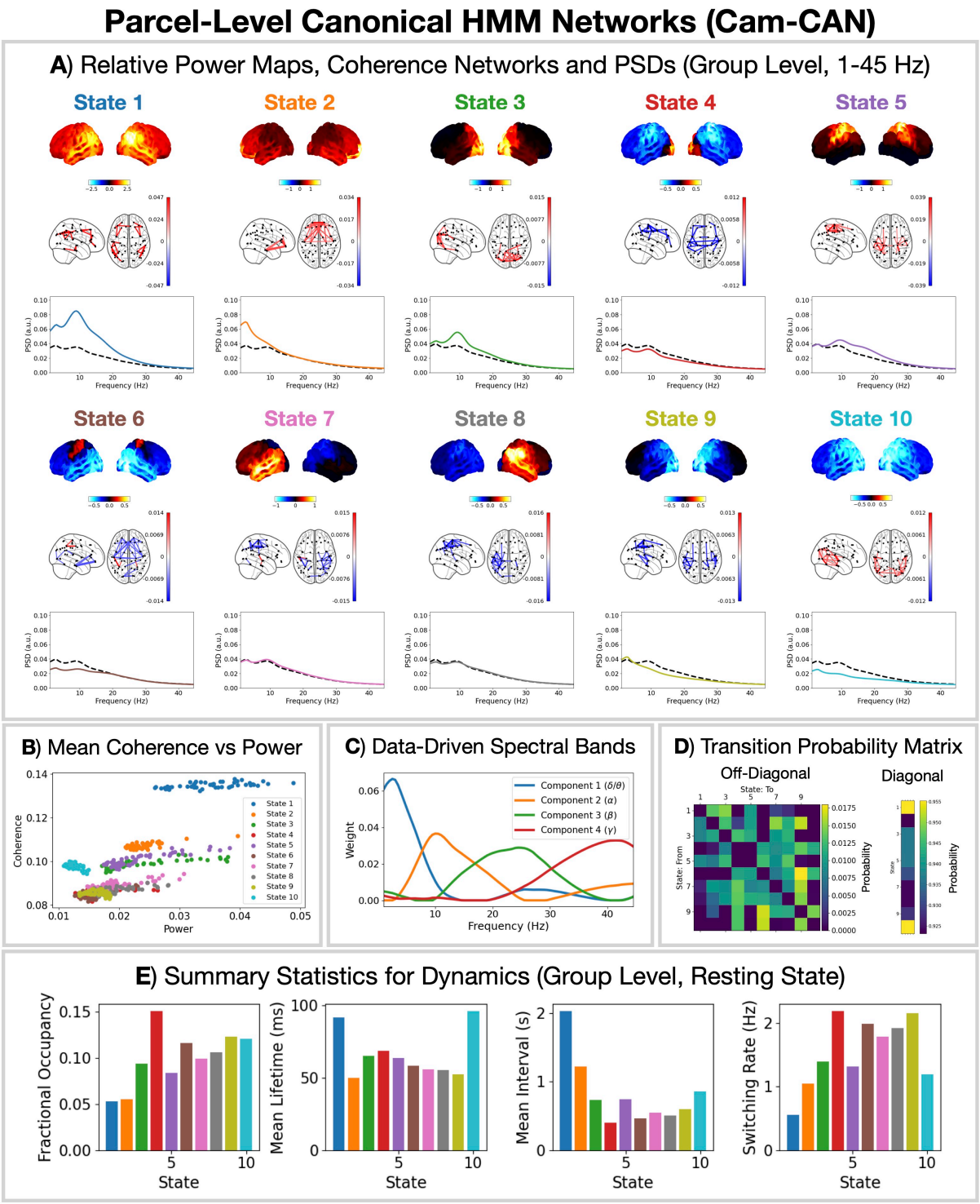

Figure 2: **10 state parcel-level canonical HMM networks inferred from Cam-CAN MEG data (N=621, rest and task)**. A) Relative state power maps, coherence networks (top 3%) and PSDs. The solid line shows the state PSD averaged over regions (parcels) and the dashed line shows the PSD averaged across states. B) Mean coherence vs power for each parcel averaged over subjects. C) Non-negative matrix factorisation components calculated from a matrix of stacked subject/state-specific coherence matrices, which indicate the networks correspond to activity in the canonical frequency bands used to study electrophysiological data (δ, θ, α, β, γ). D) State transition probability matrix. E) State summary statistics for dynamics: fractional occupancy; mean lifetime; mean interval; and switching rate.

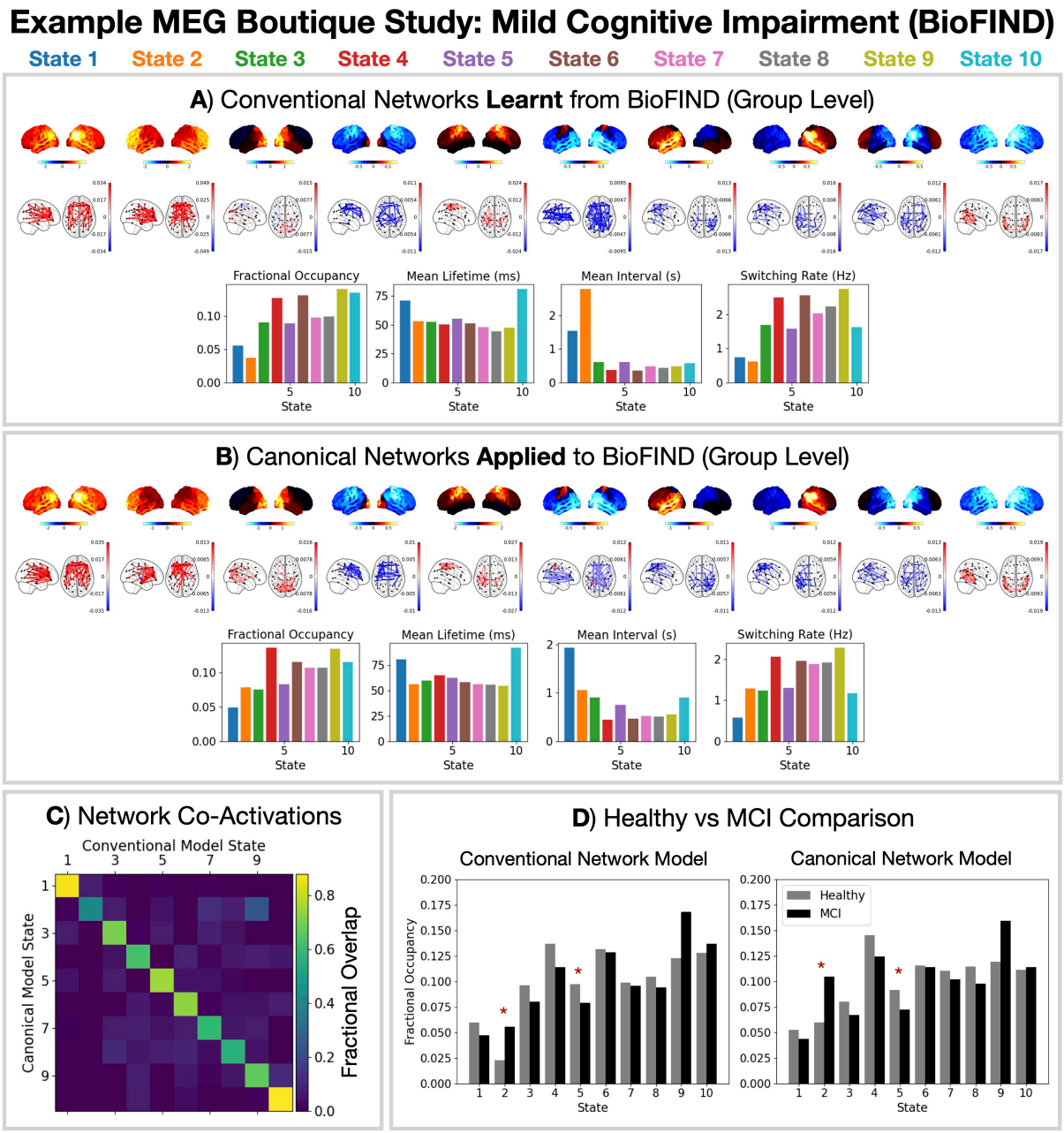

Figure 3: **Applying the canonical HMM approach provides a common set of networks that can be compared across studies and reduces the computational resources needed**. A conventional (A) and canonical (B) HMM modelling approach was applied to the BioFIND dataset. Relative state power maps (1-45 Hz, top), coherence networks (top 3%, middle) and summary statistics for dynamics (bottom) are shown. State activations from each approach show good agreement (C) and the same differences between mild cognitive impairment (MCI) and healthy control groups (D). Group differences were calculated using a General Linear Model with age and sex as confounds. Row-shuffle permutations were used for statistical significance testing with the maximum t-statistic across all features and states being used to control for multiple comparisons. The asterisk ($*$) indicates a p-value < 0.05.

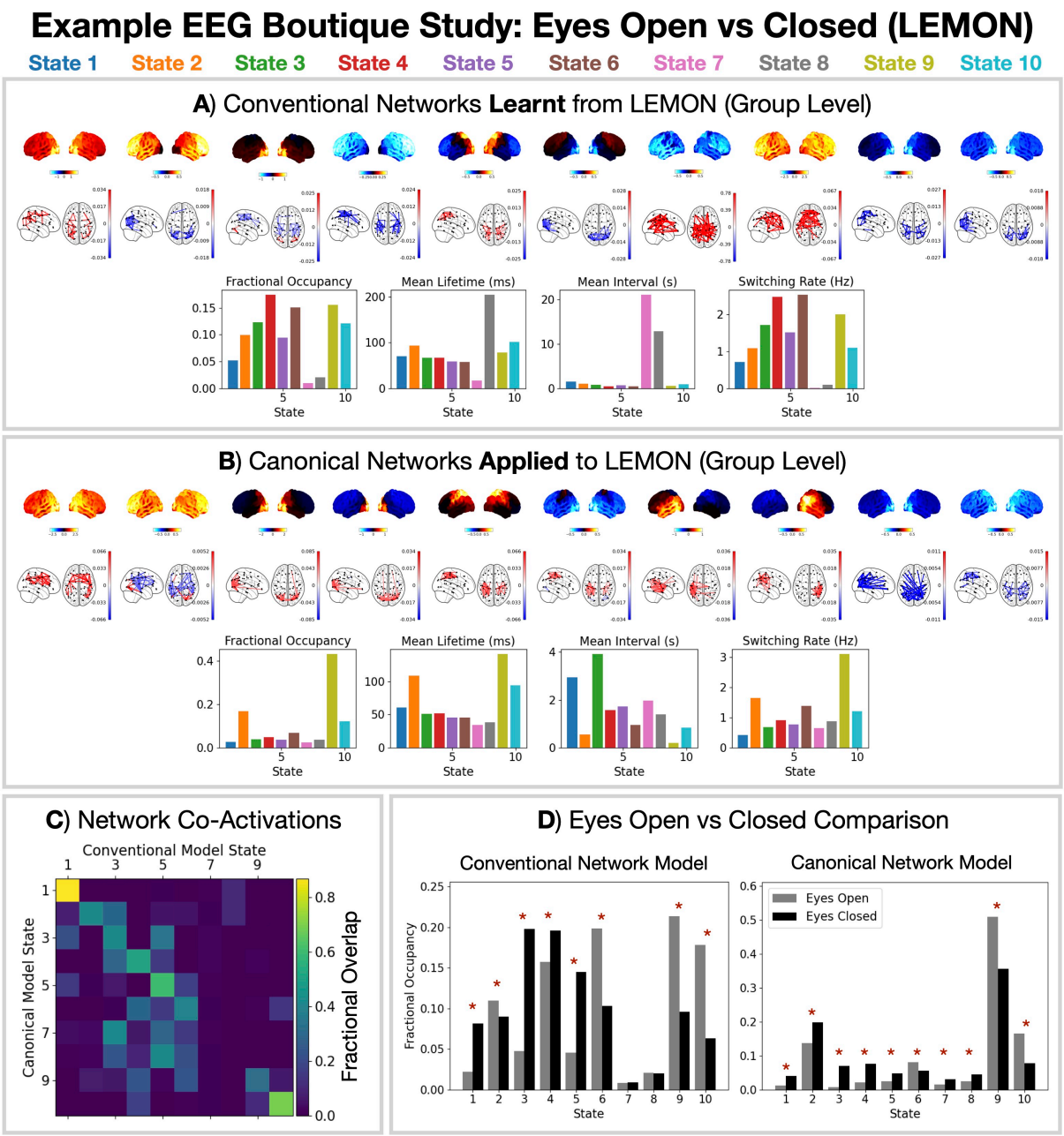

Figure 4: **MEG-like networks can be inferred from parcellated EEG data**. A) Relative power maps, coherence networks (top 3%), and summary statistics for dynamics for a conventional approach training the HMM directly on the LEMON dataset. B) Relative power maps, coherence networks (top 3%), and summary statistics for dynamics inferred in the LEMON dataset by applying the 10-state canonical HMM networks from Cam-CAN. C) Overlap of the fractional occupancy for each state from the conventional and canonical HMM network modelling approach. D) Comparison of the state fractional occupancy in eyes open and closed conditions for the conventional and canonical HMM network modelling approach. Red asterisks indicate a significant difference between eyes open and closed (p-value < 0.05).

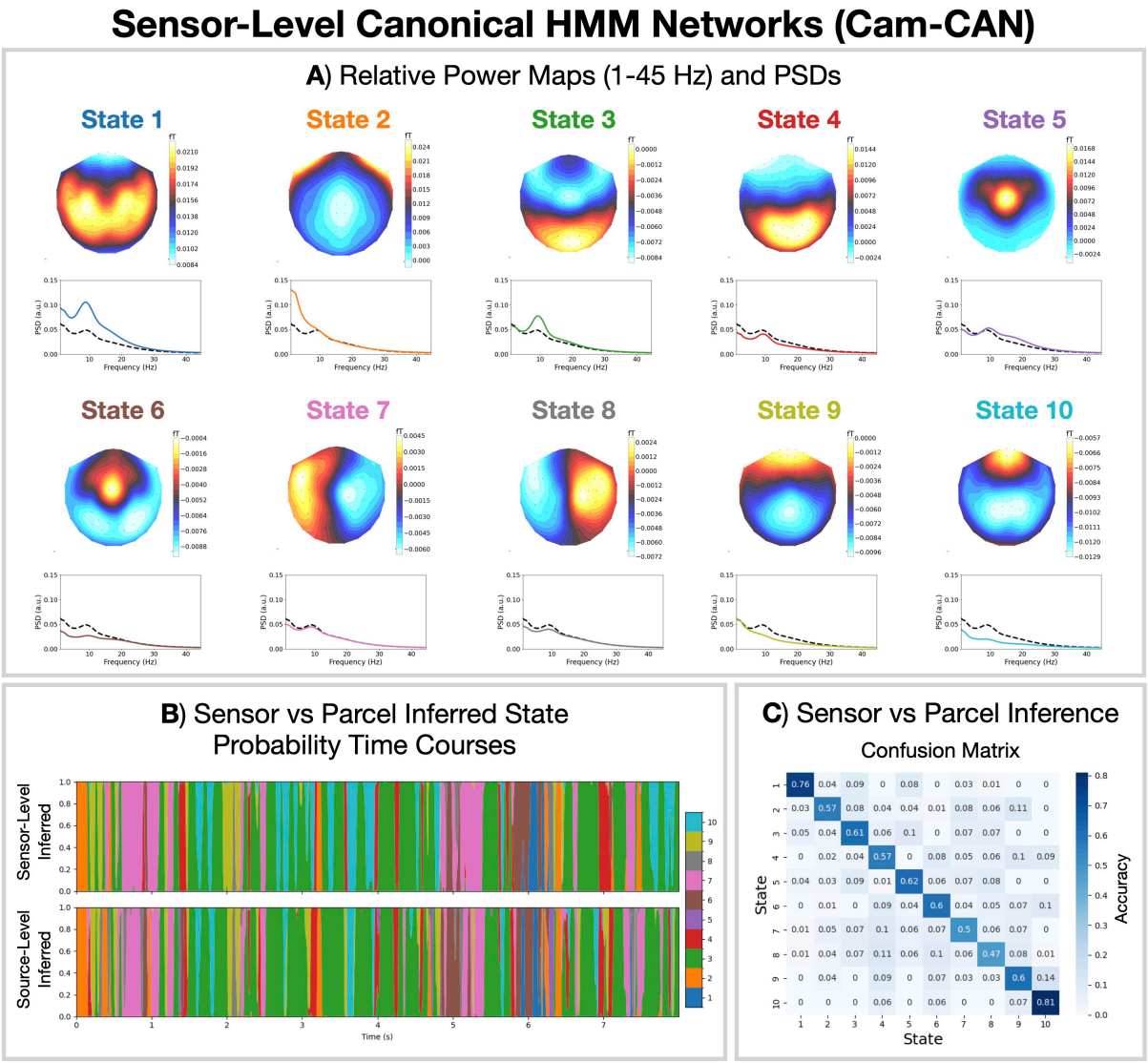

Figure 5: **The canonical HMM approach can be applied in MEG sensor space**. This allows networks to be approximately inferred when source reconstruction is not feasible. A) Topographic maps of the relative state power at each MEG sensors viewed from above and PSDs for the 10-state canonical HMM networks at the sensor level, in the same order of states as for the parcel-space canonical HMM shown in Figure 2. B) Illustrative state probabilities inferred from the first 8 seconds of the first session in Cam-CAN when applying the canonical HMM networks to the sensor-level data (top) or the (prepared) parcellated data (bottom). C) Confusion matrix: average accuracy (across participants) for the state assignment at each time point between the sensor-level and source-level canonical models.

4 Discussion

We propose a canonical HMM that can be used to study transient networks in M/EEG data. Each HMM network corresponds to a unique spatio-spectral pattern of brain activity with fast dynamics (Figure 2). The canonical HMM was trained on rest and task (visual, auditory, sensorimotor) data from a large cohort of healthy participants (N=621, 18-88 years old, Cam-CAN), capturing a diverse range of brain activity and population variability. We demonstrated how the parcel-level canonical HMM could be used to analyse a boutique MEG dataset for early Alzheimer's detection (Figure 3). We showed the canonical HMM networks could be transferred to parcellated EEG data (Figure 4). We also provided a sensor-level canonical HMM for Elekta MEG (Figure 5). Here, we discuss some considerations regarding the canonical HMM and highlight some advantages of adopting this approach.

4.1 Canonical HMM Networks as a Basis Set

The canonical HMM states presented in this work can serve as a basis set for studying M/EEG data using a set of canonical transient networks. These networks were inferred on a large MEG cohort, which is sufficient to ensure good reproducibility on unseen datasets (Figures 3 and 4). A benefit of taking this approach is it provides a description of the data in a common space, which can be useful in comparing individuals, especially in boutique or individualised (N=1) studies. The adoption of a common basis set also improves the comparability of analysis across different studies. Conventional M/EEG HMM studies typically involve training a model from scratch. This means the networks inferred in one study do not exactly match another. This can be resolved via the adoption of a common canonical HMM.

4.2 Reduction in Computational Resources Needed

A key benefit of adopting a canonical HMM approach is it means the computational resources needed to perform an HMM analysis have been greatly reduced. The end user no longer needs to train their own HMM, which can be a computationally demanding process that often needs to be repeated many times. The canonical HMMs trained in this work utilised modern GPU hardware and took on the order of 10 hours to complete a single run. In contrast, applying a trained HMM to new data (inference) requires relatively little computational resources and is quick (on the order of minutes). HMM inference can be done straightforwardly on most modern CPUs. This increases the accessibility of these tools to end users without access to significant computational resources.

4.3 Stable Network Inference

As with many machine learning approaches, the HMM can suffer from a local optima issue, where the model can converge to different final solutions (model parameters). Conventionally, this is resolved by training many HMMs from scratch and selecting the best for subsequent analysis (usually the one with the lowest variational free energy (Gohil et al., 2024a)). Adopting a canonical HMM alleviates this issue because given the trained model and data, the inference is deterministic.

4.4 Preprocessing and Source Reconstruction

To apply the parcel-level canonical HMM to new data, it must be preprocessed and source reconstructed in a similar manner to the data used to train the canonical HMM. The key data processing steps that need to be matched are downsampling to 250 Hz, LCMV beamforming, parcellation with the same atlas, orthogonalisation, sign flipping to match the template subject from Cam-CAN, number of time-delay embeddings, and PCA using the same components from

Cam-CAN. All the necessary data and example scripts for this process are provided in the code repository (see Section 5). Other steps such as the exact band-pass filtering range and data cleaning are less critical. Data collected with any scanner type/modality can be source reconstructed and be used in combination with the parcel-level canonical HMM. For scenarios where the end user does not wish to adopt the same source reconstruction, we provide a sensor-level canonical HMM for Elekta MEG.

4.5 Transferring MEG Networks to EEG

The objective of this work was to develop a canonical HMM based on MEG. However, this model can be leveraged in EEG studies. Due to differences in the underlying signal being measured (magnetic field for MEG and electric field for EEG), MEG and EEG capture different aspects of brain activity and have different spatio-spectral sensitivities (Hari and Puce, 2023). Like MEG, EEG data can be source reconstructed to estimate activity within the brain. However, EEG data is typically noisier and has a lower channel density compared to MEG. Combined with the sensitivity to changes in the electromagnetic properties of the head, this often makes EEG data more difficult to source reconstruct (Hari and Puce, 2023). Leveraging information from source reconstructed MEG could help improve the inferences in EEG data. The canonical HMM approach proposed here aims to do this. In Figure 4, we see that adopting the canonical HMM allows us to identify MEG-like networks in the parcellated EEG data. This approach could facilitate studies of large-scale cortical networks using medium/high-density EEG.

4.6 Number of States

An important choice when training an HMM is the number of states. Each specification for the number of states is a different model for the data. In this work, we use variational Bayesian inference to train the HMM, which involves optimising the *variational free energy* (Bishop and Nasrabadi, 2006). This objective function accounts for the number of states by adding a term to the training loss that penalises model complexity (i.e., a higher number of states). The variational free energy can be used to compare different models and select one for subsequent analysis (Alonso and Vidaurre, 2023). The model with the lowest value is one that best fits the data whilst accounting for the number of states chosen. Figure S8A shows the variational free energy decreases with the number of states fitted to the Cam-CAN dataset (N=621, rest and task). This indicates the more states we fit, the better we model the Cam-CAN data.

However, how well we fit a particular dataset is not the only factor we consider when selecting the number of states. We wish to identify states that exist across a large population rather than modelling noise that is only present in our particular dataset. That is, we would like to accurately describe data from participants not seen when training the model. We seek states that can be *reproduced*[4] in new data. A common approach for assessing this in neuroimaging studies is *split-half validation*[5], where a model is trained on each half of the data independently and the similarity of the trained models is assessed. States that appear in both halves are reproducible. Figure S8 shows the split-half reproducibility of the states inferred on the Cam-CAN dataset.

Each HMM state corresponds to a spatio-spectral pattern of brain activity that arises when the state is active (power map). To assess the similarity between a model from each half of Cam-CAN, we calculate the average Pearson correlation between the power maps of each state (after re-ordering to find the best pairing). We observe a drop off in the average correlation between matched states from each half at around 20 states (Figure S8B). That is, we can identify 20 states that are common in two independent fits to separate halves of the Cam-CAN dataset (Figure S8C). Figure S9 shows the split-half reproducibility using a smaller subset of Cam-

[4]Here, we refer to the same concept as *generalisation*, which is the term used in the machine learning literature.
[5]Equivalent to 2-fold cross validation.

CAN (N=50 and N=16, rest and task). We see with less data, we cannot identify as many reproducible states.

By convention, it is common to select between 8 and 12 states (Vidaurre et al., 2018b; Cho et al., 2024; Gohil et al., 2024a; Higgins et al., 2021; Gohil et al., 2024b; van Es et al., 2023; Gohil et al., 2022; Quinn et al., 2018; Vidaurre et al., 2018a, 2016; Baker et al., 2014). We make parcel and sensor-level canonical HMMs with 4 to 16 states publicly available (see Section 5).

4.7 Task-Relevance of the Canonical HMM Networks

We can use the task data from Cam-CAN to see how the different canonical HMM networks are modulated by task conditions (Van Den Heuvel and Pol, 2010; Smitha et al., 2017; Lee et al., 2013). The tasks included in Cam-CAN were a passive visual and auditory stimulus and a sensorimotor response (button press following a visuo-auditory stimulus). See (Shafto et al., 2014; Taylor et al., 2017) for further details regarding the tasks.

We can study the network-level response (Gohil et al., 2024b) to the various stimuli by epoching the state time courses around stimulus events. This is shown in Figure 6. We see multiple networks activate/deactivate in response to the stimuli. Four networks (states 1, 2, 3 and 10) show the same sequential response to the stimuli (Figure 6B). These networks correspond to higher-order cognitive function - this is discussed further in Section 4.8. We see networks that correspond to activity in unimodal primary cortices (Rolls, 2016): states 4 and 6 exhibit activity in the primary visual cortex; states 5 and 7 exhibit activity in the primary auditory cortex; and states 8 and 9 exhibit activity in the primary somatosensory/motor cortex. We can therefore associate these networks to the processing of primary sensory inputs.

The canonical HMM provide a flexible basis set that can describe a variety of cognitive functions in both task and rest data. Adopting the canonical HMM for task analysis can be useful in the study of disease. Particularly if the symptoms of disease only arise in task (Casagrande et al., 2023).

4.8 Correspondence Between MEG and fMRI Networks

Here, we compare the spatial power maps of the canonical HMM states to the fMRI networks presented in (Yeo et al., 2011), which are shown in Figure 7A.

We divide the fMRI networks into two types: those that correspond to *lower-order* cognitive function (primary sensory input, networks 1 and 2) and those that correspond to *higher-order* cognition function, i.e. functions that go beyond primary sensory input (networks 3-7) (**?**Smitha et al., 2017). Higher-order cognitive functions include processes such as executive control, memory and attention (Smitha et al., 2017; Kiely, 2014). Correlating the HMM MEG state power maps with the fMRI networks, we see the MEG networks can also be divided into those that correspond to higher-order and lower-order cognitive function (Figure 7).

Of particular interest is the *default mode* network from fMRI (network 7 in Figure 7A), which is associated with internally oriented cognitive functions (Raichle, 2015) and tends to activate in the resting state (Van Den Heuvel and Pol, 2010). In MEG, it was found that two HMM states correspond to a default mode network: a posterior and an anterior network (canonical HMM states 1 and 2 respectively) (Vidaurre et al., 2018b). This division is due to the higher temporal resolution provided by MEG (milliseconds) compared to fMRI (seconds). Combining these two HMM states we find a spatial pattern of power that correlates with the fMRI default mode network (Figure 7B).

A well-known finding in the fMRI literature is the anti-correlation between the dynamics of the default mode network (7 in Figure 7A) and the dorsal attention network (3 in Figure 7A) (Fox et al., 2005). The corresponding networks in MEG that have anti-correlated dynamics to the default mode network are states 3 and 10. This was previously shown in (Baker et al., 2014) and is reflected by the lack of transitions between these states (Figure 2D). Correlating the

combination of states 3 and 10 to the fMRI networks we see strong associations with all of the higher-order fMRI networks. Therefore, based on the spatial pattern of activity alone these states cannot be associated to a single fMRI network. However, based on the anti-correlation in terms of dynamics and the suppressed posterior-α (8-12 Hz) activity exhibited by the states, which is common marker for spatial attention (Gould et al., 2011), we consider states 3 and 10 to be attention networks in MEG.

Turning to the lower-order fMRI networks, we can see MEG states 4 and 6 correspond to the visual network (Figure 7C) and MEG states 8 and 9 correspond to the sensorimotor network (Figure 7C). In MEG, we are also able to identify a network (states 5 and 7) that correspond to auditory inputs, which is not present in the fMRI networks from (Yeo et al., 2011).

4.9 Applications in Normative Modelling

Normative modelling (Rutherford et al., 2023, 2022; Marquand et al., 2016) characterises a reference curve as a function of a covariate, commonly age. The variability around this reference curve (distribution) is also specified in the normative model. This allows us to quantity how a particular individual deviates from the reference curve as a percentile.

The canonical HMM provides a basis set for networks that can be post-hoc dual estimated for individuals (Figure S5). This means we are able to calculate individualised network characteristics, such as the region-specific power, coherence and summary statistics for dynamics for each state, based on the canonical HMM. The individualised network characteristics (in rest and task) for 621 healthy participants in the Cam-CAN study have been made publicly available (see Section 5). Such characteristics can be used to build a normative model, i.e. a model to estimate a quantify of interest as a function of age. The Cam-CAN dataset contains a wide age range (18-88 years old) with an approximately uniform distribution, which makes it ideal for use as a normative model. Age has several widespread effects on the HMM networks that are captured in the canonical model, these are studied in detail in (Gohil et al., 2024c; Tibon et al., 2021).

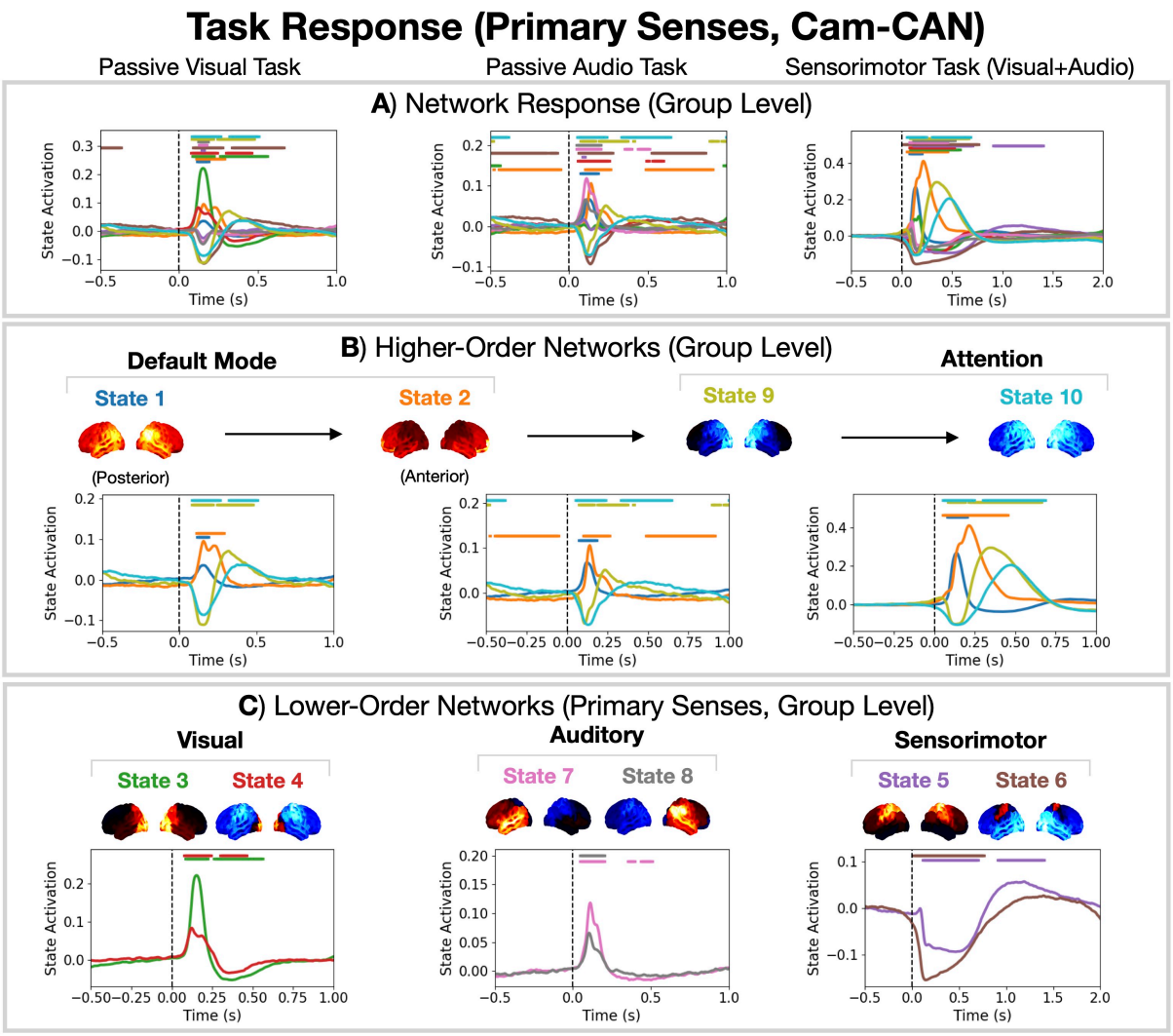

Figure 6: **The canonical HMM networks provide a useful description for a variety of tasks**. A) Network response (using the 10-state canonical HMM networks) to a passive visual task (left), passive audio task (middle) and visual/auditory sensorimotor task (right). See (Shafto et al., 2014; Taylor et al., 2017) for details regarding the tasks. B) Network response for the higher-order networks (states 1, 2, 9, 10). C) Network response for the lower-order networks corresponding to primary sense responses (states 3 and 4 for visual, states 7 and 8 for auditory, and states 5 and 6 for sensorimotor). The horizontal bars in each state activation plot indicate a p-values < 0.05.

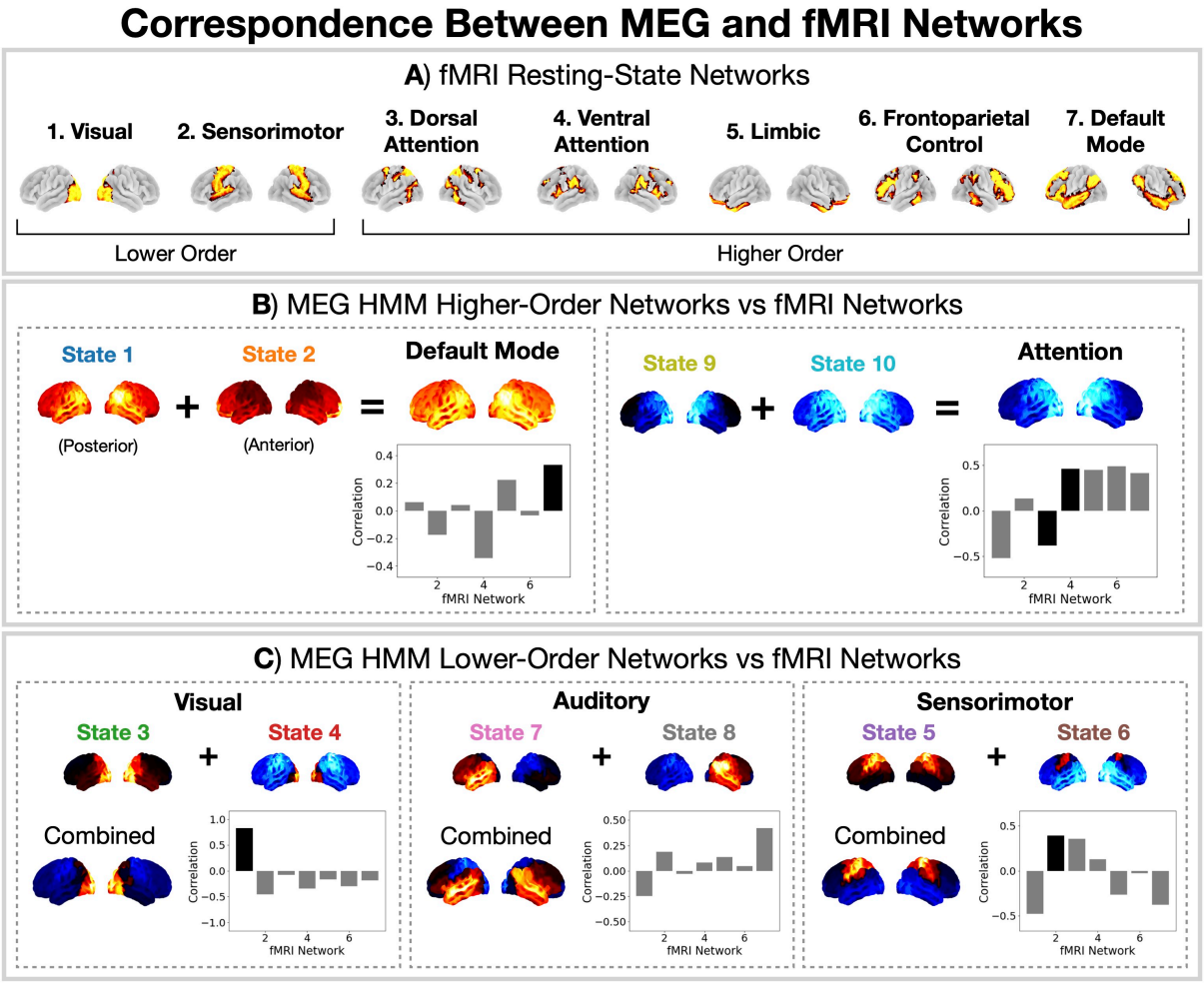

Figure 7: **Correspondence between the canonical MEG networks and fMRI resting-state networks**. A) Seven fMRI resting-state networks identified by (Yeo et al., 2011). FMRI networks 1 and 2 correspond to primary sensory input (*lower-order* cognitive function) and networks 3-7 correspond to *higher-order* cognitive function. B) Correlation between the higher-order MEG networks (default mode, left; attention, right) and the fMRI networks. C) Correlation between the lower-order networks (primary sensory input) and fMRI networks. In black we highlight the pairing of the MEG and fMRI networks.

5 Data and Code Availability

All the datasets used in this work are publicly available (see Section 2.1). Code and examples scripts for processing M/EEG data and applying the canonical HMMs have been made publicly available here: https://github.com/OHBA-analysis/Canonical-HMM-Networks.

As larger and more diverse datasets become available this resource will be actively updated. This includes providing canonical HMMs with differing model orders, different parcellations, different sensor layouts and training on larger datasets including other modalities, such as EEG and Optically Pumped Magnetometer data.

6 Funding

This research was supported by the National Institute for Health Research (NIHR) Oxford Health Biomedical Research Centre. The Wellcome Centre for Integrative Neuroimaging is supported by core funding from the Wellcome Trust (203139/Z/16/Z). CG is supported by the Wellcome Trust (215573/Z/19/Z). RH is supported by the EPSRC Centre for Doctoral Training in Health Data Science (EP/S02428X/1). MWJvE is supported by the Wellcome Trust (106183/Z/14/Z, 215573/Z/19/Z), the New Therapeutics in Alzheimer's Diseases (NTAD) supported by the MRC and the Dementia Platform UK (RG94383/RG89702). DV is supported by a Novo Nordisk Foundation Emerging Investigator Fellowship (NNF19OC-0054895), an ERC Starting Grant (ERC-StG-2019-850404), and a DFF Project 1 from the Independent Research Fund of Denmark (2034-00054B). MWW is supported by the Wellcome Trust (106183/Z/14/Z, 215573/Z/19/Z), the New Therapeutics in Alzheimer's Diseases (NTAD) study supported by UK MRC, the Dementia Platform UK (RG94383/RG89702) and the NIHR Oxford Health Biomedical Research Centre (NIHR203316). The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

7 Competing Interests

No competing interests.

References

Alonso, S. and Vidaurre, D. (2023). Toward stability of dynamic fc estimates in neuroimaging and electrophysiology: Solutions and limits. *Network Neuroscience*, 7(4):1389–1403.

Babadi, B. and Brown, E. N. (2014). A review of multitaper spectral analysis. *IEEE Transactions on Biomedical Engineering*, 61(5):1555–1564.

Babayan, A., Erbey, M., Kumral, D., Reinelt, J. D., Reiter, A. M., Röbbig, J., Schaare, H. L., Uhlig, M., Anwander, A., Bazin, P.-L., et al. (2019). A mind-brain-body dataset of mri, eeg, cognition, emotion, and peripheral physiology in young and old adults. *Scientific data*, 6(1):1–21.

Baker, A. P., Brookes, M. J., Rezek, I. A., Smith, S. M., Behrens, T., Probert Smith, P. J., and Woolrich, M. (2014). Fast transient networks in spontaneous human brain activity. *elife*, 3:e01867.

Bishop, C. M. and Nasrabadi, N. M. (2006). *Pattern recognition and machine learning*, volume 4. Springer.

Blinowska, K. and Durka, P. (2006). Electroencephalography (eeg). *Wiley encyclopedia of biomedical engineering*, 10:9780471740360.

Bommasani, R. (2021). On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.

Casagrande, C. C., Rempe, M. P., Springer, S. D., and Wilson, T. W. (2023). Comprehensive review of task-based neuroimaging studies of cognitive deficits in alzheimer's disease using electrophysiological methods. *Ageing research reviews*, 88:101950.

Cho, S., van Es, M., Woolrich, M., and Gohil, C. (2024). Comparison between eeg and meg of static and dynamic resting-state networks. *Human Brain Mapping*, 45(13):e70018.

Colclough, G. L., Brookes, M. J., Smith, S. M., and Woolrich, M. W. (2015). A symmetric multivariate leakage correction for meg connectomes. *Neuroimage*, 117:439–448.

Cooray, N., Gohil, C., Harris, B., Frost, S., and Higgins, C. (2024). Default mode network detection using eeg in real-time. In *2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1–4. IEEE.

Fauchon, C., Kim, J. A., El-Sayed, R., Osborne, N. R., Rogachov, A., Cheng, J. C., Hemington, K. S., Bosma, R. L., Dunkley, B. T., Oh, J., et al. (2022). A hidden markov model reveals magnetoencephalography spectral frequency-specific abnormalities of brain state power and phase-coupling in neuropathic pain. *Communications biology*, 5(1):1000.

Fox, M. D., Snyder, A. Z., Vincent, J. L., Corbetta, M., Van Essen, D. C., and Raichle, M. E. (2005). The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proceedings of the National Academy of Sciences*, 102(27):9673–9678.

Gohil, C., Huang, R., Roberts, E., van Es, M. W., Quinn, A. J., Vidaurre, D., and Woolrich, M. W. (2024a). osl-dynamics, a toolbox for modeling fast dynamic brain activity. *Elife*, 12:RP91949.

Gohil, C., Kohl, O., Huang, R., van Es, M. W., Parker Jones, O., Hunt, L. T., Quinn, A. J., and Woolrich, M. W. (2024b). Dynamic network analysis of electrophysiological task data. *Imaging Neuroscience*, 2:1–19.

Gohil, C., Kohl, O., Pitt, J., van Es, M. W., Quinn, A. J., Vidaurre, D., Turner, M. R., Nobre, A. C., and Woolrich, M. W. (2024c). Effects of age on resting-state cortical networks. *bioRxiv*, pages 2024–09.

Gohil, C., Roberts, E., Timms, R., Skates, A., Higgins, C., Quinn, A., Pervaiz, U., van Amersfoort, J., Notin, P., Gal, Y., et al. (2022). Mixtures of large-scale dynamic functional brain network modes. *NeuroImage*, 263:119595.

Gould, I. C., Rushworth, M. F., and Nobre, A. C. (2011). Indexing the graded allocation of visuospatial attention using anticipatory alpha oscillations. *Journal of neurophysiology*, 105(3):1318–1326.

Guichet, C., Harquel, S., Achard, S., Mermillod, M., and Baciu, M. (2025). Lifespan oscillatory dynamics in lexical production: A population-based meg resting-state analysis. *Imaging Neuroscience*, 3:imag_a_00551.

Hari, R. and Puce, A. (2023). *Meg-EEG primer*. Oxford University Press.

Henson, R. N., Mattout, J., Phillips, C., and Friston, K. J. (2009). Selecting forward models for meg source-reconstruction using model-evidence. *Neuroimage*, 46(1):168–176.

Higgins, C., Liu, Y., Vidaurre, D., Kurth-Nelson, Z., Dolan, R., Behrens, T., and Woolrich, M. (2021). Replay bursts in humans coincide with activation of the default mode and parietal alpha networks. *Neuron*, 109(5):882–893.

Hillebrand, A. and Barnes, G. R. (2005). Beamformer analysis of meg data. *International review of neurobiology*, 68:149–171.

Hughes, L. E., Henson, R. N., Pereda, E., Bruña, R., López-Sanz, D., Quinn, A. J., Woolrich, M. W., Nobre, A. C., Rowe, J. B., Maestú, F., et al. (2019). Biomagnetic biomarkers for dementia: A pilot multicentre study with a recommended methodological framework for magnetoencephalography. *Alzheimer's & Dementia: Diagnosis, Assessment & Disease Monitoring*, 11:450–462.

Hyvarinen, A. (1999). Fast and robust fixed-point algorithms for independent component analysis. *IEEE transactions on Neural Networks*, 10(3):626–634.

Jenkinson, M. (2005). Bet2: Mr-based estimation of brain, skull and scalp surfaces. In *Eleventh Annual Meeting of the Organization for Human Brain Mapping, 2005*, volume 17, page 167.

Kiely, K. M. (2014). Cognitive function. In *Encyclopedia of quality of life and well-being research*, pages 974–978. Springer.

Kohl, O., Gohil, C., Zokaeia, N., Hu, M. T., Nobrea, A. C., Woolrich, M., and Quinn, A. (2024). Changes in sensorimotor network dynamics in resting-state recordings in parkinson's disease. *medRxiv*, pages 2024–01.

Kohl, O., Woolrich, M., Nobre, A., and Quinn, A. (2023). Glasser52: a parcellation for meg-analysis.

Lee, M. H., Smyser, C. D., and Shimony, J. S. (2013). Resting-state fmri: a review of methods and clinical applications. *American Journal of neuroradiology*, 34(10):1866–1872.

Makkinayeri, S., Guidotti, R., Basti, A., Woolrich, M. W., Gohil, C., Pettorruso, M., Ermolova, M., Ilmoniemi, R. J., Ziemann, U., Romani, G. L., et al. (2025). Investigating brain network dynamics in state-dependent stimulation: A concurrent electroencephalography and transcranial magnetic stimulation study using hidden markov models. *Brain Stimulation*, 18(3):800–809.

Marquand, A. F., Rezek, I., Buitelaar, J., and Beckmann, C. F. (2016). Understanding heterogeneity in clinical cohorts using normative models: beyond case-control studies. *Biological psychiatry*, 80(7):552–561.

Perrin, F., Pernier, J., Bertrand, O., and Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and clinical neurophysiology*, 72(2):184–187.

Quinn, A., van Es, M., Gohil, C., and Woolrich, M. (2022). Ohba software library in python (osl). *Zenodo https://doi. org/10.5281/ZENODO*, 6875060.

Quinn, A. J., Vidaurre, D., Abeysuriya, R., Becker, R., Nobre, A. C., and Woolrich, M. W. (2018). Task-evoked dynamic network analysis through hidden markov modeling. *Frontiers in neuroscience*, 12:603.

Rabiner, L. and Juang, B. (2003). An introduction to hidden markov models. *ieee assp magazine*, 3(1):4–16.

Raichle, M. E. (2015). The brain's default mode network. *Annual review of neuroscience*, 38(1):433–447.

Rolls, E. T. (2016). *Cerebral cortex: principles of operation*. Oxford University Press.

Rosner, B. (1983). Percentage points for a generalized esd many-outlier procedure. *Technometrics*, 25(2):165–172.

Rossi, C., Vidaurre, D., Costers, L., Akbarian, F., Woolrich, M., Nagels, G., and Van Schependom, J. (2023). A data-driven network decomposition of the temporal, spatial, and spectral dynamics underpinning visual-verbal working memory processes. *Communications Biology*, 6(1):1079.

Rutherford, S., Barkema, P., Tso, I. F., Sripada, C., Beckmann, C. F., Ruhe, H. G., and Marquand, A. F. (2023). Evidence for embracing normative modeling. *Elife*, 12:e85082.

Rutherford, S., Kia, S. M., Wolfers, T., Fraza, C., Zabihi, M., Dinga, R., Berthet, P., Worker, A., Verdi, S., Ruhe, H. G., et al. (2022). The normative modeling framework for computational psychiatry. *Nature protocols*, 17(7):1711–1734.

Scheltens, P., De Strooper, B., Kivipelto, M., Holstege, H., Chételat, G., Teunissen, C. E., Cummings, J., and van der Flier, W. M. (2021). Alzheimer's disease. *The Lancet*, 397(10284):1577–1590.

Seedat, Z. A., Rier, L., Gascoyne, L. E., Cook, H., Woolrich, M. W., Quinn, A. J., Roberts, T. P., Furlong, P. L., Armstrong, C., St. Pier, K., et al. (2023). Mapping interictal activity in epilepsy using a hidden markov model: A magnetoencephalography study. *Human Brain Mapping*, 44(1):66–81.

Shafto, M. A., Tyler, L. K., Dixon, M., Taylor, J. R., Rowe, J. B., Cusack, R., Calder, A. J., Marslen-Wilson, W. D., Duncan, J., Dalgleish, T., et al. (2014). The cambridge centre for ageing and neuroscience (cam-can) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. *BMC neurology*, 14(1):204.

Sitnikova, T. A., Hughes, J. W., Ahlfors, S. P., Woolrich, M. W., and Salat, D. H. (2018). Short timescale abnormalities in the states of spontaneous synchrony in the functional neural networks in alzheimer's disease. *NeuroImage: Clinical*, 20:128–152.

Smith, S. M. (2002). Fast robust automated brain extraction. *Human brain mapping*, 17(3):143–155.

Smith, S. M., Fox, P. T., Miller, K. L., Glahn, D. C., Fox, P. M., Mackay, C. E., Filippini, N., Watkins, K. E., Toro, R., Laird, A. R., et al. (2009). Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the national academy of sciences*, 106(31):13040–13045.

Smitha, K. A., Akhil Raja, K., Arun, K., Rajesh, P., Thomas, B., Kapilamoorthy, T., and Kesavadas, C. (2017). Resting state fmri: A review on methods in resting state connectivity analysis and resting state networks. *The neuroradiology journal*, 30(4):305–317.

Tait, L., Özkan, A., Szul, M. J., and Zhang, J. (2021). A systematic evaluation of source reconstruction of resting meg of the human brain with a new high-resolution atlas: Performance, precision, and parcellation. *Human Brain Mapping*, 42(14):4685–4707.

Taylor, J. R., Williams, N., Cusack, R., Auer, T., Shafto, M. A., Dixon, M., Tyler, L. K., Henson, R. N., et al. (2017). The cambridge centre for ageing and neuroscience (cam-can) data repository: Structural and functional mri, meg, and cognitive data from a cross-sectional adult lifespan sample. *neuroimage*, 144:262–269.

Tibon, R., Tsvetanov, K. A., Price, D., Nesbitt, D., Can, C., and Henson, R. (2021). Transient neural network dynamics in cognitive ageing. *Neurobiology of aging*, 105:217–228.

Vaghari, D., Bruna, R., Hughes, L. E., Nesbitt, D., Tibon, R., Rowe, J. B., Maestu, F., and Henson, R. N. (2022). A multi-site, multi-participant magnetoencephalography resting-state dataset to study dementia: The biofind dataset. *Neuroimage*, 258:119344.

Van Den Heuvel, M. P. and Pol, H. E. H. (2010). Exploring the brain network: a review on resting-state fmri functional connectivity. *European neuropsychopharmacology*, 20(8):519–534.

van Es, M. W., Gohil, C., Quinn, A. J., and Woolrich, M. W. (2025). osl-ephys: A python toolbox for the analysis of electrophysiology data. *Frontiers in Neuroscience*, 19:1522675.

van Es, M. W., Higgins, C., Gohil, C., Quinn, A. J., Vidaurre, D., and Woolrich, M. W. (2023). Large-scale cortical networks are organized in structured cycles. *bioRxiv*, pages 2023–07.

Vidaurre, D., Abeysuriya, R., Becker, R., Quinn, A. J., Alfaro-Almagro, F., Smith, S. M., and Woolrich, M. W. (2018a). Discovering dynamic brain networks from big data in rest and task. *NeuroImage*, 180:646–656.

Vidaurre, D., Hunt, L. T., Quinn, A. J., Hunt, B. A., Brookes, M. J., Nobre, A. C., and Woolrich, M. W. (2018b). Spontaneous cortical activity transiently organises into frequency specific phase-coupling networks. *Nature communications*, 9(1):2987.

Vidaurre, D., Quinn, A. J., Baker, A. P., Dupret, D., Tejero-Cantero, A., and Woolrich, M. W. (2016). Spectrally resolved fast transient brain states in electrophysiological data. *Neuroimage*, 126:81–95.

Vidaurre, D., Smith, S. M., and Woolrich, M. W. (2017). Brain network dynamics are hierarchically organized in time. *Proceedings of the National Academy of Sciences*, 114(48):12827–12832.

Wang, E. Y., Fahey, P. G., Ding, Z., Papadopoulos, S., Ponder, K., Weis, M. A., Chang, A., Muhammad, T., Patel, S., Ding, Z., et al. (2025). Foundation model of neural activity predicts response to new stimulus types. *Nature*, 640(8058):470–477.

Yeo, B. T., Krienen, F. M., Sepulcre, J., Sabuncu, M. R., Lashkari, D., Hollinshead, M., Roffman, J. L., Smoller, J. W., Zöllei, L., Polimeni, J. R., et al. (2011). The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology*.