NeuroImage 260 (2022) 119489

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

Modes of cognition: Evidence from metastable brain dynamics

Katerina Capouskova [a,*], Morten L. Kringelbach [b,c], Gustavo Deco [a,d,e,f]

[a] Center for Brain and Cognition, Computational Neuroscience Group, Department of Information and Communication Technologies, Universitat Pompeu Fabra, Ramon Trias Fargas 25-27, Barcelona 08005, Spain
[b] Department of Psychiatry, University of Oxford, Oxford, United Kingdom
[c] Center for Music in the Brain, Department of Clinical Medicine, Aarhus University, Aarhus, Denmark
[d] Department of Neuropsychology, Max Planck Institute for Human Cognitive and Brain Sciences, Leipzig, Germany
[e] Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain
[f] Turner Institute for Brain and Mental Health, Monash University, Melbourne, VIC, Australia

ARTICLE INFO

Keywords:
Functional connectivity
Brain states
fMRI
HCP data set
Classification
Entropy
DADA

ABSTRACT

Managing cognitive load depends on adequate resource allocation by the human brain through the engagement of metastable substates, which are large-scale functional networks that change over time. We employed a novel analysis method, deep autoencoder dynamical analysis (DADA), with 100 healthy adults selected from the Human Connectome Project (HCP) data set in rest and six cognitive tasks. The deep autoencoder of DADA described seven recurrent stochastic metastable substates from the functional connectome of BOLD phase coherence matrices. These substates were significantly differentiated in terms of their probability of appearance, time duration, and spatial attributes. We found that during different cognitive tasks, there was a higher probability of having more connected substates dominated by a high degree of connectivity in the thalamus. In addition, compared with those during tasks, resting brain dynamics have a lower level of predictability, indicating a more uniform distribution of metastability between substates, quantified by higher entropy. These novel findings provide empirical evidence for the philosophically motivated cognitive theory, suggesting on-line and off-line as two fundamentally distinct modes of cognition. On-line cognition refers to task-dependent engagement with the sensory input, while off-line cognition is a slower, environmentally detached mode engaged with decision and planning. Overall, the DADA framework provides a bridge between neuroscience and cognitive theory that can be further explored in the future.

1. Introduction

In cognitive processing, the brain forms temporal functional unions of distinct brain areas to process information. This switching between functional networks occurs in not only cognitively demanding conditions but also less challenging basal states, such as meditation or rest (Cabral et al., 2011). The rapid changing of functional networks provides cognitive flexibility to couple with the environment and fulfill a demand or decouple from the environment during rest to integrate information and prepare for the subsequent cognitive task requirements (Sadaghiani et al., 2010; Sadaghiani and Kleinschmidt, 2013). The activation and deactivation of different functional networks are supported by the structural connections of the brain's connectome, and therefore the structural constraints play an important role in the pattern of switching (Sporns, 2013). There are two main aspects of the functional connectivity (FC) patterns: the temporal, which determines the length of each pattern, and the spatial, which determines the possible topological structure of the functional connectivity.

Within the fields of cognitive science, psychology, and philosophy, it has been debated that there exist two fundamentally distinct modes of cognition: on-line and off-line (Costa and Rocha, 2005). On-line cognition involves momentary input from our immediate environment and is concerned with the task at present, which requires rapid embodied processing (Iverson and Thelen, 1999). When decoupled from the environment, cognition switches to a slower, off-line mode that enables more deliberate decisions or the planning of future behaviors (Wilson, 2002). In everyday activities, one is more engaged in an on-line cognition mode; however, when deeper understanding is needed or a hypothetical situation is being constructed, one switches into off-line mode.

In the DADA framework, within the use of temporal and spatial elements of functional connectivity patterns, we define brain states in a probabilistic metastable substate space (Alderson et al., 2020). These substates correspond to stochastic subclasses of standard and continual brain states (Vila-Vidal et al., 2019). Identified substates appear repeatedly and can be characterized by their probability of incidence, time spans, and spatial characteristics (Deco et al., 2019). Here, we hypoth-

* Corresponding author.
E-mail address: katerina.capouskova@upf.edu (K. Capouskova).

https://doi.org/10.1016/j.neuroimage.2022.119489.
Received 20 May 2021; Received in revised form 12 July 2022; Accepted 15 July 2022
Available online 23 July 2022.

esize that the stochastic substates represent the continual changes between the two cognitive modes, namely on-line and off-line. The on-line and off-line modes switch rapidly and often; however, during task activity, one is more engaged with the environment, and the substates representing on-line mode are present more often compared with rest. During rest, the substates corresponding to the off-line mode of internal planning have a higher occurrence rate. This would correspond to the hypothesis that the substates with higher functional connectivity, lower entropy, low modularity, and a high clustering coefficient should be recognized in tasks with a higher frequency compared with rest, as they represent the focus, precision, and execution of a cognitive task.

In different cognitive tasks, higher variability in FC predicts lower attention (Fong et al., 2019). Entropy, as a global measure of switching between metastable substates, should be higher in less attention-demanding modalities such as rest compared with cognitive tasks that require a higher cognitive load (Hellyer et al., 2015, 2014). Previous studies have demonstrated that canonical resting-state networks, consistent spatiotemporal patterns across different participants and data sets identified in rest, are relatively stable even under task conditions (Harmelech and Malach, 2013; Kieliba et al., 2019). However, a more precise DADA division of FC patterns along the spatial and temporal axes should help to redefine finer metastable substates with different probabilities and lifetimes in various cognitive modalities. In rest, FC should visit different substates more often, as it requires higher cognitive flexibility because it serves as a preparatory stage for higher specialization in more attention-demanding tasks.

The standard FC analysis sliding window partitions the temporal axis by the length parameter of the window, and inside each window interval, connectivity is computed as a Pearson correlation coefficient. The resulting set of connectivity matrices describes the time course of whole-brain functional connectivity (Hutchison et al., 2013). However, the primary shortcoming of the sliding window is the undefined length parameter, which can vary and affect the time resolution and statistical validation. Some methods have been proposed in different recording modalities to overcome the sliding-window drawback, such as phase coherence connectivity or time-frequency analysis (Chang and Glover, 2010; Vila-Vidal et al., 2020), which adapts the window length to the frequency content; however, it adds a new dimension to the parameter space. The leading eigenvector method (LEIDA*)* for dimension reduction as a proposed plan to overcome multidimensional issues (Cabral et al., 2017) is, in DADA, replaced by a more sophisticated non-linear dimension reduction autoencoder algorithm (Hinton and Salakhutdinov, 2006). Autoencoder, which is a certain type of artificial neural network, in DADA operates on the functional connectome of BOLD phase coherence matrices.

Dimensionality reduction can be effective if there is an underlying low dimensional structure. If the structure occupies non-linearity or curvature in the feature space, then autoencoder is a better choice as it can encode more information with less dimensions compared to traditional linear methods like PCA or LEIDA. It has been hypothesized that brain dynamics occupy low-dimensional smooth manifold despite the large number of its neuronal components and the multi-dimensionality of neuroimaging data (Gao and Ganguli, 2015; Jazayeri and Afraz, 2017; Jirsa, 2020; Pillai and Jirsa, 2017; Shine et al., 2019). Numerous studies have shown that not only in tasks (Chaudhuri and Fiete, 2016; Gallego et al., 2017; Williams et al., 2018) but also in rest or sleep, which are known to have rich neuronal activity repertoires, the structure of the dynamics is organized and can be characterized as a low-dimensional intrinsic manifold (Chaudhuri et al., 2019; Rué-Queralt et al., 2021). Given this evidence, within the context of the curved manifold, an autoencoder is the best option for encoding the data structure while preserving most of the features' relationship with each other. In this study, the autoencoder's more precise encoding provides a better separation of different substates and their evaluation.

The proposed DADA methodology (Fig. 1) includes BOLD signal transformation into phase via the Hilbert transform to pre-process the fMRI data for computing the phase coherence between all areas pairwise at each time point (Goelman et al., 2017), yielding a dynamic FC with size NxNxT, where N represents the number of brain areas, and T is the number of recorded time points. The reduction of dimensions is performed by an autoencoder. The functional connectivity dynamics of each time step are estimated by cosine similarity (Hansen et al., 2015). After performing clustering (e.g. k-means, among others), the mean lifetime of a substate, the probability of each substate, the switching rate, and the graph metrics were computed. For a global measure of switching between metastable substates, we employed entropy and Kullback-Leibler divergence and evaluated the performance of the DADA algorithm in terms of clustering.

2. Methods

2.1. Ethics

Full informed consent was acquired from all participants prior to the study according to the Washington University–University of Minnesota (WU-Minn HCP) Consortium research procedures and ethical guidelines. Moreover, the study was authorized by the Washington University review board.

2.2. Participants

The data set used for this study was selected from the March 2017 public data release from the Human Connectome Project (HCP), from which we selected a representative of 1003 participants. From this large portion of participants, we made another choice to reproduce in the smaller down sample of 100 unrelated participants (54 females, 46 males, mean age = 29.1+/−3.7 years). This subset of subjects administered by the HCP ensures they are not members of the same family, as this benchmark was crucial for excluding possible distinguishable confounds and the obligation for family-structure co-variables in the study.

2.3. Neuroimaging acquisition for fMRI HCP

The 1003 HCP participants were scanned using a 3-T connectome-Skyra scanner (Siemens).

One session performed on the same day included resting state fMRI acquisition lasting approximately 15 min, with eyes open and adopting a relaxed fixation on a projected bright cross on a dark background. The other sessions comprised seven different cognitive tasks. The full data set containing all the details about the participants, the study protocol, and the preprocessing of the data for all tasks and rest sessions can be obtained at the HCP website (http://www.humanconnectome.org/).

2.4. HCP tasks

The HCP data set contains seven tasks: working memory, motor, gambling, language, social, emotional, and relational, which are explained in greater detail on the HCP website (Barch et al., 2013). The HCP protocol consisted of two separate sessions, with the first session covering working memory, gambling, and motor tasks and the second session covering language, social cognition, relational processing, and emotion processing tasks. We selected the six most distinct tasks from the HCP battery of tasks. The relational processing task includes shape identification, which is likewise represented in the emotion processing task.

2.5. Parcellations

Canonical cortical parcellations were used to process all neuroimaging data with added subcortical regions. For its coarser parcellation, the Mindboggle-modified Desikan-Killiany parcellation (Desikan et al., 2006) was used for a less refined parcellation (compared with the

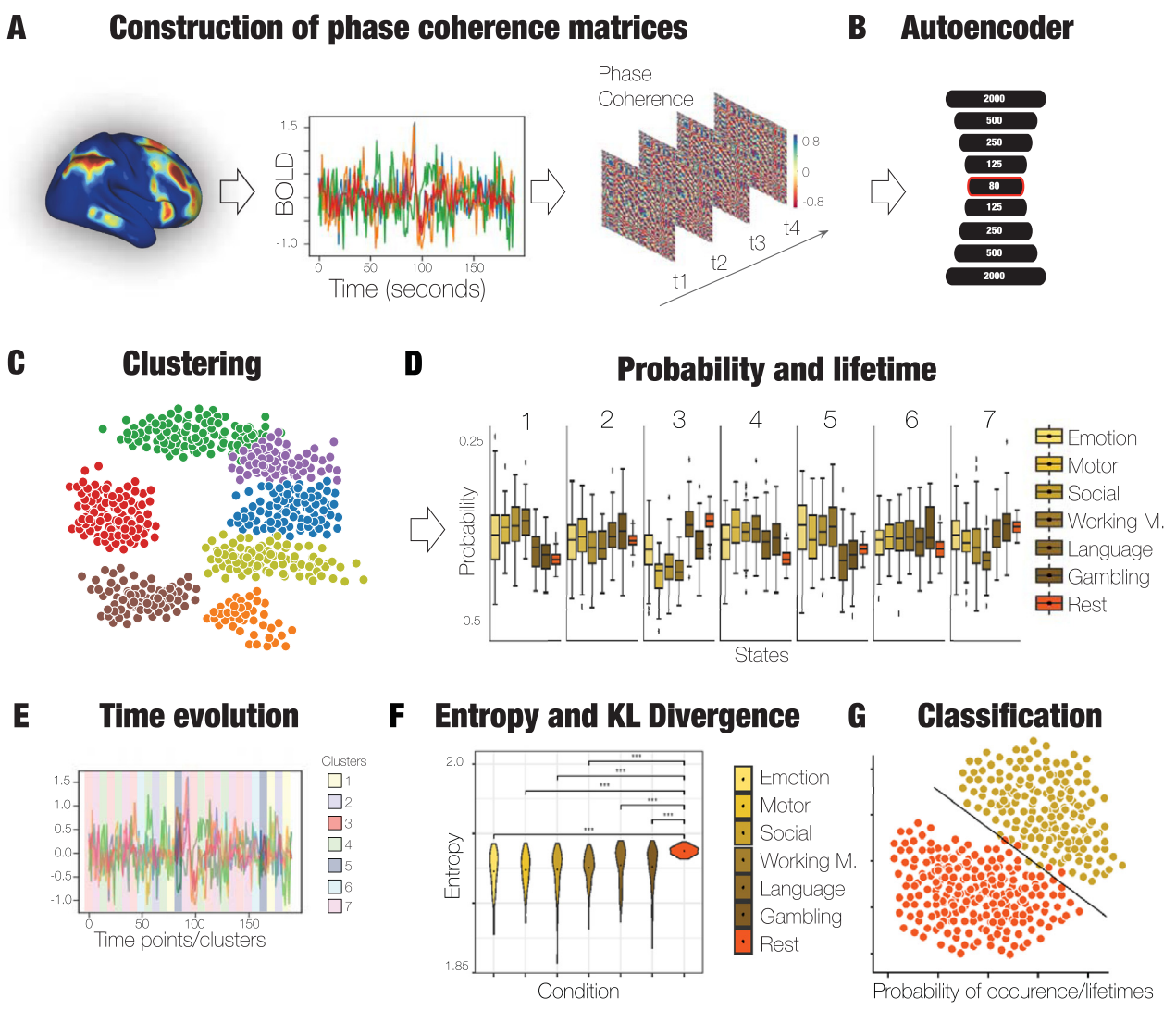

Fig. 1. *General deep autoencoder dynamical analysis (DADA) representation of the detection of dFC states from BOLD using the autoencoder and clustering model.* (A) BOLD signals from one subject at various brain areas (N = 80). BOLD signals are further smoothed with a second-order Butterworth filter and transformed into phases of a BOLD signal via the Hilbert transform, and with BOLD phase coherence connectivity, dFC is computed. At each time point t, the cosine function of the BOLD phase difference between brain area b1 and brain area b2 yields a symmetric NxN dFC(b1, b2, t) matrix. (B) A nine-layer-deep autoencoder allows for representing the connectivity vector of dFC(t) with reduced dimensionality to 80 features (previously 80 × 80 features). (C) K-means clustering model detects dFC metastable substates with the optimal number of clusters with a high silhouette score and a low sum of squared distances. (D) The mean probability and mean lifetime for each metastable substate (cluster) and each task and rest condition are obtained. (E) Time evolution schematic illustration representing the BOLD signal and metastable substates as clusters in time. (F) Entropy and Kullback-Leibler divergence represent the global whole-brain measure for each subject and condition. (G) Classification of each condition in the seven metastable substates with probability and lifetime features (individual measurable properties).

Glasser parcellation (Glasser et al., 2016), as further dimensionality reduction was performed) with a total of 62 cortical regions, 31 regions per hemisphere (Klein and Tourville, 2012). Moreover, 18 subcortical regions were added, 9 regions per hemisphere: hippocampus, amygdala, subthalamic nucleus (STN), globus pallidus internal segment (GPi), globus pallidus external segment (GPe), putamen, caudate, nucleus accumbens, and thalamus. This addition formed a final parcellation consisting of 80 regions in the DBS80 parcellation, with an exact definition in the common HCP CIFTI 'grayordinates' standard space.

2.6. Pre-processing and extraction of functional time series in fMRI resting state and task data

The detailed pre-processing protocol of the HCP resting state and task data sets is explained in detail on the HCP website (https://github.com/Washington-University/HCPpipelines/wiki/Installation-and-Usage-Instructions). To briefly summarize, the standard HCP functional pre-processing pipeline was used for the rest and task data, which employs standardized methods using FSL (FMRIB Software Library), FreeSurfer, and the Connectome Workbench software (Glasser et al., 2013; Smith et al., 2013). The pre-processing procedures contained a correction for spatial and gradient distortions and head motion, intensity normalization and bias field removal, registration to the T1 weighted structural image, transformation to the 2 mm Montreal Neurological Institute (MNI) space, and in rest data the use of the FIX artifact removal procedure (Schröder et al., 2015; Smith et al., 2013). Parameters of head motion were regressed out. Moreover, in rest data head motion noise was regressed out and structured artifacts were eliminated via ICA analysis denoising and FIX method (independent component analysis followed by FMRIB's ICA-based X-noisifier (Griffanti et al., 2014; Salimi-Khorshidi et al., 2014)). More information about the ICA-FIX method can be accessed via: https://github.com/Washington-University/HCPpipelines/blob/

master/ICAFIX/README.md. The pre-processed time series of all 'grayordinates' are in HCP CIFTI 'grayordinates' standard space and available in the surface-based CIFTI file for each participant for the resting state and each task. We employed a customised Matlab script using the ft_read_cifti function from the Fieldtrip toolbox (Oostenveld et al., 2011) to extract the average time series of all the 'grayordinates' in each region of the Glasser and DBS80 parcellations, which are represented in the HCP CIFTI 'grayordinates' standard space. The BOLD signal from both task and rest data was filtered by 2nd order Butterworth filter in the span of 0.008–0.08 Hz.

2.7. Dynamic functional connectivity

The time considerate dynamic FC matrix (dFC) is computed using BOLD phase coherence connectivity (Deco et al., 2017; Deco and Kringelbach, 2016; Ponce-Alvarez et al., 2015). The dFC has a size of NxNxT, where N is the number of brain areas—80 in the case of the DK80 parcellation—and T is the total number of time points recorded (add a time point for each task and rest). First, we filtered the BOLD signal using a second-order Butterworth filter (Fig. 1A). Then, for the phase coherence estimation, we applied the Hilbert transform to calculate the phase of the filtered BOLD signals in all areas n, $\theta(n, t)$. The phase coherence between two brain areas n and p at time t, *dFC(n, p, t),* is computed using the following equation:

$$dFC(n, p, t) = \cos\left(\theta(n, t) - \theta(p, t)\right)$$

where cos() represents the cosine function (Fig. 1A). The cosine function is useful for data rescaling, as cos(0) = 1 for areas that have a momentarily analogous phase, their BOLD signals are coordinated, and their dFC(n, p, t) is near 1. However, in the opposite case when two BOLD signals are not aligned, they are orthogonal, and their dFC(n, p, t) is approximately 0. As the phase coherence calculation does not provide directionality, the resultant matrix is undirected. Therefore, the matrix is symmetric across the diagonal, and all the significant values can be apprehended from the upper or lower triangular numbers of the matrix.

2.8. Autoencoder

To extract the dFC substates for the temporal axis, a method for their classification must be proposed. We used an unsupervised machine learning method k-means clustering (more details below). However, the dFC(t) matrix for the clustering algorithm is high-dimensional NxN and represents a 'curse of dimensionality' problem that is primarily represented by nonintuitive geometry, sparsity, and a high number of irrelevant features (individual measurable properties of observables) in high-dimensional data sets (Ronan et al., 2016). In previous studies (Deco and Kringelbach, 2016; Hansen et al., 2015; Preti et al., 2017), only the upper or lower triangular parts of the dFC(t) matrix were extracted, as the matrix is symmetric. In a more recent study (Cabral et al., 2017), the LEIDA method of the leading eigenvector $V_1(t)$ of each dFC(t) method was proposed. The leading eigenvector $V_1(t)$ possesses lower dimensionality (Nx1) compared with triangular elements (N(N-1)/2) and serves as a representant of the dominant connectivity pattern of dFC(t).

To further improve the features' representation in lower-dimensional space (see a comparison in the results), in DADA, we employed non-linear data transformation 10-layer-deep autoencoder algorithm (Fig. 1B). The artificial neural network is prepared with random weights trained together by minimizing the difference between the original data and its reconstruction. The network consists of three parts: the encoder, the code, and the decoder. In the encoder part, the model is trained to reduce the dimensions of the input data into an encoded representation. Meanwhile, the code is a layer that consists of the compressed representation of the input data. Finally, in the decoder, the model is trained to reconstruct the input data from the code layer as close to the original input as possible. Training of the network is achieved by optimizing weights with backpropagating error derivatives through the decoder and then encoder layers of the autoencoder. The encoder input space ($\mathcal{X}$) and decoder input space ($\mathcal{F}$) are defined as transitions (ϕ, ψ) minimising the reconstruction loss so that the output data difference with respect to the input data is minimised and can be represented as follows:

$$\phi : \mathcal{X} \to \mathcal{F}$$

$$\psi : \mathcal{F} \to \mathcal{X}$$

$$\phi, \psi = \arg\min_{\phi,\psi} \left\| X - \left(\psi^{\circ}\phi\right) X \right\|^2$$

The autoencoder was structured as a 2000-500-250-125-80-125-250-500-2000 schema inspired by Hinton and Salakhutdinov (2006), with the middle layer serving as the final output layer of the encoded vector E(t). We performed optimization using a gradient-based Adam optimizer and employed a mean squared error loss function (Kingma and Ba, 2014). Before training the autoencoder, we used a random undersampler algorithm to balance the data set (Lemaître et al., 2017).

2.9. Metastable substates—K-means clustering

By applying unsupervised machine learning methods on all the encoded vectors E(t) across time points, subjects, and conditions, we identified a discrete number of dFC substates.

We implemented k-means clustering with k clusters, namely seven (Figs. 1A and 4A), as seven clusters proved to be appropriate according to the list of our evaluation method which includes having several runs with a high silhouette score, short fit time, and small sum of squared distances to the closest centroid for all observations (Fig. 4B). Nevertheless, we do not, of course, claim that this number of states is a biological ground-truth. Other possible strategies exist that can guide the number of states using, for example, a measure such as free energy (Bishop, 2006) or using non-parametric methods tapping the power of infinite Hidden Markov models (Beal et al., 2001). Importantly, as shown by Vidaurre et al. (2018) the different number of states just offer different levels of detail in the brain dynamics with solutions using fewer clusters simply merging part of solutions with more clusters. We included entropy analysis for 6 and 8 states in the Supplementary material (Fig. 5) showing similar results with different ranges of entropy values.

From the clustering method, we acquired k different functional connectivity recurrent patterns. We used the silhouette score (Rousseeuw, 1987) to assess the quality and validity of the clustering separation. For a reasonable clustered structure, the average silhouette score should be at least 0.51 (Kaufman and Rousseeuw, 2005).

2.10. Modularity and average clustering coefficient

To obtain modularity measurement, we first used the Louvain algorithm for community detection in a weighted undirected graph with the best possible grouping of nodes in a network. We then optimized for modularity to measure the frequency of connections inside communities compared with connections between communities. Modularity is defined as follows:

$$Q = \frac{1}{2m} \sum_{ij} \left(A_{ij} - \frac{k_i k_j}{2m} \right) \delta\left(c_i,\ c_j\right);$$

where m stands for the number of edges, A_{ij} is the adjacency matrix, k_i is the degree of i, and δ is 1 if i and j belong to the same community and is 0 otherwise.

For each of dFCs resulting from analyzing every time point in each subject and each condition, we defined the clustering of a node in a weighted undirected graph (dFC considered as an adjacent matrix) as

the geometric average of the subgraph edge weights (Onnela et al., 2005):

$$c_u = \frac{1}{deg(u)(\deg(u)-1)} \sum_{vw} \left(w'_{uv} w'_{vw} w'_{uw}\right)^{1/3}$$

where deg(u) is the degree of u, and w' is for normalized weights by the maximum weight in the network. After obtaining the clustering coefficient for each node in a network (our dFC), we computed the average to estimate the average clustering coefficient for each dFC. All the measures were computed using the NetworkX Python package (Hagberg et al., 2008).

2.11. Entropy and Kullback-Leibler divergence

For each subject and condition, we calculated the entropy of the probability of a substate (cluster) occurrence (Fig. 1E). The entropy is computed using the following equation:

$$H(X) = -\sum p(x)\ log\ p(x)$$

where p(x) is the probability of a substate. From this, we obtained a global measure of the nature of changing from one substate to another. The higher the entropy, the less predictable the pattern of occupancy in all the substates.

For each pair of conditions, we measured the Kullback-Leibler divergence using the following equation:

$$D_{KL}(P\|Q) = \sum_{x\in X} P(x) log\left(\frac{P(x)}{Q(x)}\right)$$

where P(x) is a probability distribution for one condition, and Q(x) is a probability distribution for the second condition. We symmetrized the divergence between the two conditions as follows:

$$D_{KL} = \left(D_{KL}(P\|Q) + D_{KL}(Q\|P)/2\right)$$

From measuring the Kullback-Leibler divergence, we estimated the difference between the probability distributions for all pairs of conditions.

2.12. Classification

We run a multiclass logistic regression classification algorithm to assess the separability of each task based on two features (individual measurable properties): the probability and lifetimes for each metastable substate (Fig. 1G). We selected logistic regression, as it performed better than the k-neighbor and support vector classifier in terms of accuracy. Moreover, we employed a grid search to optimize for the best parameters and used k-fold cross-validation (five folds) to score the performance on the test (20%) and training (80%) data. The number of data points in each class was balanced with the random under-sampler algorithm (Lemaître et al., 2017).

2.13. Between conditions comparisons

To detect significant differences between groups, we utilized a permutation-based paired *t*-test. This is a non-parametric two-sample hypothesis test that applies permutations of group labels to assess the null distribution contrary to depending on the test-type standard distributions. The null distribution is estimated individually for each condition. A *t*-test is used to analyze the group comparison for each of 5000 permutations.

2.14. Data and code availability statement

The data for all tasks and rest sessions are publicly available at the HCP website (http://www.humanconnectome.org/). The codes are publicly available at https://github.com/katerinaC/brain_dynamics.

3. Results

3.1. Metastable substates

After we applied the k-means clustering (which performed better compared with the ward and dbscan clustering methods) algorithm on all the output encoded vectors E(t), we assigned each dFC to seven FC pattern groups for all subjects ($n = 100$) in all tasks and rest. From this assignment, we could obtain a mean FC pattern for each cluster—or metastable substate (Fig. 2). For each of the mean FC patterns, we calculated the normalized degree connectivity by summing all the connections for one brain area (represented by a single row in the FC matrix) divided by the number of all possible connections (number of columns in the FC matrix). By estimating the normalized degree, we could extract the connectedly dominant brain areas for each of the seven metastable substates (Supplementary Fig. 2).

The most connected substate with the highest normalized degree connectivity (0.8) in the right thalamus was substate #5. Then followed substate #1, with its highest normalized degree (0.7) in the right thalamus as well. Next was substate #4, with its highest normalized degree (0.6) in the right thalamus. Following this was substate #6, with its highest normalized degree (0.5) in the right thalamus. Regarding the less connected substates, next was substate #2, with its normalized degree (0.3) in the right thalamus. Then came substate #7, with its normalized degree (0.2) in the right superior parietal lobule. Finally, the least connected substate was represented by #3, with its highest normalized degree (0.2) in the right putamen. Overall, the most dominant brain areas with high normalized degrees in all metastable substates were the right and left thalami, right and left superior frontal lobules, and right and left precunei.

The distribution of metastable substates occurrences is represented here by the average probability of exhibiting each of the seven defined brain FC configurations in each modality (i.e. rest, language task, gambling task, social task, working memory task, emotion task, and motor task) for each subject (Fig. 3A and D). The highest probability of expressing the most connected substate, namely substate #5, was in the working memory (mean ± standard error of the mean = 0.165 ± 0.00039), social (0.158 ± 0.0004), and emotion tasks (0.163 ± 0.0006). Meanwhile, the lowest probability of occurrence was detected in the rest ($0.138 \pm 7.06 \times 10^{-5}$) and language tasks (0.126 ± 0.00048). A similar distribution revealed metastable substate #1 to have the second highest connectivity and highest prevalence for the working memory (0.169 ± 0.00027), social (0.166 ± 0.0004), and motor tasks (0.161 ± 0.00035). Correspondingly, rest ($0.126 \pm 7.\ 03 \times 10^{-5}$) still exhibited this FC pattern with the lowest probability. On the contrary, metastable substate #3, which had the lowest connectivity, exhibited the highest probability of emergence in the rest ($0.168 \pm 9.08 \times 10^{-5}$) and language (0.166 ± 0.0003) tasks and the lowest probability in the motor task (0.106 ± 0.00036). The second least connected metastable substate, namely substate #7, appeared most in the rest ($0.163 \pm 5.9 \times 10^{-5}$) and gambling tasks (0.167 ± 0.0004) and least in the working memory task (0.124 ± 0.0002). The moderately connected metastable substate #6 arose in all the modalities with fair similarity, save for the rest ($0.137 \pm 9.8 \times 10^{-5}$) and gambling tasks (0.157 ± 0.00047), which significantly differed from the remaining tasks.

Similar patterns could be observed in the metastable substate lifetimes (Fig. 3B a D). The longest mean lifetime in the most connected metastable substate, namely substate #5, was in the working memory (0.878 ± 0.00066) and motor (0.851 ± 0.0009) tasks and shortest in the language (0.816 ± 0.0009) and rest tasks (0.83 ± 0.0002). In the least connected metastable substate, namely substate #3, the longest lifetime was in the rest (0.871 ± 0.0002) and language (0.859 ± 0.0008) tasks compared with the shortest lifetime in the motor task (0.783 ± 0.0007). Furthermore, in the median metastable substate #6, a significant difference in lifetime was demonstrated between the gambling task (0.855 ± 0.00095) and rest (0.835 ± 0.0002).

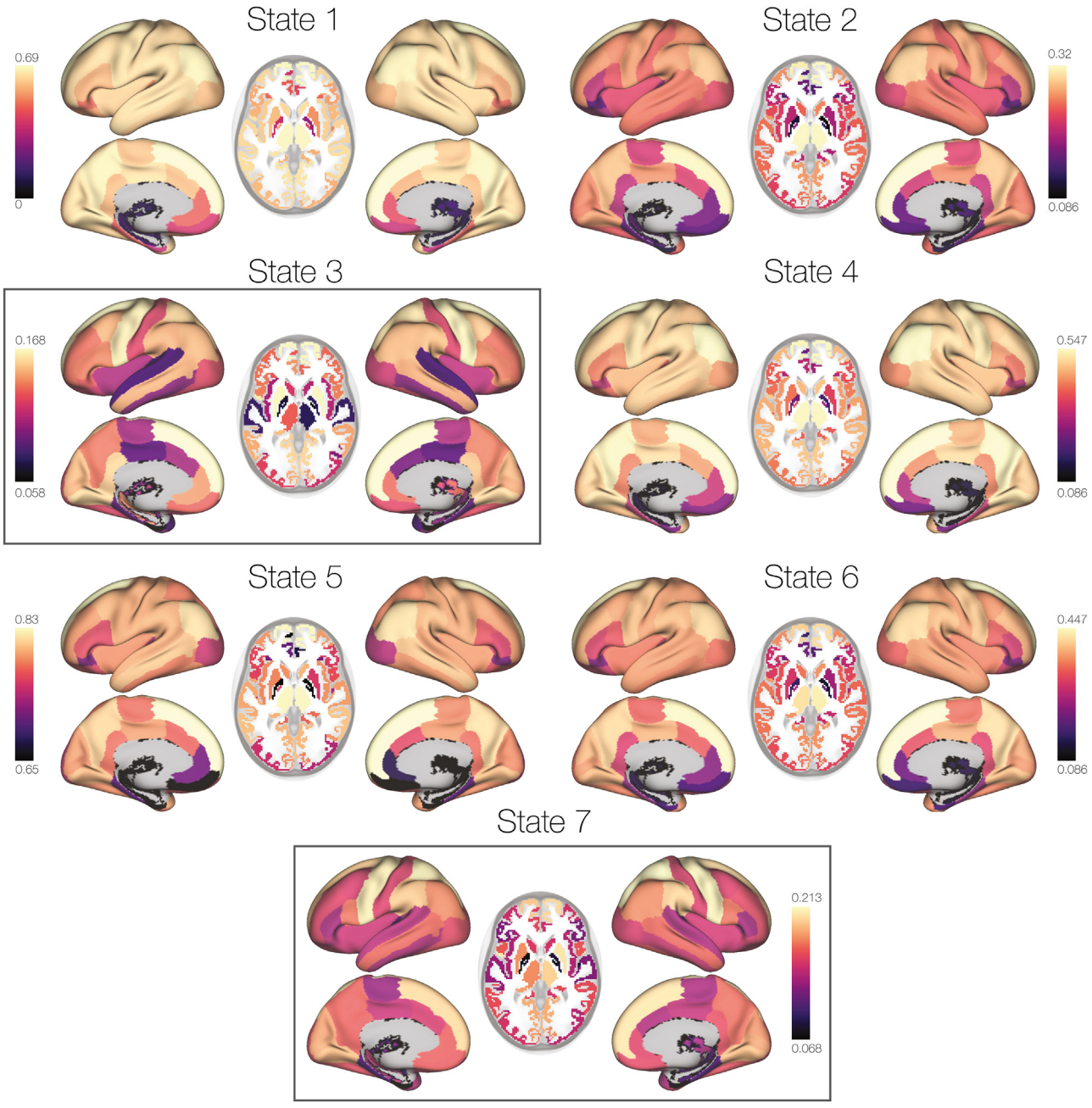

Fig. 2. *FC patterns.* Seven reccurring metastable substates, as identified by clustering the reduced DADA's vectors of all participants and time points. Each surface brain render represents each of the seven states as a normalized connectivity degree. State renders 1 and 5 depict connected states across modalities compared with state renders 3 and 7, which denote the most disconnected states that are most pronounced in rest compared with during cognitive tasks. Rendered with Connectome Workbench (available at https://www.humanconnectome.org/software/connectome-workbench).

By comparing the DADA's autoencoder dimensionality reduction with the previously suggested LEIDA method (Cabral et al., 2017; Deco et al., 2019), we could assess the capacity and effectiveness of the clustering separation (Fig. 4C and D). The average silhouette score for the seven clusters with the autoencoder (0.54) was above the threshold of 0.51 necessary for a reasonable clustered structure (Fig. 4A). Meanwhile, when using the LEIDA method (0.098), the value is below the point (Fig. 4B).

3.2. Switching between metastable substates

We calculated the probability of transitioning from one metastable substate to another for each cognitive task and rest. These calculated probabilities are represented in transition matrices, where rows (i) represent the probability transition from substate i to another substate j represented by columns (j); therefore, each element of position (i, j) represents the transition probability p_{ij} (Fig. 5A).

In rest, the highest probability was found for switching to substates #3 and #7 (0.16 and 0.17), while the lowest probability was found for switching to substates #1 and #4 (0.12 and 0.13). In the emotion task, the lowest probability was to remain in substate #3 (0.11), while the highest probability was in transitioning from substate #3 to substate #5 (0.18). Meanwhile, in gambling task, the least probable transition was from substate #5 to #1 (0.1), while the most probable transition was from substates #1, #3, and #5 to #7 (0.17). During the language task, the highest probability of transitioning was from substates #1 and

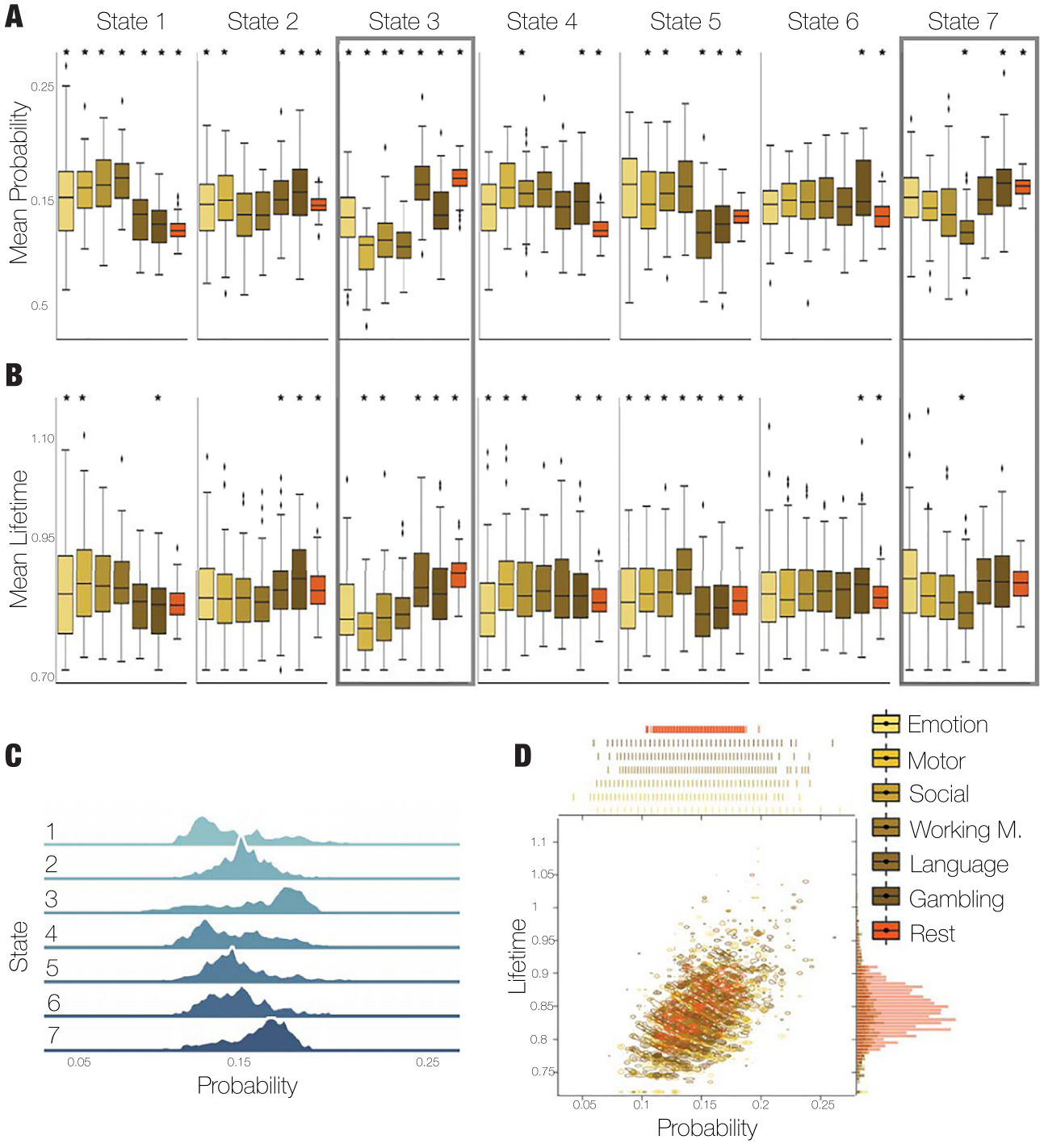

Fig. 3. *Probabilities and lifetimes of metastable substates under different conditions.* (A) Boxplot representing probability of a metastable substate occurring for each modality (i.e. rest, language task, gambling task, social task, working memory task, emotion task, and motor task), each substate, and each subject. Metastable substates #1, #4, and #5 are less likely for rest compared with other tasks, especially the motor, working memory, and social tasks; however, metastable substates #3 and #7 are highly likely for the rest, language, and gambling tasks compared with the other tasks. Stars indicate a significant difference from all the other conditions (permutation-based paired *t*-test [$p < 0.005$] with Bonferroni correction) (B) Boxplot representing lifetime of a metastable substate for each tested modality. The longest dwell time for the rest and language tasks is in state #3. Overall, the main significant differences, as estimated by a permutation-based paired *t*-test with Bonferroni correction, in the probabilities and lifetimes of states are between the group of tasks and the group of rest. Stars indicate a significant difference from all the other conditions (permutation-based paired *t*-test [$p < 0.005$] with Bonferroni correction) (C) Probability distributions for each of the seven metastable substates with shifted distributions in states #3 and #7. (D) Substate probabilities plotted against substate lifetimes, demonstrating the different distributions in rest compared with the tasks, as rest distribution is more compact in terms of both probability and lifetime, making it more easily distinguishable from the task distributions.

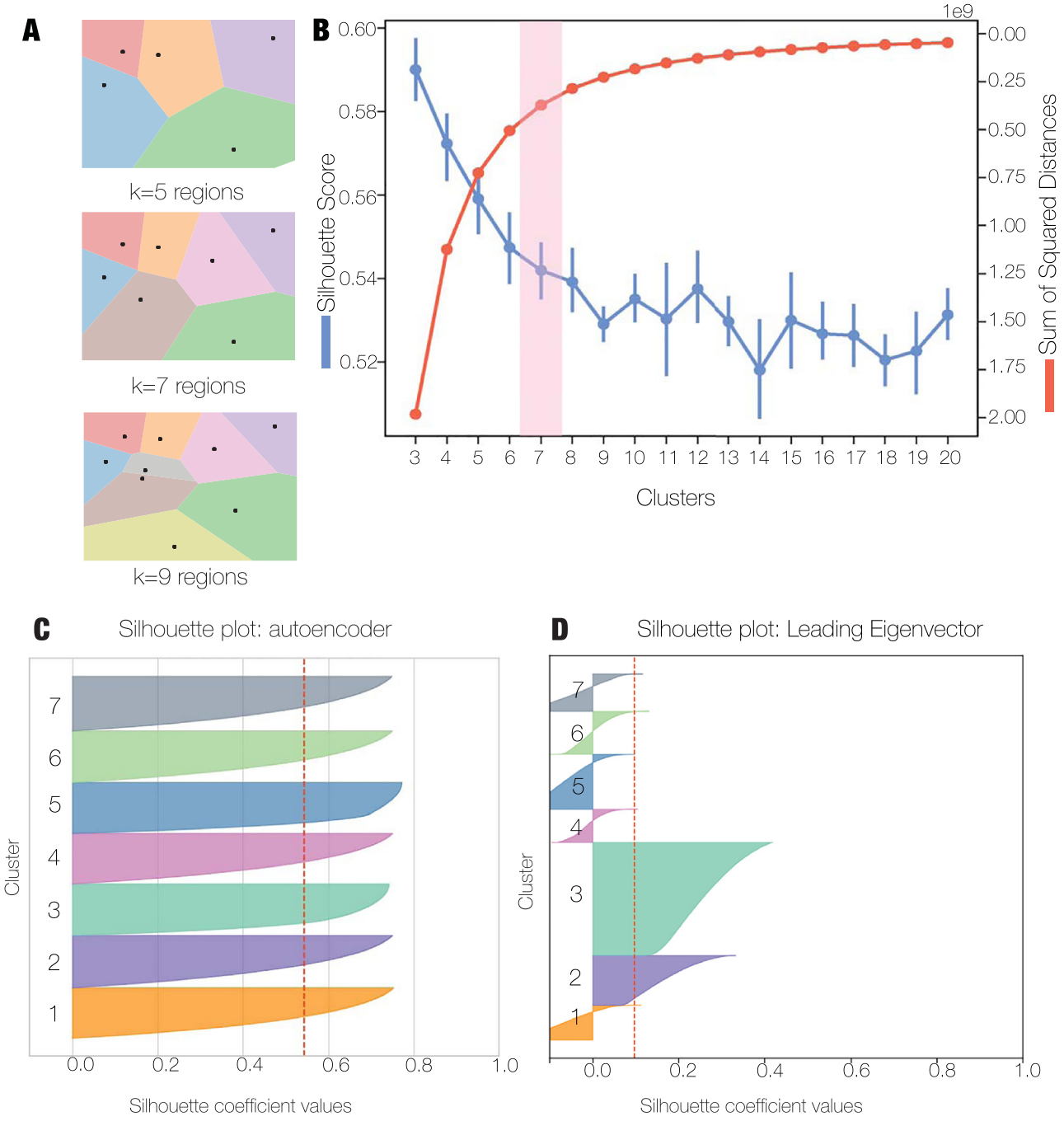

Fig. 4. *Clustering.* (A) Voronoi diagram. Schematic representation of the partition of a 2D plane into k regions, in which each point in each cell is closer to its cluster centre than any other centre. The seven cluster regions represent the most evenly spread partition, which likewise performed the best with the HCP data, as observable from the high silhouette coefficient. (B) The score of two quality measurements from 10 runs: the sum of squared distances and silhouette score. (C) Average silhouette coefficient values of the k-means clustering algorithm in DADA with the use of the autoencoder for dimension reduction (0.54) (D) Average silhouette coefficient values of the k-means clustering algorithm in LEIDA with the use of the leading eigenvector for dimension reduction (0.098).

#6 to #3 and from #5 to #2 (0.17), while the lowest probability of transitioning was from substates #4 and #5 to #5 (0.11). Next, in the motor task, the lowest probability was in switching from substate #3 to #3 (0.08), while the highest probability was in switching from substate #3 to #1 and #4 (0.17). Furthermore, in the social task, the most probable transition occurred from substates #3 and #5 to #1 (0.17), while the least probable transition occurred from substate #3 to #3 (0.1). Last, in the working memory task, the highest probability was in transitioning from substate #5 to #5 (0.18), while the lowest probability was in transitioning to substate #3 (0.11 and 0.12).

3.3. Modularity and average clustering coefficient

We calculated graph metrics for all weighted FC matrices for each condition and substate, finding many statistical differences between tasks and rest in each substate that pass a permutation-based *t*-test with $p < 0.005$ (Supplementary Table 1). Modularity measures the structure of a network by dividing the network into communities called modules (Zamora-López et al., 2016) (Supplementary Fig. 3A). High modularity indicates high connectedness between nodes within the same module and was present in substates #2, #3, #6, and #7, which have low global

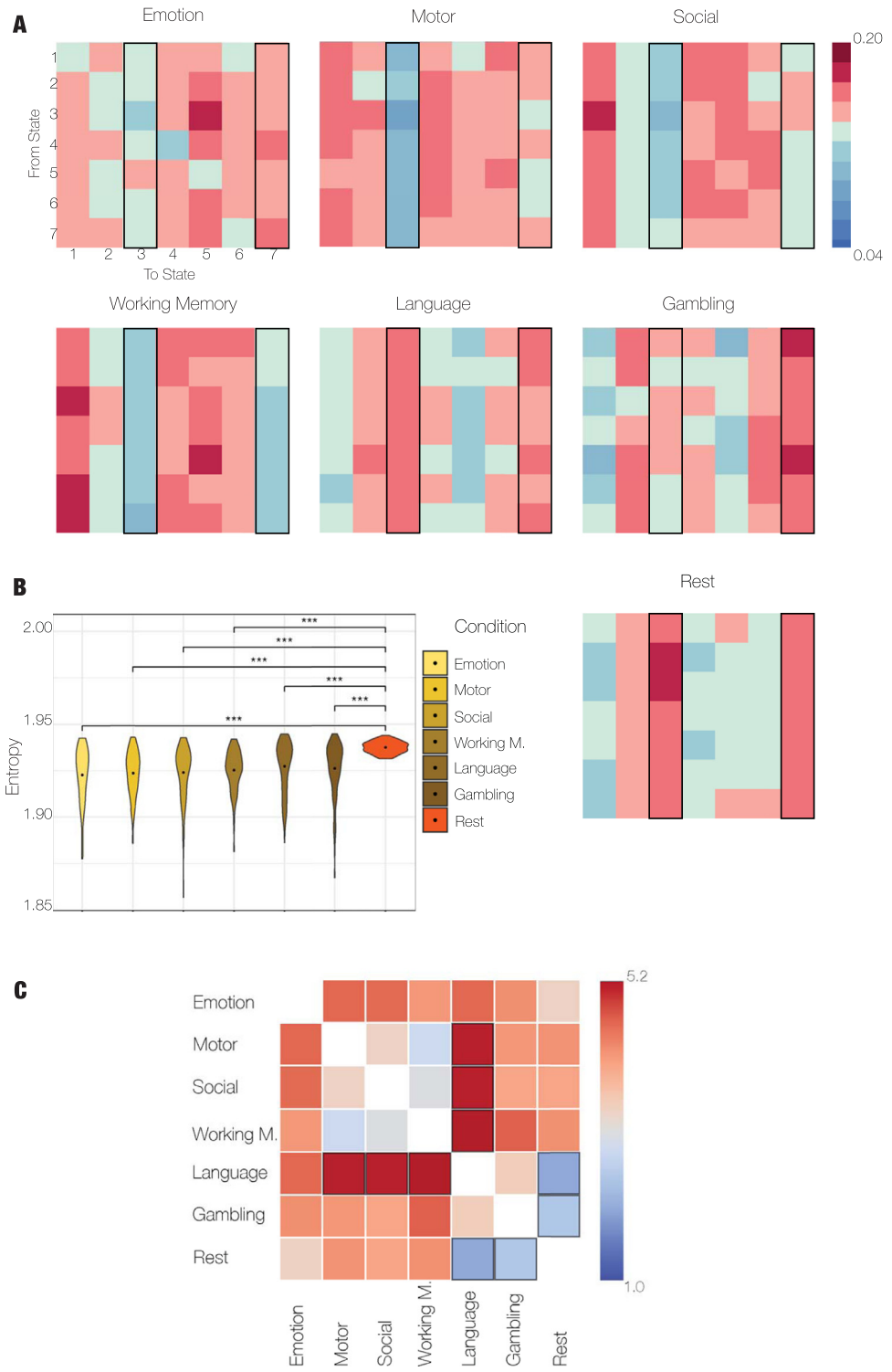

Fig. 5. *Transitions, entropy, and Kullback-Leibler divergence*. (A) In transition matrices, rows display the probability of being in each substate, while columns indicate the probability of switching to any of the other seven substates. In rest, the highest probability of switching is to substates #3 and #7 (from 0.15 to 0.17). Meanwhile, in the motor, social, and working memory tasks, the switching probabilities are low in substate #3 and higher in substates #1 and #4. Furthermore, the working memory task switching probabilities are the most opposite to those of rest, with the most likely switching to substates #1, #4, and #5 and the least likely switching to substates #3 and #7. In the language, emotion, and gambling tasks, these differences are more subtle, especially in the language task, which has the switching probabilities most similar to those of rest. In the emotion task, the highest probability of switching is from substate #3 to substate #5 (0.18). Meanwhile, in the language task, the most plausible switching is from substates #1 and #5 to substate #7 (0.17). (B) The entropy of the probability of a metastable substate occurrence was calculated for each subject and condition. A significant difference was present in terms of entropy (permutation-based paired *t*-test, $p < 0.005$, Bonferroni correction) between the rest and the task conditions. High entropy and a larger dynamical repertoire at rest suggest a greater preparedness than that observed during a task. (C) Estimation of the Kullback-Leibler divergence obtains the differences between probability distributions for all pairs of conditions. The language task demonstrates the largest distance of probability distribution from all the tasks, with the most distant distribution to the motor, social, and working memory tasks (5.2). Meanwhile, the rest probability distribution is most similar to the language (2.0) and gambling (2.5) tasks.

connectivity. Meanwhile, low modularity was displayed in substates #1 and #5, which have high global connectivity (Supplementary Fig. 3B and C).

The clustering coefficient measures the number of closed triplets over the number of all possible triplets in a graph (Zamora-López et al., 2011) (Supplementary Fig. 4A) and indicates how close the graph is to a complete graph. For our weighted undirected graph, the clustering was characterized as the geometric average of the subgraph edge weights. The highest average clustering coefficient was detected in substates #1 and #5, which had high global connectivity (Supplementary Fig. 4B and C).

3.4. Entropy and Kullback-Leibler divergence

Considering global measures such as entropy and Kullback-Leibler divergence, we found a statistical significance in terms of entropy between rest (mean ± standard error = 1.94 ± 0.0003, $p < 0.005$) and all the cognitive tasks (i.e. emotion, gambling, language, social, emotion, and working memory) (Fig. 5B).

The language task exhibited the largest distance of probability distribution from all the tasks as calculated by Kullback-Leibler divergence, with the furthest distance to the motor, social, and working memory

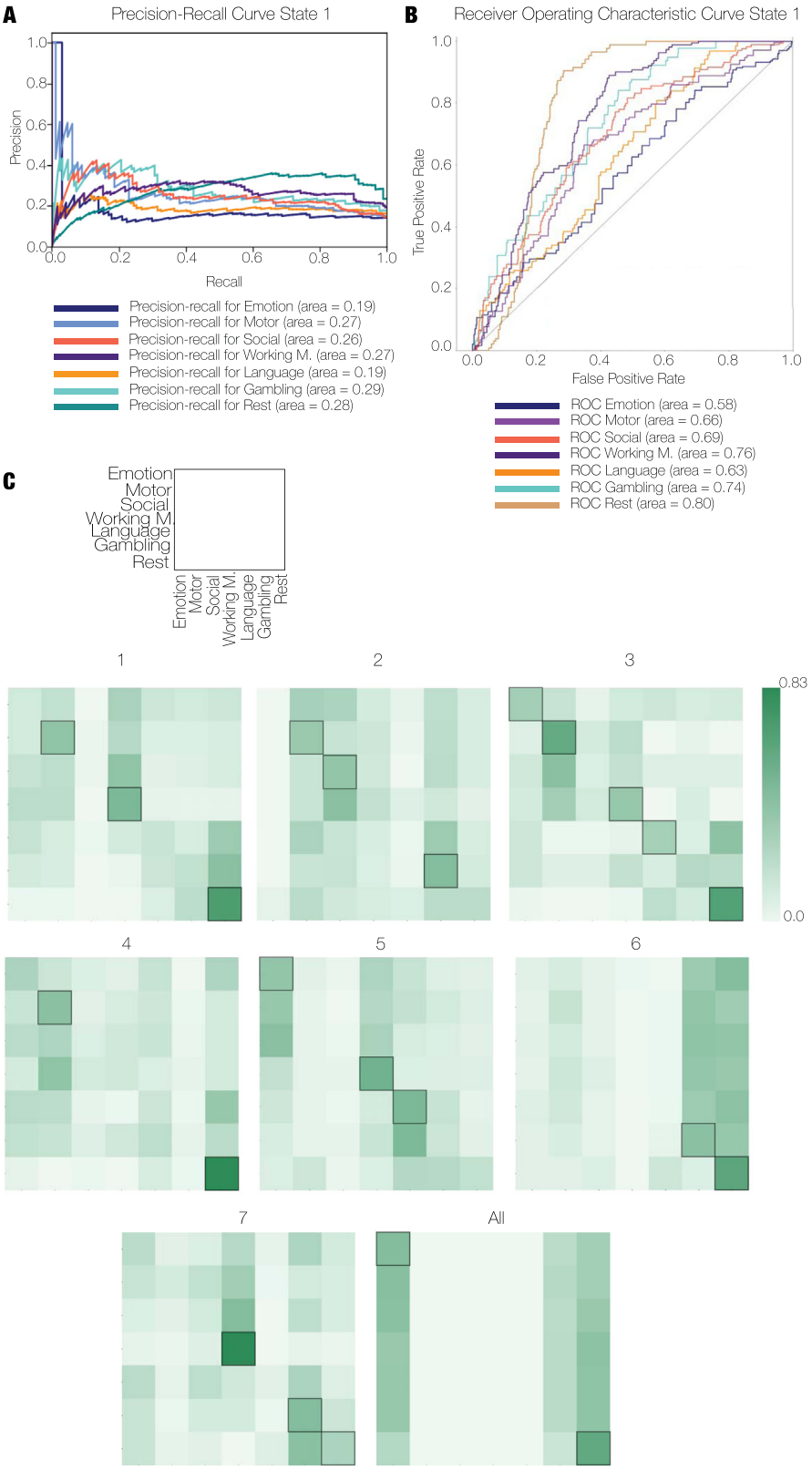

Fig. 6. *Classification metrics*. (A) Precision-recall curve for multi-class classification for substate #1: precision = true positives / true positives + false positives; recall = true positives / true positives + false negatives). The best performance in the precision-recall was the gambling task class with an area under the curve (AUC) = 0.29 (B) Receiver operating characteristic curve for multi-class classification for the conditions classified within the substate #1. Presenting the true positive rate (sensitivity = true positive / (true positive + false negatives)) over the false positive rate (expectancy = false positive / (false positive + true negative)). The best performing was rest, with AUC = 0.80, while the second best was the working memory task, with AUC = 0.76. Both results are at an acceptable discrimination level. More precision-re- call curves and ROC are presented in the supplementary materials. (C) Legend of the confusion matrices with task order. Normalized confusion matrices displaying the true labels (rows) against the predicted labels (columns) for each substate and task. The model could very well distinguish rest in substates #1 (0.69), #3 (0.66), #4 (0.83), #6 (0.62), and all substates (0.6). Moreover, it distinguished the language task with high precision in substates #1 (0.49), #5 (0.53), and #7 (0.83). Meanwhile, the motor task was well separated in substates #3 (0.59) and #4 (0.43).

tasks (5.2). On the other hand, the rest probability distribution was closest to the language (2.0) and gambling (2.5) tasks (Fig. 5C).

3.5. Classification

Good separation between rest and the tasks with the probability of occurrence feature was present within substates #1 (Fig. 6A and B), #3, #4, and #6 (supplementary material) when the classification was performed for each substate separately; however, there was also fair separation for rest when the classification was performed on all substates (supplementary material). The degree of correctly predicted classes within normalized confusion matrices (Fig. 6C) for the rest condition in these substates was greater than 60%: in particular, #1 (69%), #3 (0.66%),

#4 (83%), #6 (62%), all (60%). Other notable predicted labels occurred within the gambling task in substates #2 (46%), #7 (46%), #6 (43%), the motor task in substate #3 (59%), the working memory task in substate #5 (53%), the language task in substate #5 (49%), and the emotion task within all substates (46%). Overall, the cross-validated accuracies of test data for each substate classification passed the baseline (14.3%) of random prediction: #1 (29.9%), #2 (20.6%), #3 (37.9%), #4 (26.9%), #5 (26.7%), #6 (18.5%), #7 (28.7%). In the case of all substates, the performance was lower (16.5%), though still above the baseline level. More classification performance scores are presented in the supplementary materials (precision-recall plots and receiver operating characteristic curves).

4. Discussion

Here, we provide empirical evidence for the prediction of a recent cognitive theory of the higher cognitive load, postulating that fulfilling tasks offers higher information integration (Gilson et al., 2018a). We found that while performing various cognitive tasks, there was a higher probability of more connected metastable substates compared with in rest (i.e. substates with a higher clustering coefficient and lower modularity are more probable in a task than in rest).

These results can be considered in the context of the common distinction in cognitive theory between on-line and off-line cognition modes. Rest represents the off-line cognition mode with less connected substates with higher modularity (separation into communities) and a lower clustering coefficient (local connectivity), which suggests an increase in the number of major hubs in these more globally disconnected substates. In contrast, the on-line cognitive mode is found during tasks characterized by more globally connected substates with a higher clustering coefficient and lower modularity.

The novel DADA framework makes it possible to identify the metastable substates and how their probability changes during tasks and rest. This is made possible via the implementation of improved dimensionality reduction in the form of a multilayer deep autoencoder. This method proved to reduce the dimensions and better prepare the data for further clustering compared with the previously introduced LEIDA method. The clustering silhouette score performance indicated acceptable division into clusters that served to detect dominating discrete dFC patterns in time, even if the dFC evolution switches more fluently. By using an autoencoder algorithm, the DADA method lowers the signal-to-noise ratio in the FCD and provides insight into repeating topological changes in the FC.

Our results of the highest normalized degrees in each of the metastable substates revealed that the thalamus plays a major role in all the substates that represent higher connectivity. This corroborates the notion of the thalamus being an important hub that connects major cognitive areas, such as the hippocampus, striatum, medial prefrontal cortex, and regions that have a regulatory role, such as the septum, hypothalamus, and areas in the brainstem. Hypothalamus cells serve as a regulating agent in the flow of information in the forebrain as a capacity of internal states, such as emotion and sleep. Moreover, the thalamus controls episodic memory, which is a prerequisite for many cognitive roles, such as problem-solving, inference, and memory consolidation (Varela, 2014; Varela et al., 2014). Other identified regions with high normalized degrees, such as the (left and right) superior frontal gyri, precunei, superior parietal gyri, putamina, and lingual gyri, were previously recognized as structural hubs of the adult human brain (data from group averaged connectome from the HCP data set) that play an important role in information integration and communication between brain regions (Bell and Shine, 2016; Gong et al., 2009; Hagmann et al., 2008; Oldham and Fornito, 2019; van den Heuvel and Sporns, 2013, 2011). These findings validate the robustness of our method, which can detect important functional and structural hubs.

We identified seven prevalent stochastic metastable substates of typical and perpetual brain states. They appear recurrently and have different probabilities of switching from one substate to another. However, the number of substates and their respective FC representations as detected by our method depend on the dimensionality of the input data—particularly the parcellation and temporal scale. We used coarse parcellation with 80 regions of interest with a repetition time (TR) of 0.72 s. The appropriate number of metastable substates and their respective pattern might change with different imaging and post-process conditions.

Lower modularity and a higher clustering coefficient in highly connected substates describe high global connectedness; however, a low clustering coefficient and high modularity in low connected substates indicate the importance of the identified hubs that are locally connected and create more separate communities. In tasks, the brain more often visits substates that are highly connected with lower modularity and a higher clustering coefficient compared with in rest. This indicates brain activity in more rigid resting state networks (RSNs) or intrinsic connectivity networks (ICNs) (Seeley et al., 2007) during rest, while during tasks, the brain overcomes these networks and forms new connections between different brain areas, creating novel networks with lowered modularity. There are many significant differences in both the modularity and clustering coefficient metrics between the task and rest conditions, so they represent a promising method for distinguishing between cognitive modalities. Further graph analysis with a dynamic aspect (Gilson et al., 2019, 2018b) of the distinct metastable substate patterns deserves greater attention in future studies.

Transition probability indicates the preferred switching from one state to another or eventually dwelling in one state during a prolonged period. Moreover, it suggests how much energy is necessary for such switching to occur. The more probable the transition from one state to another, the less energy required to make the transition, while less probable transitions require more energy consumption (Gu et al., 2015). In tasks, there is a higher preference for certain states compared with rest, with a higher probability of switching to or remaining in these states indicating fewer energy demands. However, if switching to a less probable state in a task, more energy is required compared with rest, where the probability of switching to different states is more similar. These results indicate the preference of a specific substate or set of substates in tasks compared with rest that is more prone to switching between different substates.

The significantly largest entropy in rest endorses findings from previous studies (Escrichs et al., 2019; Saenger et al., 2018). In rest, the brain's regime is orchestrated by increased metastability, meaning that the system's dynamics are more unpredictable or random and span a wider dynamical regime than in cognitive tasks. Furthermore, higher entropy means that sampling for the rest condition would yield higher uncertainty regarding the system's state at a particular time-point (Carhart-Harris et al., 2014). This result likewise aligns with computational models of the whole-brain network dynamics that were optimally stimulated for the rest condition, when the parameters of the system are in a critical range, while the system is constantly repelled away from equilibrium (Breakspear, 2017; Cabral et al., 2014; Deco and Jirsa, 2012). Higher entropy might appear to be disadvantageous with increased randomness, but in a condition when no instant reaction is required, such a state is desirable, as higher entropy allows for more easily learning a new task with lower specialized performance (Fong et al., 2019; Nogueira et al., 2018; Yang et al., 2019). Rest can, therefore, be viewed as a preparatory off-line state that provides the grounds for more efficient performance during a specific on-line task.

The results of Kullback-Leibler divergence revealed that the language task probability distribution is the furthest from other tasks but is not that far from the rest condition. This might be due to the characteristics of the task (Binder et al., 2011), where half of the time, participants listen to short stories and must then answer some semantical questions about the story. This scenario appears to be cognitively similar to resting when one is occupied with flickering thoughts that might generate short stories. The rapid thought switches that occur during rest can be

related to the language task with its second control part consisting of basic arithmetic calculations.

The classification scores indicate that our method is suitable for distinguishing between tasks and rest; however, it is not fine enough to classify between cognitive tasks. Yet, there is a higher-performing prediction within the substates compared with all substates overall, in which only three classes dominated, indicating promising results for future analyses.

In conclusion, we have presented empirical evidence for a major cognitive theory. Specifically, we have demonstrated that the human adult brain during rest tends to occupy metastable substates with lower connectivity, higher modularity, and a low clustering coefficient. On the other hand, during attention-demanding tasks, the brain is more likely to shift to highly connected metastable substates. Thus, all conditions visit each of the defined metastable substates, indicating a rearranging from on-line to off-line cognition. High entropy during rest implies a preparatory period with high amenability towards the upcoming on-line cognitive task.

Data and code availability statement

The data for all tasks and rest sessions are publicly available at the HCP website (http://www.humanconnectome.org/).

The codes are publicly available at https://github.com/katerinaC/brain_dynamics.

Declaration of Competing Interest

The authors declare no competing financial interests.

Credit authorship contribution statement

Katerina Capouskova: Conceptualization, Methodology, Software, Validation, Formal analysis, Investigation, Data curation, Writing – original draft, Visualization. **Morten L. Kringelbach:** Conceptualization, Data curation, Writing – review & editing, Supervision, Visualization. **Gustavo Deco:** Conceptualization, Methodology, Formal analysis, Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Data availability

Data are publicly available.

Acknowledgments

G.D. was supported by the Spanish Research Project AWAKENING: using whole-brain models perturbational approaches for predicting external stimulation to force transitions between different brain states, ref. PID2019-105772GB-I00 /AEI/10.13039/501100011033, financed by the Spanish Ministry of Science, Innovation and Universities (MCIU), State Research Agency (AEI); and K.C and G.D. were also supported by the European Union's Horizon 2020 FET Flagship Human Brain Project (grant agreement number 945539, HBP SGA3). M.L.K. was supported by the ERC Consolidator Grant: CAREGIVING (No. 615539), Center for Music in the Brain, funded by the Danish National Research Foundation (DNRF117), and Centre for Eudaimonia and Human Flourishing funded by the Pettit and Carlsberg Foundations. The funding bodies had neither influence in the research design, data collection, and analysis nor will in preparing and publishing the manuscript.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2022.119489.

References

Alderson, T.H., Bokde, A.L.W., Kelso, J.A.S., Maguire, L., Coyle, D., 2020. Metastable neural dynamics underlies cognitive performance across multiple behavioural paradigms. Hum. Brain Mapp. 41, 3212–3234. doi:10.1002/hbm.25009.

Barch, D.M., Burgess, G.C., Harms, M.P., Petersen, S.E., Schlaggar, B.L., Corbetta, M., Glasser, M.F., Curtiss, S., Dixit, S., Feldt, C., Nolan, D., Bryant, E., Hartley, T., Footer, O., Bjork, J.M., Poldrack, R., Smith, S., Johansen-Berg, H., Snyder, A.Z., Van Essen, D.C., 2013. Function in the human connectome: task-fMRI and individual differences in behavior. Neuroimage 80, 169–189. doi:10.1016/J.NEUROIMAGE.2013.05.033.

Beal, M., Ghahramani, Z., Rasmussen, C., Dietterich, T., Becker, S., Ghahramani, Z., 2001. The infinite Hidden Markov model. Advances in Neural Information Processing Systems. MIT Press.

Bell, P.T., Shine, J.M., 2016. Subcortical contributions to large-scale network communication. Neurosci. Biobehav. Rev. 71, 313–322. doi:10.1016/J.NEUBIOREV.2016.08.036.

Binder, J.R., Gross, W.L., Allendorfer, J.B., Bonilha, L., Chapin, J., Edwards, J.C., Grabowski, T.J., Langfitt, J.T., Loring, D.W., Lowe, M.J., Koenig, K., Morgan, P.S., Ojemann, J.G., Rorden, C., Szaflarski, J.P., Tivarus, M.E., Weaver, K.E., 2011. Mapping anterior temporal lobe language areas with fMRI: a multi-center normative study. Neuroimage 54, 1465. doi:10.1016/J.NEUROIMAGE.2010.09.048.

Bishop, C.M., 2006. Pattern Recognition and Machine Learning. Springer New York, NY, New York.

Breakspear, M., 2017. Dynamic models of large-scale brain activity. Nat. Neurosci. 20, 340–352. doi:10.1038/nn.4497.

Cabral, J., Hugues, E., Sporns, O., Deco, G., 2011. Role of local network oscillations in resting-state functional connectivity. Neuroimage 57, 130–139. doi:10.1016/j.neuroimage.2011.04.010.

Cabral, J., Kringelbach, M.L., Deco, G., 2014. Exploring the network dynamics underlying brain activity during rest. Prog. Neurobiol. doi:10.1016/j.pneurobio.2013.12.005.

Cabral, J., Vidaurre, D., Marques, P., Magalhães, R., Moreira, P.S., Miguel Soares, J., Deco, G., Sousa, N., Kringelbach, M.L., 2017. Cognitive performance in healthy older adults relates to spontaneous switching between states of functional connectivity during rest. Sci. Rep. 7, 5135. doi:10.1038/s41598-017-05425-7.

Carhart-Harris, R.L., Leech, R., Hellyer, P.J., Shanahan, M., Feilding, A., Tagliazucchi, E., Chialvo, D.R., Nutt, D., 2014. The entropic brain: a theory of conscious states informed by neuroimaging research with psychedelic drugs. Front. Hum. Neurosci. 8 (20). doi:10.3389/fnhum.2014.00020.

Chang, C., Glover, G.H., 2010. Time–frequency dynamics of resting-state brain connectivity measured with fMRI. Neuroimage 50, 81–98. doi:10.1016/j.neuroimage.2009.12.011.

Chaudhuri, R., Fiete, I., 2016. Computational principles of memory. Nat. Neurosci. 19, 394–403. doi:10.1038/nn.4237.

Chaudhuri, R., Gerçek, B., Pandey, B., Peyrache, A., Fiete, I., 2019. The intrinsic attractor manifold and population dynamics of a canonical cognitive circuit across waking and sleep. Nat. Neurosci. 22, 1512–1520. doi:10.1038/s41593-019-0460-x.

Costa, F.A.E., Rocha, L.M., 2005. Introduction to the special issue: embodied and situated cognition. Artif. Life 11, 5–11. doi:10.1162/1064546053279035.

Deco, G., Cabral, J., Woolrich, M.W., Stevner, A.B.A., van Hartevelt, T.J., Kringelbach, M.L., 2017. Single or multiple frequency generators in on-going brain activity: a mechanistic whole-brain model of empirical MEG data. Neuroimage 152, 538–550. doi:10.1016/J.NEUROIMAGE.2017.03.023.

Deco, G., Cruzat, J., Cabral, J., Tagliazucchi, E., Laufs, H., Logothetis, N.K., Kringelbach, M.L., 2019. Awakening: predicting external stimulation to force transitions between different brain states. Proc. Natl. Acad. Sci. USA 116, 18088–18097. doi:10.1073/pnas.1905534116.

Deco, G., Jirsa, V.K., 2012. Ongoing cortical activity at rest: criticality, multistability, and ghost attractors. J. Neurosci. 32, 3366–3375. doi:10.1523/JNEUROSCI.2523-11.2012.

Deco, G., Kringelbach, M.L., 2016. Metastability and coherence: extending the communication through coherence hypothesis using a whole-brain computational perspective. Trends Neurosci. 39, 432. doi:10.1016/J.TINS.2016.04.006.

Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., Killiany, R.J., 2006. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31, 968–980. doi:10.1016/J.NEUROIMAGE.2006.01.021.

Escrichs, A., Sanjuán, A., Atasoy, S., López-González, A., Garrido, C., Càmara, E., Deco, G., 2019. Characterizing the dynamical complexity underlying meditation. Front. Syst. Neurosci. 13, 27. doi:10.3389/fnsys.2019.00027.

Fong, A.H.C., Yoo, K., Rosenberg, M.D., Zhang, S., Li, C.S.R., Scheinost, D., Constable, R.T., Chun, M.M., 2019. Dynamic functional connectivity during task performance and rest predicts individual differences in attention across studies. Neuroimage 188, 14–25. doi:10.1016/J.NEUROIMAGE.2018.11.057.

Gallego, J.A., Perich, M.G., Miller, L.E., Solla, S.A., 2017. Neural manifolds for the control of movement. Neuron 94, 978–984. doi:10.1016/j.neuron.2017.05.025.

Gao, P., Ganguli, S., 2015. On simplicity and complexity in the brave new world of large-scale neuroscience. Curr. Opin. Neurobiol. 32, 148–155. doi:10.1016/j.conb.2015.04.003.

Gilson, M., Deco, G., Friston, K.J., Hagmann, P., Mantini, D., Betti, V., Romani, G.L., Corbetta, M., 2018a. Effective connectivity inferred from fMRI transition dynamics during movie viewing points to a balanced reconfiguration of cortical interactions. Neuroimage 180, 534–546. doi:10.1016/J.NEUROIMAGE.2017.09.061.

Gilson, M., Kouvaris, N.E., Deco, G., Mangin, J.F., Poupon, C., Lefranc, S., Rivière, D., Zamora-López, G., 2019. Network analysis of whole-brain fMRI dynamics:

a new framework based on dynamic communicability. Neuroimage 201, 116007. doi:10.1016/J.NEUROIMAGE.2019.116007.
Gilson, M., Kouvaris, N.E., Deco, G., Zamora-López, G., 2018b. Framework based on communicability and flow to analyze complex network dynamics. Phys. Rev. E 97, 052301. doi:10.1103/PhysRevE.97.052301.
Glasser, M.F., Coalson, T.S., Robinson, E.C., Hacker, C.D., Harwell, J., Yacoub, E., Ugurbil, K., Andersson, J., Beckmann, C.F., Jenkinson, M., Smith, S.M., Van Essen, D.C., 2016. A multi-modal parcellation of human cerebral cortex. Nature 536, 171–178. doi:10.1038/nature18933.
Glasser, M.F., Sotiropoulos, S.N., Wilson, J.A., Coalson, T.S., Fischl, B., Andersson, J.L., Xu, J., Jbabdi, S., Webster, M., Polimeni, J.R., Van Essen, D.C., Jenkinson, M., 2013. The minimal preprocessing pipelines for the Human Connectome Project. Neuroimage 80, 105–124. doi:10.1016/J.NEUROIMAGE.2013.04.127.
Goelman, G., Dan, R., Růžička, F., Bezdicek, O., Růžička, E., Roth, J., Vymazal, J., Jech, R., 2017. Frequency-phase analysis of resting-state functional MRI. Sci. Rep. 7, 43743. doi:10.1038/srep43743.
Gong, G., He, Y., Concha, L., Lebel, C., Gross, D.W., Evans, A.C., Beaulieu, C., 2009. Mapping anatomical connectivity patterns of human cerebral cortex using *in vivo* diffusion tensor imaging tractography. Cereb. Cortex 19, 524–536. doi:10.1093/cercor/bhn102.
Griffanti, L., Salimi-Khorshidi, G., Beckmann, C.F., Auerbach, E.J., Douaud, G., Sexton, C.E., Zsoldos, E., Ebmeier, K.P., Filippini, N., Mackay, C.E., Moeller, S., Xu, J., Yacoub, E., Baselli, G., Ugurbil, K., Miller, K.L., Smith, S.M., 2014. ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging. Neuroimage 95, 232–247. doi:10.1016/J.NEUROIMAGE.2014.03.034.
Gu, S., Pasqualetti, F., Cieslak, M., Telesford, Q.K., Yu, A.B., Kahn, A.E., Medaglia, J.D., Vettel, J.M., Miller, M.B., Grafton, S.T., Bassett, D.S., 2015. Controllability of structural brain networks. Nat. Commun. 6, 8414. doi:10.1038/ncomms9414.
Hagberg, A.A., Schult, D.A., Swart, P.J., Varoquaux, G., Vaught, T., Millman, J., 2008. Exploring network structure, dynamics, and function using NetworkX. In: Proceedings of the 7th Python in Science Conference. Los Alamos National Lab, Los Alamos, NM, United States, pp. 11–15.
Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C.J., Wedeen, V.J., Sporns, O., 2008. Mapping the structural core of human cerebral cortex. PLoS Biol. 6, e159. doi:10.1371/journal.pbio.0060159.
Hansen, E.C.A., Battaglia, D., Spiegler, A., Deco, G., Jirsa, V.K., 2015. Functional connectivity dynamics: modeling the switching behavior of the resting state. Neuroimage 105, 525–535. doi:10.1016/J.NEUROIMAGE.2014.11.001.
Harmelech, T., Malach, R., 2013. Neurocognitive biases and the patterns of spontaneous correlations in the human cortex. Trends Cogn. Sci. 17, 606–615. doi:10.1016/j.tics.2013.09.014.
Hellyer, P.J., Scott, G., Shanahan, M., Sharp, D.J., Leech, R., 2015. Cognitive flexibility through metastable neural dynamics is disrupted by damage to the structural connectome. J. Neurosci. 35, 9050–9063. doi:10.1523/JNEUROSCI.4648-14.2015.
Hellyer, P.J., Shanahan, M., Scott, G., Wise, R.J.S., Sharp, D.J., Leech, R., 2014. The control of global brain dynamics: opposing actions of frontoparietal control and default mode networks on attention. J. Neurosci. 34, 451–461.
Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks. Science 313, 504–507 (80).
Hutchison, R.M., Womelsdorf, T., Allen, E.A., Bandettini, P.A., Calhoun, V.D., Corbetta, M., Della Penna, S., Duyn, J.H., Glover, G.H., Gonzalez-Castillo, J., Handwerker, D.A., Keilholz, S., Kiviniemi, V., Leopold, D.A., de Pasquale, F., Sporns, O., Walter, M., Chang, C., 2013. Dynamic functional connectivity: promise, issues, and interpretations. Neuroimage 80, 360–378. doi:10.1016/J.NEUROIMAGE.2013.05.079.
Iverson, J.M., Thelen, E., 1999. Hand, mouth and brain: the dynamic emergence of speech and gesture. J. Conscious. Stud. 6, 19–40.
Jazayeri, M., Afraz, A., 2017. Navigating the neural space in search of the neural code. Neuron 93, 1003–1014. doi:10.1016/j.neuron.2017.02.019.
Jirsa, V., 2020. Structured flows on manifolds as guiding concepts in brain science. In: Selbstorganisation – Ein Paradigma für Die Humanwissenschaften. Springer, Fachmedien Wiesbaden, pp. 89–102. doi:10.1007/978-3-658-29906-4_6 Wiesbaden.
Kaufman, L., Rousseeuw, P.J., 2005. Finding Groups in Data: An Introduction To Cluster Analysis. Wiley Wiley Series in Probability and Statistics doi:10.1002/9780470316801.
Kieliba, P., Madugula, S., Filippini, N., Duff, E.P., Makin, T.R., 2019. Large-scale intrinsic connectivity is consistent across varying task demands. PLoS ONE 14, e0213861. doi:10.1371/journal.pone.0213861.
Kingma, D.P., Ba, J., 2014. Adam: a method for stochastic optimization. International Conference on Learning Representations.
Klein, A., Tourville, J., 2012. 101 labeled brain images and a consistent human cortical labeling protocol. Front. Neurosci. 6, 171. doi:10.3389/fnins.2012.00171.
Lemaître, G., Nogueira, F., Aridas, C.K., 2017. Imbalanced-learn: a Python toolbox to tackle the curse of imbalanced datasets in machine learning. J. Mach. Learn. Res. 18, 1–5.
Schröder, T.N., Haak, K.V, Zaragoza Jimenez, N.I., Beckmann, C.F., Doeller, C.F., 2015. Functional topography of the human entorhinal cortex. eLife 4. doi:10.7554/eLife.06738.
Nogueira, R., Lawrie, S., Moreno-Bote, R., 2018. Neuronal variability as a proxy for network state. Trends Neurosci. 41, 170–173. doi:10.1016/J.TINS.2018.02.003.
Oldham, S., Fornito, A., 2019. The development of brain network hubs. Dev. Cogn. Neurosci. 36, 100607. doi:10.1016/J.DCN.2018.12.005.
Onnela, J.P., Saramäki, J., Kertész, J., Kaski, K., 2005. Intensity and coherence of motifs in weighted complex networks. Phys. Rev. E. Stat. Nonlinear Soft Matter Phys. 71, 65103. doi:10.1103/PhysRevE.71.065103.
Oostenveld, R., Fries, P., Maris, E., Schoffelen, J.M., 2011. FieldTrip: open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. Comput. Intell. Neurosci. 2011, 1–9. doi:10.1155/2011/156869.
Pillai, A.S., Jirsa, V.K., 2017. Symmetry breaking in space-time hierarchies shapes brain dynamics and behavior. Neuron 94, 1010–1026. doi:10.1016/j.neuron.2017.05.013.
Ponce-Alvarez, A., Deco, G., Hagmann, P., Romani, G.L., Mantini, D., Corbetta, M., 2015. Resting-state temporal synchronization networks emerge from connectivity topology and heterogeneity. PLOS Comput. Biol. 11, e1004100. doi:10.1371/journal.pcbi.1004100.
Preti, M.G., Bolton, T.A., Van De Ville, D., 2017. The dynamic functional connectome: state-of-the-art and perspectives. Neuroimage 160, 41–54. doi:10.1016/J.NEUROIMAGE.2016.12.061.
Ronan, T., Qi, Z., Naegle, K.M., 2016. Avoiding common pitfalls when clustering biological data. Sci. Signal. 9, re6. doi:10.1126/scisignal.aad1932.
Rousseeuw, P.J., 1987. Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. J. Comput. Appl. Math. 20, 53–65. doi:10.1016/0377-0427(87)90125-7.
Rué-Queralt, J., Stevner, A., Tagliazucchi, E., Laufs, H., Kringelbach, M.L., Deco, G., Atasoy, S., 2021. Decoding brain states on the intrinsic manifold of human brain dynamics across wakefulness and sleep. Commun. Biol. 4, 854. doi:10.1038/s42003-021-02369-7.
Sadaghiani, S., Hesselmann, G., Friston, K.J., Kleinschmidt, A., 2010. The relation of ongoing brain activity, evoked neural responses, and cognition. Front. Syst. Neurosci. doi:10.3389/fnsys.2010.00020.
Sadaghiani, S., Kleinschmidt, A., 2013. Functional interactions between intrinsic brain activity and behavior. Neuroimage 80. doi:10.1016/j.neuroimage.2013.04.100.
Saenger, V.M., Ponce-Alvarez, A., Adhikari, M., Hagmann, P., Deco, G., Corbetta, M., 2018. Linking entropy at rest with the underlying structural connectivity in the healthy and lesioned brain. Cereb. Cortex 28, 2948. doi:10.1093/CERCOR/BHX176, (New York, NY).
Salimi-Khorshidi, G., Douaud, G., Beckmann, C.F., Glasser, M.F., Griffanti, L., Smith, S.M., 2014. Automatic denoising of functional MRI data: combining independent component analysis and hierarchical fusion of classifiers. Neuroimage 90, 449–468. doi:10.1016/j.neuroimage.2013.11.046.
Seeley, W.W., Menon, V., Schatzberg, A.F., Keller, J., Glover, G.H., Kenna, H., Reiss, A.L., Greicius, M.D., 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. J. Neurosci. 27, 2349–2356. doi:10.1523/JNEUROSCI.5587-06.2007.
Shine, J.M., Breakspear, M., Bell, P.T., Martens, K.A.E., Shine, R., Koyejo, O., Sporns, O., Poldrack, R.A., 2019. Human cognition involves the dynamic integration of neural activity and neuromodulatory systems. Nat. Neurosci. 22, 289–296. doi:10.1038/s41593-018-0312-0.
Smith, S.M., Beckmann, C.F., Andersson, J., Auerbach, E.J., Bijsterbosch, J., Douaud, G., Duff, E., Feinberg, D.A., Griffanti, L., Harms, M.P., Kelly, M., Laumann, T., Miller, K.L., Moeller, S., Petersen, S., Power, J., Salimi-Khorshidi, G., Snyder, A.Z., Vu, A.T., Woolrich, M.W., Xu, J., Yacoub, E., Uğurbil, K., Van Essen, D.C., Glasser, M.F., 2013. Resting-state fMRI in the Human Connectome Project. Neuroimage 80, 144–168. doi:10.1016/J.NEUROIMAGE.2013.05.039.
Sporns, O., 2013. Structure and function of complex brain networks. Dialogues Clin. Neurosci. 15, 247–262.
van den Heuvel, M.P., Sporns, O., 2013. Network hubs in the human brain. Trends Cogn. Sci. 17, 683–696. doi:10.1016/J.TICS.2013.09.012.
van den Heuvel, M.P., Sporns, O., 2011. Rich-club organization of the human connectome. J. Neurosci. 31, 15775–15786. doi:10.1523/JNEUROSCI.3539-11.2011.
Varela, C., 2014. Thalamic neuromodulation and its implications for executive networks. Front. Neural Circuits 8, 69. doi:10.3389/fncir.2014.00069.
Varela, C., Kumar, S., Yang, J.Y., Wilson, M.A., 2014. Anatomical substrates for direct interactions between hippocampus, medial prefrontal cortex, and the thalamic nucleus reuniens. Brain Struct. Funct. 219, 911. doi:10.1007/S00429-013-0543-5.
Vidaurre, D., Hunt, L.T., Quinn, A.J., Hunt, B.A.E., Brookes, M.J., Nobre, A.C., Woolrich, M.W., 2018. Spontaneous cortical activity transiently organises into frequency specific phase-coupling networks. Nat. Commun. 9, 1–13. doi:10.1038/s41467-018-05316-z, 2018 91.
Vila-Vidal, M., Capouskova, K., Atasoy, S., Kringelbach, M.L., Deco, G., 2019. Uncovering the spatiotemporal scales of common neuro-mental constructs. Phys. Life Rev. doi:10.1016/j.plrev.2019.10.004.
Vila-Vidal, M., Pérez Enríquez, C., Principe, A., Rocamora, R., Deco, G., Tauste Campo, A., 2020. Low entropy map of brain oscillatory activity identifies spatially localized events: a new method for automated epilepsy focus prediction. Neuroimage 208, 116410. doi:10.1016/J.NEUROIMAGE.2019.116410.
Williams, A.H., Kim, T.H., Wang, F., Vyas, S., Ryu, S.I., Shenoy, K.V., Schnitzer, M., Kolda, T.G., Ganguli, S., 2018. Unsupervised discovery of demixed, low-dimensional neural dynamics across multiple timescales through tensor component analysis. Neuron 98, 1099–1115.e8. doi:10.1016/j.neuron.2018.05.015.
Wilson, M., 2002. Six views of embodied cognition. Psychon. Bull. Rev. 9, 625–636. doi:10.3758/BF03196322.
Yang, G.R., Cole, M.W., Rajan, K., 2019. How to study the neural mechanisms of multiple tasks. Curr. Opin. Behav. Sci. 29, 134–143. doi:10.1016/J.COBEHA.2019.07.001.
Zamora-López, G., Chen, Y., Deco, G., Kringelbach, M.L., Zhou, C., 2016. Functional complexity emerging from anatomical constraints in the brain: the significance of network modularity and rich-clubs. Sci. Rep. 6, 38424. doi:10.1038/srep38424.
Zamora-López, G., Zhou, C., Kurths, J., 2011. Exploring brain function from anatomical connectivity. Front. Neurosci. 5, 83. doi:10.3389/fnins.2011.00083.