Graph models of brain state in deep anaesthesia reveal sink state dynamics of reduced spatiotemporal complexity and integration

Deep anaesthetisation is an important surgical and explorative tool in the study of consciousness. Much work has been done to connect the deeply anaesthetised condition with decreased functional complexity. However, anaesthesia-induced unconsciousness is also a dynamic condition in which functional activity and complexity may fluctuate, being perturbed by internal or external (e.g. noxious) stimuli. We use fMRI data from a cohort undergoing ultra-slow propofol induction and a dynamic graph modelling framework which characterises changes in functional activity and connectivity as changes in brain state. We examine the group-level dynamics of brain state activity, complexity, integration and modularization in deep anaesthesia and wakefulness and propose novel measures for temporal complexity of brain state dynamics. We find that deep anaesthesia states are less temporally complex and less functionally-related to one another than wakeful states. Anaesthesia dynamics are dominated by a handful of sink states that act as low-complexity attractors to which subjects repeatedly return. Our analysis suggests that dynamic functional organisation in anaesthesia can be characterised by temporally stable, modular communities that change little in time, stratifying regional activity and functionally isolating cortical from both subcortical, and in particular, thalamic brain areas. Using data from a large human structural tractography atlas, we show that functional modularisation in deep anaesthesia more closely resembles normative structural organisation than does functional organisation in wakefulness. We find evidence suggesting that brain state trajectories and in particular, the dominance of low complexity attractor-like sink states in anaesthesia, appear to depend on subject-specific age and anaesthesia susceptibility factors.

James Wilsenach[1], Charlotte M. Deane[2], Gesine Reinert[2], Katie Warnaby[3]

General anaesthesia is an invaluable surgical tool that reduces the risk of complications and shields patients from immediate pain while being generally safe and effective [1]. However, anaesthetic affects are patient specific and incorrect monitoring of anaesthesia risks delayed recovery of consciousness, post-operative delirium and psychological trauma through post-surgical recollection of pain, helplessness due to paralysis or fear. In the long-term, over anaesthetisation has also been associated with cognitive decline in older patients who are also at higher risk of post-operative delirium [2][3][4][5] These factors motivate the importance of understanding the different possible trajectories, both into and out of unconsciousness, where anaesthesia has been shown to be immensely useful as a tool for controlled exploration of the conscious and unconscious state spaces [6].

Other important avenues of studying unconscious brain state change include the study of deep sleep alongside the so-called Disorders of Consciousness (DOCs). Here, Hidden Markov Models (HMMs) are a class of models which have been useful in this exploration with applications to sleep (using fMRI) and to DOCs (using EEG data) [7][8]. In HMMs neural activity are modelled as observations from a sequence of underlying, hidden brain states, hidden in the sense of being observed only indirectly through electrophysiological or haemodynamic recordings. The future trajectory of hidden states, depends only on the current state (see Supplementary Section S.1 for a brief introduction). HMMs and other state-space models have previously been used to demonstrate a decrease in temporal complexity across unconscious and minimally conscious conditions (such as sleep) using a variety of models and measures [7][9] This decrease has been characterised, in part, as a lowered dynamic

[1]Alan Turing Institute, British Library, 96 Euston Road, London NW1 2DB

[2]Department of Statistics, University of Oxford, 24-29 St Giles', Oxford OX1 3LB

[3]Oxford Centre for Functional MRI of the Brain, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU

switching rate between states relative to wakefulness. We provide a new characterisation in terms of the information (entropy) rate of the underlying state switching dynamics and relate it to established measures of temporal complexity.

Altered thalamocortical connectivity has previously been observed in simultaneous fMRI/EEG recordings. Thalamic disconnection has been hypothesised to be an important part of the transition to anaesthetic induced unconsciousness [10] and is strongly associated with disruption of the resting state default mode, sensory and salience networks [11]. Dose dependent cortical changes in response to pain and auditory stimuli are also observed in deep anaesthesia with the posterior prefrontal cortex retaining some level of activity, at moderate to low doses, in conjunction with the parietal lobe [12]. These areas are commonly thought to be involved in self-perception and awareness and may remain partially active forming a possible conserved network in deep anaesthesia. These effects are mediated by factors such as age which can be a significant determiner of anaesthesia susceptibility, brain state complexity and post-operative delirium [13][14][15]. We show links between brain state activity, age and time at loss of responsiveness, which can be used as a proxy for anaesthetic susceptibility [16].

General decreases across several measures have been observed in anaesthesia including signal-based entropy and complexity measures (such as Lempel-Ziv complexity) [17][18][19] Signal complexity and in particular Kolmogrov complexity, is of particular interest as it is thought to be a possible neural correlate of consciousness and is notable for its relationship to other proposed neural correlates [20][21][22][19], though its accurate quantification in practice, in particular in recordings of brain activity, remains difficult [23]. In the special case where the dynamics can be modelled as a Markov process the information required to encode the observed state, termed the information or entropy rate, is exactly proportional to the Kolmogorov complexity [24]. Along with related measures that are just now being proposed [25] we believe this is the first time that the information rate has been put forward as a viable measure of dynamic complexity in the study of consciousness.

In order to understand the spatiotemporal modularisation of activity in both wakefulness and anaesthesia, we employ the Hidden Markov Graph Model (HMGM) framework put forward in [26]. From the HMGM we develop a new way of studying the relationship between regional and temporal brain state activity. Previous studies demonstrate large scale changes in the brain state dynamics under anaesthesia, notably reduced functional excitability and connectivity in brain states when compared to healthy resting wakefulness [12] We corroborate these findings and show here that global reductions in dynamic complexity may be due to both global reductions in functional integration across brain states and the emergence of a small set of low complexity states that act as sinks dominating unconscious dynamics.

Previous work has shown that certain features of the structural connectome, notably measures derived from white matter tractography, resemble connectivity features in both deep anaesthesia induced unconsciousness and DOCs [9]. However, functional connectivity is known to change dynamically with brain state across conscious conditions [27][28][29][30]. We expand upon previous community detection methods [31][26] and apply them to the HMGMs for both wakefulness and anaesthesia, thereby inferring the dynamic functional module organisation of these conditions. We use these new methods to show that spatiotemporal functional stratification occurs in anaesthesia-induced unconsciousness [32]. We also compare these dynamic functional modules to structural

modules derived from a large proxy structural graph model based on a large cohort of adult tractography data, and show that anaesthesia functional organisation more closely aligns with the modular organisation of the structural connectome.

Our results suggest new avenues for exploration in the study of anaesthesia, in particular how age and anaesthesia susceptibility may modulate patient trajectories both into and out of anaesthesia-induced unconsciousness. Moreover, we demonstrate that unconsciousness dynamics may involve both dissolution of important functional modules and loss of the brain's capacity for the dynamic processing of multiple streams of information.

1 Materials and Methods

Fig. 1 shows the experimental, pre-processing and modelling frameworks we apply in fitting the HMGM (a dynamic model of brain state activity) and applying the robust consensus community detection method [33] These steps are central to our brain state modeling and analysis and which has been further developed here and in [26]. Data acquisition, preprocessing and experimental design broadly follow the methods as in [12], while new information rate and state comparison methods are detailed below. Methods to determine statistical dependencies between graphs use the so-called distance correlation (see Section 1.7 for details), including between the states themselves and when determining structure-function relationships. In all other cases (e.g. when computing differences in temporal dynamics), the Pearson correlation coefficient is used.

1.1 Data Acquisition

The experimental paradigm is shown in Fig. 1A; fMRI data was acquired over the course of a four phase experiment in which 16 initial subjects received a phase-specific dose of the GABA-ergic general anaesthetic propofol *[12]*. In the first phase, baseline wakeful activity was recorded for 10 minutes. Subjects were asked to remain still with eyes closed. In the second phase, propofol was gradually infused to a maximum Effective Concentration (ESC) of $4\mu g/ml$, which is well above the minimal dose previously found to be effective in inducing loss of consciousness in subjects and is sufficient for inducing sleep-like slow waves [34]. In phase three the maximum ESC was maintained for 10 minutes leading to a stable deep anaesthesia state. After this period, the propofol dose was gradually lessened until subjects regained full responsiveness.

We focus on the 10 minutes pre-induction (phase 1) and the 10 minutes post-induction recordings (phase 3), when subjects are held at the maximum dosage. The high dosage results in a condition of deep anaesthesia-induced unconsciousness and non-responsiveness to conventional rousing stimuli. fMRI data was acquired at a frequency of one 3D full volume scan of the head every three seconds at a 3D voxel resolution of $2mm^3$. All brain volumes were artefact corrected and registered to a single $1mm^3$, subject-specific high-resolution volume-scan in FSL [35]. The resulting volumes were then mapped into the MNI152 MRI standard space [36], see [12] for further details.

1.2 Subject Exclusion Criteria

This study focuses on comparison with propofol-induced deep anaesthetic states while attempting to exclude subjects whose trajectory is dominated by atypical or potentially harmful activity. Burst-suppression is a state of deep anaesthesia that has been particularly associated with poor post-

operative outcomes including post-operative delirium and cognitive decline [37]. It is characterised by a global pattern of alternating low and high amplitude periods of neural activity (see Supplementary Figure 1 for EEG activity example) [38]. We thus excluded subjects on the basis of pervasive burst-suppression (>2 minutes) during simultaneous EEG recordings. Recordings of subjects exhibiting brief periods of burst suppression (less than a minute across the recording) were not altered or removed. One subject was removed due to an incomplete anaesthesia recording, two subjects were removed on the basis of extensive burst suppression. This resulted in a reduced set of $N=13$ deeply anaesthetised and awake subject recordings (see Supplementary Table 1) which were used in the construction of new wakefulness and anaesthesia HMGMs. Th age of the final subjects ranged between 19 and 34 with a median of 25.

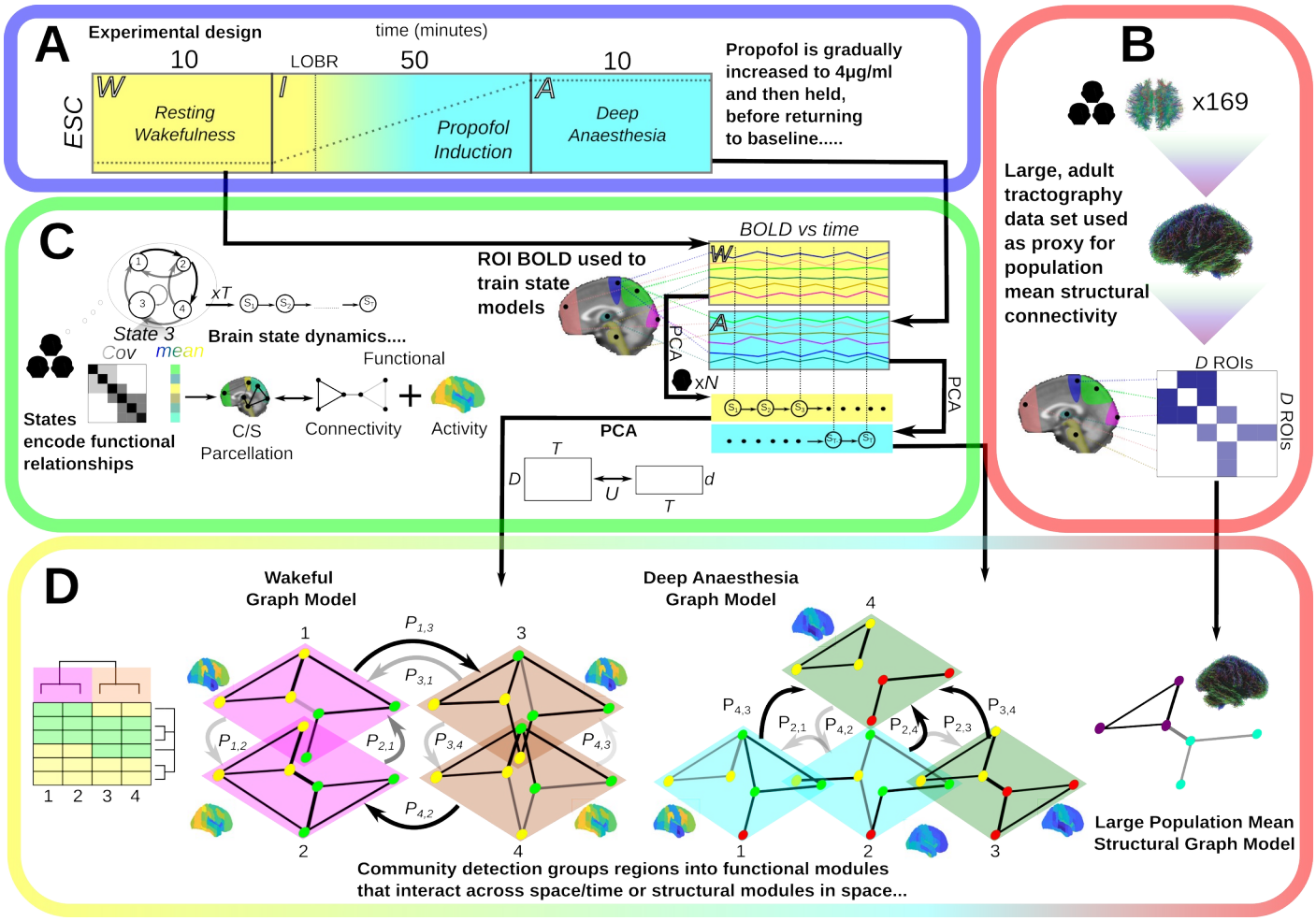


*Figure 1: This figure shows our experimental and analysis framework. (**A**) Subject fMRI is recorded at resting wakefulness, yellow (W) followed by ultra-slow propofol induction (I), where the estimated Effect Site Concentration (ESC) is gradually increased to 4μg/ml, and then held during the deep anaesthesia, cyan, phase (subjects are then allowed to recover). Loss of Behavioural Responsiveness (LOBR) during induction is marked with a dotted line. (**B**) An independent tractography dataset from a large cohort of adult subjects was used to derive a structural connectome model (here showing a simple hypothetical parcellation of six ROIs). (**C**) In our model framework, brain dynamics arise from a sequence of hidden brain states. Occurrence of the next state is probabilistically dependent on the current state in the sequence. These probabilities are shown with an arrow (darker, higher probability). A simple four state model is sampled repeatedly to generate a sequence of length T. Each state encodes expected functional connectivity and activity relationships that are mapped into HO-parcellation space through the state mean (blue, low) and covariance (dark, high) parameters. Models are trained using both wakeful and anaesthesia fMRI recordings for all N subjects. ROI BOLD time series, across all subjects were first dimensionality reduced by Principal Component Analysis (PCA) separately across each condition through the eigenmatrix (U). (**D**) We combine functional and dynamic information into a multiplex graph model (dark arrows, high probability transition), with each region belonging to a dynamically changing functional community that can be used to hierarchically cluster functionally related states (shaded trees) and regions. Similarly, a population mean structural graph-community model is created from the independent adult tractography dataset.*

1.3 Structural Graph Model

Fig. 1B shows how structural graph model construction was carried. This construction utilised a parcellation derived from the combined Harvard-Oxford cortical and subcortical (HO) atlas developed in [39] [26] This was used due to its broad coverage and relatively small number of $D=63$ ROIs to fit. We accessed a large structural dataset combining data from 169 normative adult tractographies based on multi-shell Diffusion Tensor Imaging re-registered to MNI152 standard space [40]. The tractographies were constructed using the Gibbs tracking algorithm [41], which were then co-registered in standard space as outlined in [42], and all tracts were added to a single large group-level tractography file in the TrackVis file format [43]. A single-layer structural graph model was then fit using DSI-studio's structural connectome tool [44] and the HO atlas with default parameters and median streamline length normalisation, producing an undirected, weighted graph model proxy of subject structural connectivity [45]. Robust consensus modularity community detection was applied to determine structural modules (see Supplementary Section S.2) [31].

We used the generalised Louvain algorithm to optimise the modularity score across all possible partitions of the graph models which is a stochastic method [46]. We utilise consensus community detection which involves iteratively re-partitioning the data into communities and aggregating these to obtain a single partition that is representative of all likely partitions of brain regions across a range of resolutions. This method has been shown to improve robustness of the communities to noise in the connectome [31].

1.4 Preprocessing, Model Selection and Training

Model selection and training for both anaesthesia and wakefulness is carried out as detailed in the methodology paper [26] using the HMM model training procedure defined in [47] The Hidden Markov interpretation of brain dynamics and model training procedure are summarised in Fig. 1C. The number of initial states are selected by maximum entropy after dimensionality reduction using Parallel Analysis for principal component selection as in [26]. States were removed if they appeared in less than 20% of subjects and transition probabilities were re-standardised so that each row sums to one. The probability of each state occurrence at each time point is then calculated, which determines the Fractional Occupancy (FO) of each state. The FO represents the expected fraction of time spent in each state over the length of the recording. Lastly, mean and covariance relationships were calculated by backprojecting into the original D-dimensional HO parcellation space, resulting in a model for each condition with a total of $|S|=K$ states, D brain regions, and mean and covariance for each state.

1.5 Graph Model and Measurement of Functional Connectivity

We use the the absolute Pearson correlation as a measure of functional connectivity [48]. This was chosen to simplify the interpretation and allow for comparison with established results which use this measure. More formally, we measure the functional connectivity as the absolute value of the Pearson correlation coefficient between regions calculated from the covariance matrix of each state [49].

1.6 Multiplex Graph Model and Spatiotemporal Consensus Community Detection

The HMGM models of both wakefulness and deep anaesthesia is a dynamic graph model. More formally, the model is a multiplex graph in which the interlayer edges are the probabilities of transition between states and the intralayer edges are determined by the functional connectivity

between regions (nodes). This structure can be seen in Fig. 1D; each layer is given a colour according to its relatedness to the community composition of other states. Communities partition the nodes (regions) into functionally related sets that coordinate activity.

Here, we determine community membership spatiotemporally so that regions that tend to coordinate across space and time are placed in the same communities (denoted by node colour). We measure the goodness of fit of a given community partition using the multiplex modularity score that takes into account both functional connectivity and state transition relationships between states [50][46]. Our method modifies the traditional modularity score (see Supplementary Section S.3) by using the transition probability between states to inform the level of coupling between regions in different states. In this way communities can represent likely functional module reconfigurations that are required to perform a coordinated sequence of tasks or effect complex behaviours. The modularity is optimised using the generalised Louvain algorithm [46].

1.7 Sink States, Switching and Information Rates

Markov models are probabilistic state space models in which state transitions occur probabilistically depending only on the previous state in the sequence. In a sufficiently regular Markov model, the stationary probability of a state is the long term probability of occupying any one state over a sufficiently long time (see Supplementary Section S.3). Here, we use the stationary probability to define a centrality measure over states; a sink state is a state with high long-run occupancy probability. This quantity is closely related to the PageRank centrality on directed graphs and we term the stationary probability of each state the sink centrality in this context [51]. Here, we define a strong sink state to be one that has a sink centrality of at least 0.05.

Fig. 2 provides a sketch of how HMGM analysis can provide important axes along which to examine brain state activity by contrasting model behaviours. The switching rate is a well established measure of the temporal complexity of brain state models. It is the frequency at which state change occurs in the (a postereori) most probable state sequence that explains the data; such a sequence was calculated using the so-called Viterbi algorithm [52]. Fig. 2A shows four different four-state Markov models where node size is proportional to the sink centrality. A probable state sequence is generated from each model from which the state switching events are determined and the switching rate (in Hz) is calculated. The FO is also shown, which is based on the expected time spent in each state. The FO can be viewed as the distribution describing the probability of observing a subject in a given state during recording (see Supplementary Section S.1).

The switching rate is an important measure but does not take into account the diversity of states visited. To factor in this diversity, the information rate measures the expected information gained by sampling from a Markov model over a sufficiently long state sequence [53]. It is measures the expected surprise and thus the overall realisability of all states in the model, providing an estimate of the temporal information capacity or complexity of the underlying process (see Supplementary Section S.4). A model can have high switching rate while having relatively low complexity or be high in both measures. On a subject-specific level, brain state diversity can also be measured directly using the entropy of the distribution defined by the proportion of time subjects are likely to spend in any one state.

1.7 Comparing States by Functional and Community Activity

HMM state functional connectivity and activity provide important modes to compare conditions of consciousness; brain activity may range from excitable (high absolute mean functional activity) to quiescent. We use paired distance correlation between states to compare states across both functional mean activity as well as functional connectivity. The states themselves encode complex connectivity relationships; these high dimensional relationships are often non-linear and so our comparison benefits from using a measure that takes into account such non-linear dependencies [54]. Fig. 2B shows a range from dark blue (quiescent) to yellow (excitable) brain states. In addition, we apply the so-called, signed distance correlation (using standard Pearson correlation) to provide additional information on state dependencies [55].

The level of information integration within each state was assessed using the average local clustering coefficient of the state functional connectivity. Lastly, state modularity is measured by the modularity score [56], which measures the capacity of the state to independently process multiple complex streams of information (see Supplementary Section S.4). Modularity and integration are both necessary conditions for the brain to perform complex tasks which require integration of information from multiple sources.

1.8 Clustering Communities to Characterise Brain Dynamics

Fig. 2C shows the results of multiplex community analysis for a model. Expected community spatial activity is projected onto the brain using the stationary distribution to provide the fractional membership of each region. We perform hierarchical clustering using Ward's algorithm to group related states and brain regions across time and space in order to characterise these dynamic communities [57], as well as to characterise the FO across subjects and regions (see Supplementary Section S.5).

1.9 Comparing Functional to Structural Connectivity

In order to test whether deep anaesthesia or wakeful state functional connectivity is more related to structural connectivity, we compared the degree and level of integration of each state across conditions to the structural connectome using distance correlation. We also used the Adjusted Mutual Information (AMI) to compare the community composition of structural and dynamic graph models. The AMI is a measure of agreement between two partitions of brain regions that is more robust to differences in partition size than other measures such as the Adjusted Rand Index (see Supplementary Section S.6) [58].

2 Results

In this section we observe key differences in static activity between wakefulness and anaesthesia conditions. We then train and analyse two dynamic models using data from each condition.

Prior to pre-processing, we characterised static activity by comparing functional activity and connectivity across subjects in anaesthesia and wakefulness. On a subject level there is high Pearson correlation between mean functional activity in wakeful and anaesthetised conditions ($r = 0.8760$,

$p<0.001$, t-test, two-sided) as measured by subject mean BOLD signal. This is contrasted against a much lower, but still significant, level of correlation at the level of functional connectivity ($r=0.3908$, $p<0.001$, t-test, two-sided) as measured from region-wise absolute Pearson cross-correlation in BOLD signal.

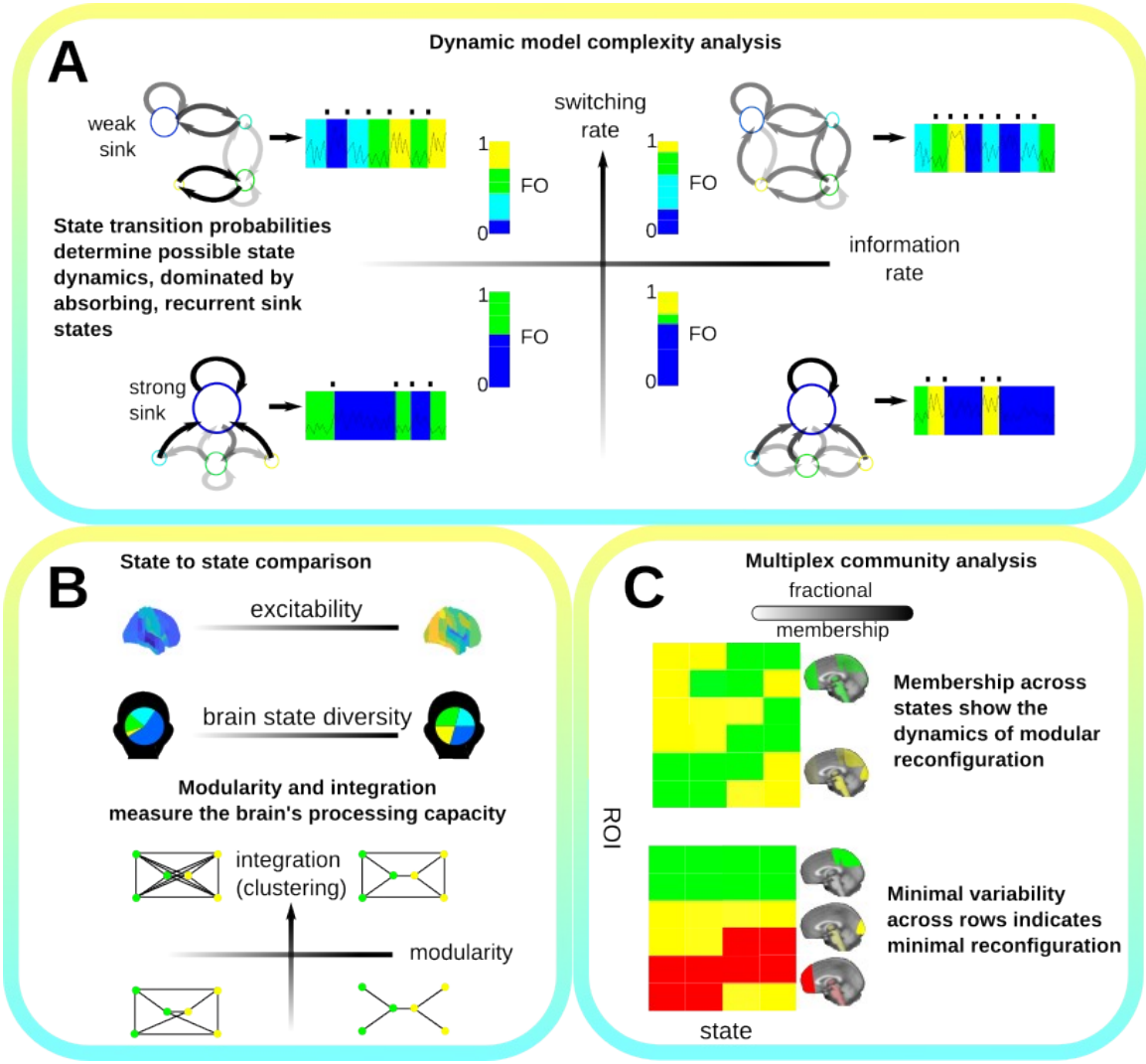


Figure 2: This figure shows our analysis framework (A) Here, we use information theoretic comparisons of state space complexity and depict four models, one with switching rate increasing from top to bottom. Larger nodes are more central states that occur more often and are called sinks. For each recording we show the probability of being in any one state, the Fractional Occupancy (FO), as a bar plot with length of each coloured bar corresponding to the FO. (B) Mean functional activity is an important characteristic of brain states which may vary from low to high excitability. Similarly, the diversity of states visited over is another important measure of temporal complexity. The community and connectivity structure also informs the complexity of state information processing along multiple axes including modularisation (capacity for independent processing) and integration. (C) Multiplex community analysis allows for states to be partitioned into functionally related modules that share members across temporally related states. These communities may be characterised spatially by their fractional membership, the expected proportion of time a region spends in each community, here shown by colour opacity from light (low membership) to dark (high membership).

2.1 Selection of Models

We select the number of states for both wakeful and deep anaesthesia models based on the maximum entropy of state occurrence (maximum entropy). This results in two different models that maximise the entropy with 33 states for wakefulness and 19 for anaesthesia respectively. After exclusion of sporadic states that featured in less than 25% of subject trajectories, the resulting models had

$K_{wake}=32$ and $K_{anaes}=18$ states. For validation purposes, we also trained one combined model using both anaesthesia and wakeful data concatenated to produce a single model with the same number of initial states as wakefulness $K_{comb}=33$ (after applying the same state exclusion criterion).

2.2 Deep Anaesthesia Brain State Transition Graph

Fig. 3 shows the transition probabilities between states as the edges in a directed graph of states (nodes) for both anaesthesia and wakefulness. States are organised by the integration of their functional connectivity (as measured by local clustering coefficient) and the size of states is proportional to their sink centrality. The two graph models show clear separation along the state integration axis, with the least integrated states being anaesthesia states ($p<0.001$, Mann-Whitney U Test, two-sided). The two least integrated states, anaesthesia states A-12 and A-17, are also the strongest sinks. In Fig. 3 we see that these sink states have many high probability in-edges. In particular, A-17 is reachable with relatively high probability from states within its own cluster, and is unlikely to be transitioned to from external states, except A-3 and A-9 which both act as bridges from the other cluster. Similarly, transitions to A-12 the second strongest sink state generally occur from within its own cluster but there are a few states in the A-17 dominated cluster that preferentially transition to A-12 acting as bridges between clusters. State A-5 has no strong connection to either cluster.

Wakeful dynamics are more diverse with no dominant sink state and a number of clusters. Mean activity maps for sink states also show larger ranges (higher excitability) than in anaesthesia across all states ($p=0.017<0.05$, Mann-Whitney U Test, two sided). Overall, the simpler modular structure of the anaesthesia model suggests lower complexity dynamics and indeed the anaesthesia information rate of 0.940 is less than the wakeful information rate of 1.82. Supplementary Figure 2 shows that when varying the number of states in the model; a difference between the two models persists, indicating that the difference in complexity is not simply due to a difference in the number of states.

2.2 Sink State Dominance and Anaesthesia Quiescence

Fig. 4A and Fig. 4B show the stationary distribution across all subjects as pie charts with wedges representing sink states (with sink centrality 0.05 or more). Again, we see strong sink state dominance in anaesthesia and a more uniform stationary distribution in wakefulness. Fig. 4C and 4D show the top 50% most extreme regional activity for each sink state in wakefulness and anaesthesia respectively. All regions shown for A-12 and A-17 have below mean regional BOLD activity. In general, wakeful states have higher absolute mean activity across regions than in anaesthesia ($p<0.001$, Mann-Whitney U Test, two-sided). This also holds for the combined trial validation model (Supplementary Figure 3) in which anaesthesia trials were dominated by a single state (state 29) which had lower than baseline mean absolute functional activity.

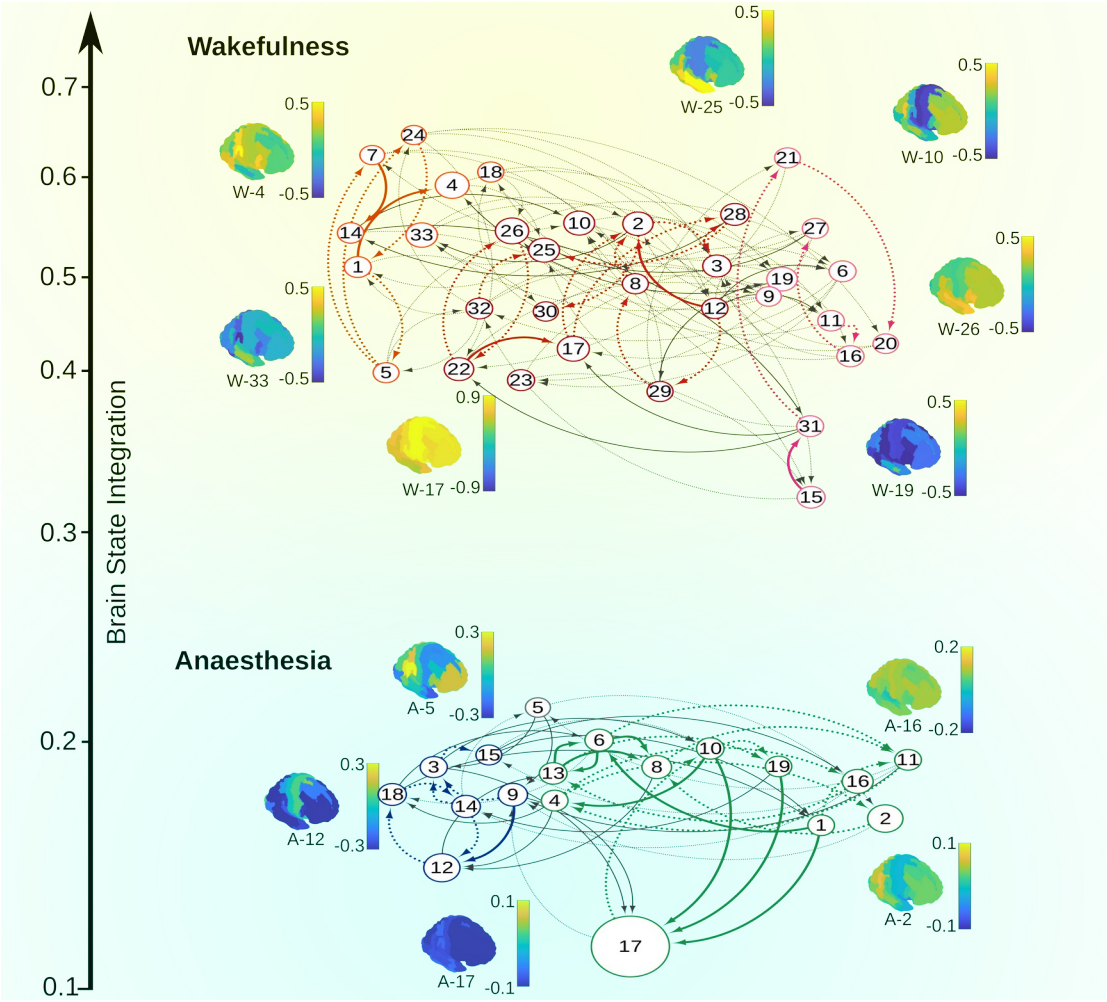


Figure 3: In the HMGM states are related by the probability of transition between them. This figure shows two graphs representing the respective state dynamics of subjects when awake or anaesthetised. Each node is a state the size of which is proportional to the state sink centrality. Edges represent that there is a non-negligible Transition Probability (TP) from one state to another. For visualisation purposes, only the two strongest edges from or to a node are shown with weaker probability edges (TP<0.05) denoted by dotted lines. The level of brain state integration (clustering coefficient) is used to order the states on the γ-axis on a $\log_{10}$-scale. Background shading indicates the change from high integration (yellow, bottom) wakeful states to significantly less integrated (cyan, blue) anaesthesia states (p<0.001, Mann-Whitney U Test). States are organised horizontally into communities with coloured edges represent connections between states that share similar functional connectivity and dynamics. Here, we show brain maps for all states that account for 0.05 or more of the stationary distribution. We show brain maps of mean functional activity for all sink states (and associated colourbar), as well as state A-5 which is a transient outlier which does not belong to any cluster.

2.3 Subject Specific Dynamics in Wakefulness and Anaesthesia

While there is some overlap in FO between subjects, Fig. 5A shows that the two sink states separate subjects into two groups depending on which sink state dominates. The majority of subjects are dominated by A-17, the remainder have a high A-12 occupancy. The FO of the two states (A-12 and A-17) have a strong negative Pearson correlation of $\rho = -0.922$ ($p < 0.001$, permutation test). Together A-12 and A-17 account for $> 50\,\%$ of the mean FO across all subjects in deep anaesthesia, indicating that a majority of subject time is spent in one of these two states. This is in comparison to wakefulness in which no two states can account for more than 20% of mean FO. While it is not clear why subjects 10 and 4 are primarily dominated by other states, we note that the two subjects are the youngest in the study. In the validation models using either the combined data (including both wakeful and anaesthesia trials in a single model) or an alternative model with $K = 20$ states, anaesthesia trials were also dominated by sink states (states 29 and 20 respectively, see Supplementary Figures 3 and 4).

Fig. 5B shows a statistical breakdown of the state switching dynamics across all subjects with higher switching in wakefulness ($p < 0.001$, Wilcoxon Signed-rank test, two sided) and a higher state diversity, as measured by the normalised entropy of FO ($p < 0.001$, Wilcoxon Signed-rank test, two sided). Together these results suggest that dynamics under deep anaesthesia are not only simpler but also less dynamic, with a lower fraction of states participating in the dynamics of most subjects, as would be expected for dynamics with a lower information rate. However, there is also some evidence that subjects are more variable, likely due to the subject-specific anaesthesia depth which varies even when subjects are held at the same ESC [59].

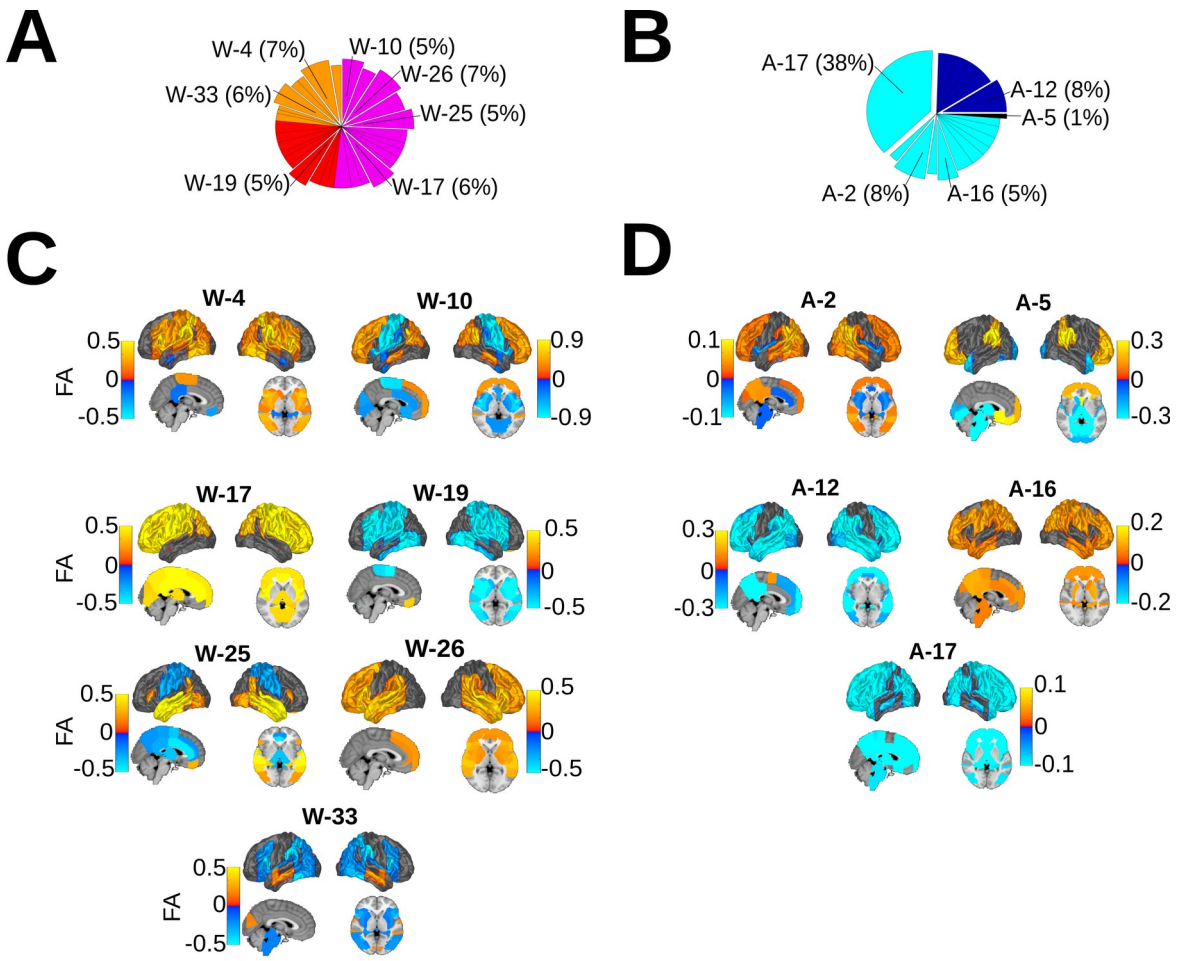


*Figure 4: This figure summarizes the state activity of important states, highlighting the sink states as well as the outlying state A-5 and the prominent areas of surface Functional Activity (FA) for each. The stationary probability is shown for **(A)** wakefulness and **(B)** anaesthesia, with wedges indicating the sink states. In **(C)**, the top 50% mean absolute functional activity in wakefulness is shown as a surface plot with a variety of high and low activity regions. **(D)** shows activity in the anaesthesia sink states, with much lower absolute mean functional activity across these states, indicative of lower absolute mean activity in anaesthesia as a whole (p<0.001, Mann-Whitney U test, two-sided).*

2.4 Stratification of Functional Connectivity in Anaesthesia

Spatiotemporal consensus clustering combining functional and temporal relationships was performed (see Section 1.6). In the wakeful case, the resulting model shows highly distributed dynamics with each community including diverse and dynamically changing regional membership (see Fig. 6A). We characterise these communities as recruiting from regions including visual, salience-default mode and sensory networks respectively. This is most apparent in in Fig. 6B where regional switching between communities is evident. The anaesthesia model (see Fig. 6C and D) reveals strikingly different results, with a large number of less dynamic communities that are primarily restricted to neighbouring regions in the same area. There is far less change of membership between states in anaesthesia in comparison to wakefulness resulting in stratified and static dynamics. Even though we do see less switching in anaesthesia dynamics, sporadic co-recruitment of regions commonly associated with the so-called Resting State Networks does occur for example in A-16 (exhibiting Default Mode-like activity in community 3).

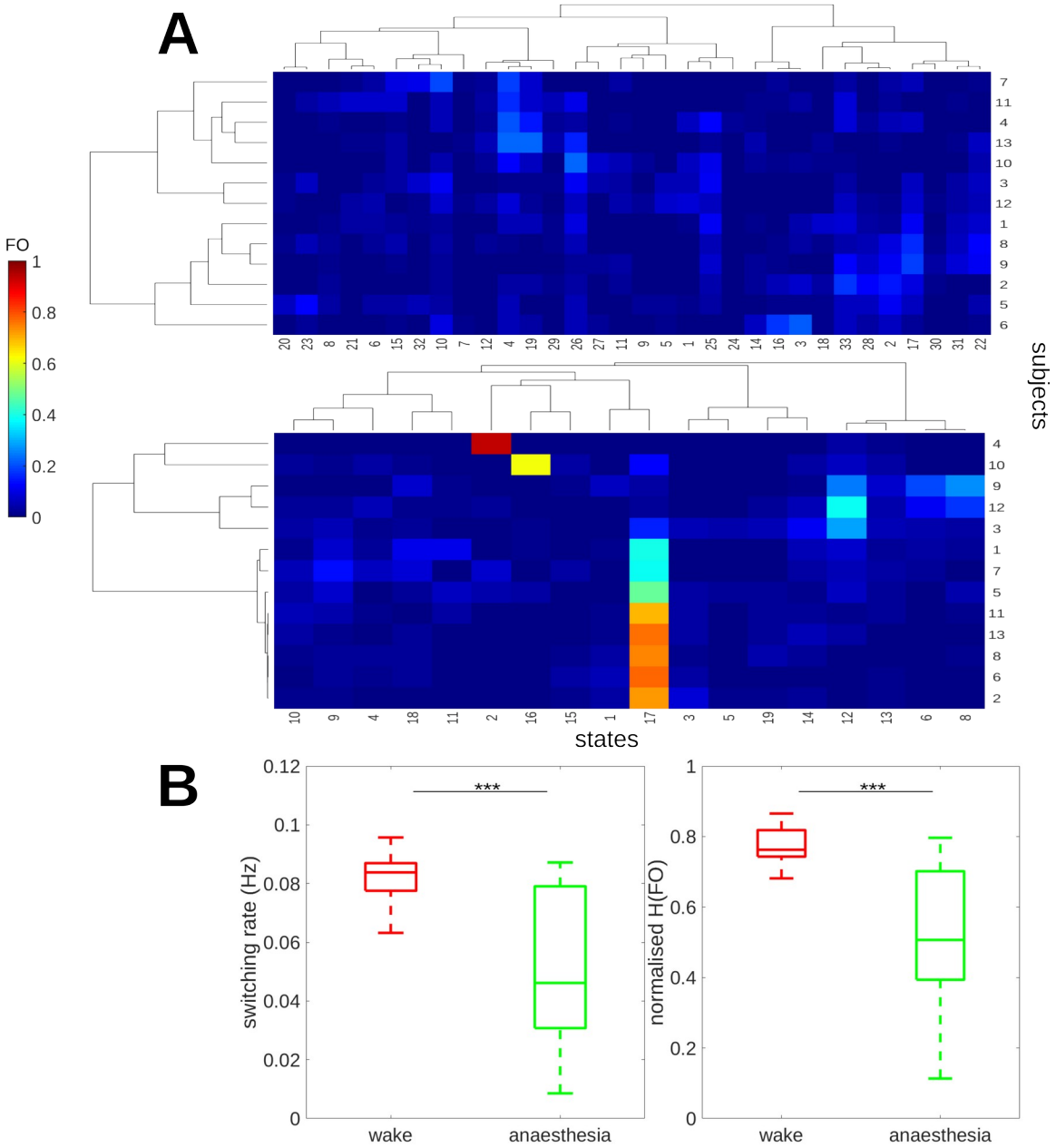


Figure 5: Subject specific state dynamics is shown for both anaesthesia and wakefulness. ***(A)*** *shows the FO heat map as a clustergram where subjects (rows) and states (columns) have been hierarchically clustered by Pearson correlation using Ward's algorithm. In anaesthesia, dynamics tend to be dominated by a subset of sink states, including subjects 4 and 10 which show a strong preference for A-2 and A-16 respectively.* ***(B)*** *Shows a boxplot comparing wakeful and and anaesthetised switching rates and normslised occupancy entropy (H(FO)). The state switching rate (Hz) is more varied but significantly lower in anaesthesia (p<0.001, Wilcoxon Signed-rank test, two-sided) The H(FO) shows a similar pattern to the switching rate with highly variable but overall lower state diversity present in anaesthetised versus awake subjects (p<0.001, Wilcoxon Signed-rank test).*

Furthermore, the consensus clustering in anaesthesia suggests that subcortical areas, most notably the thalamus and neighbouring regions (including the caudate), generally thought of as an information bottleneck for sensory inputs reaching the brain [60], are isolated from cortical brain regions in almost all states in anaesthesia. This is not seen in wakefulness in which thalamic regions share and switch membership with multiple cortical regions depending on state, suggesting dynamic involvement of the thalamus in wakefulness.

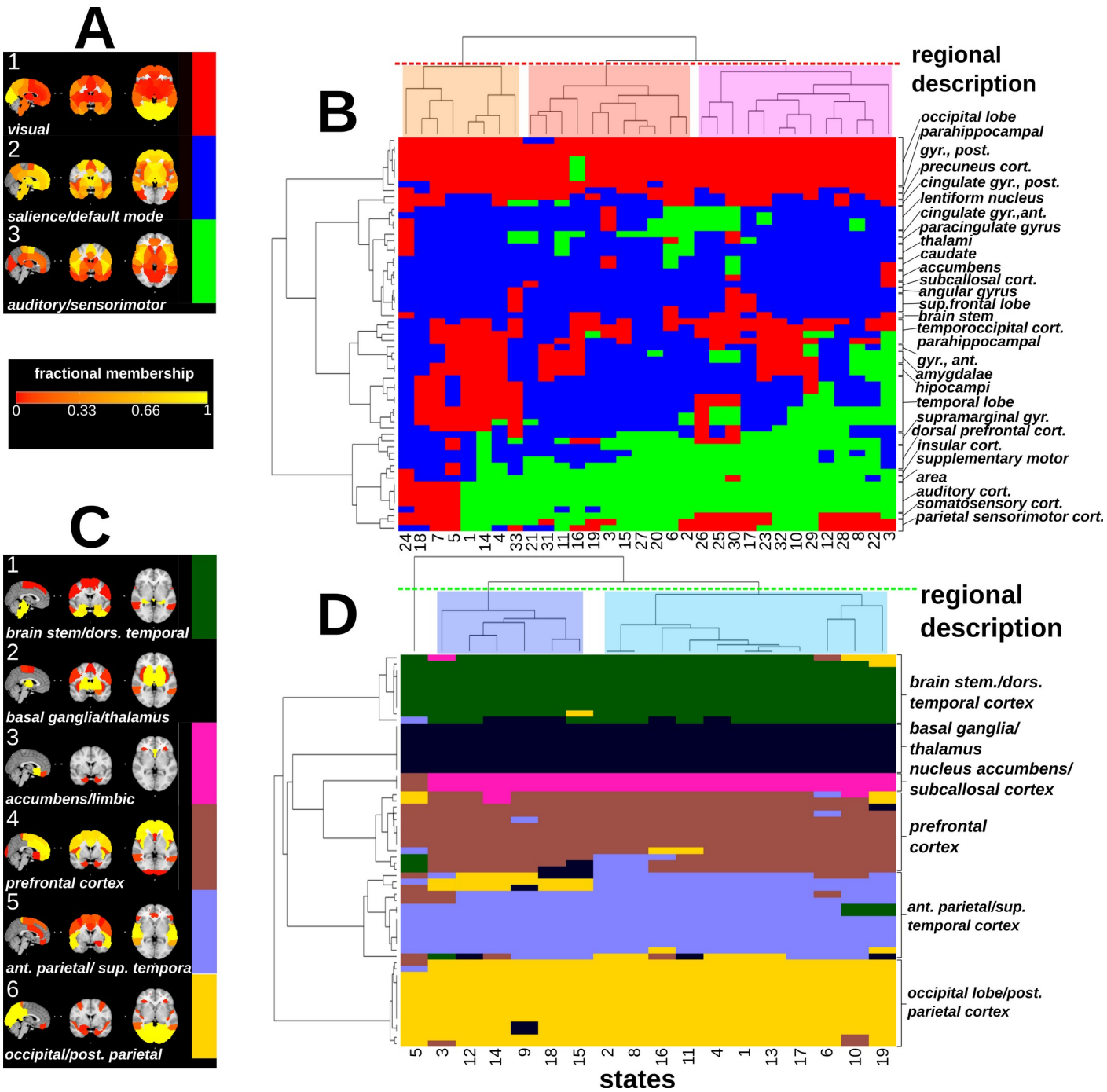


Figure 6: This figure summaries the results of multiplex consensus community detection across regions (rows) and states (columns). Each region is coloured according to its community membership. In addition, regions and states have been hierarchically clustered in both deep anaesthesia and wakefulness using Ward's algorithm, with Jaccard index similarity so that states with similar community composition and regions with similar membership are co-clustered (single colour patches). ***(A)*** *this plot shows the regional composition of each wakeful community as a brain map where colour represent the fractional membership of the region. Communities have been broadly described according to the text in white.* ***(B)*** *the clustergram of wakeful communities shows a heterogeneous pattern of changing community membership.* ***(C)*** *shows brain maps for the deep anaesthesia condition. Brain regions are largely static (fractional membership near to one) and have a relatively restricted anatomically organisation with more total communities.* ***(D)*** *the anaesthesia clustergram shows regional stratification with minimal reorganisation between states.*

2.6 Hub State Functional Connectivity and Homogeneity

Plots of the functional connectome thresholded to the top 5% strongest (by magnitude) of each of the sink states are shown in Fig. 7A, with mean activity maps for each shown vertically. The heatmaps represent the pairwise distance correlation (thresholded by 0.5) between states acrosss both conditions. Very few wakeful states share a significant relationship with any wakeful state on the basis of either functional activity (bottom left) or connectivity (top right). Fig. 7B shows that pairwise distance correlation of functional connectivity between wakeful states is significantly higher than in anaesthesia ($p<0.001$, Mann-Whitney U, two-sided) suggesting that, on average, anaesthesia states resemble neither each other nor are directly related to wakeful states. However, A-17, is at least somewhat related to all other anaesthesia states. A similar but weaker relationship holds for functional activity ($p<0.001$ Mann-Whitney U, two-sided) with mean functional activity correlation lower in anaesthesia than in wakefulness in Fig. 7C.

2.7 Brain State Functional to Structural Connectivity Comparison

Fig. 8A shows the results of comparing structural modules, as determined by consensus multi-resolution community analysis (Supplementary Figure 5 for these modules). Anaesthesia state community composition is a moderate but consistent match for the structural modules with an average community similarity, as measured by AMI, of 0.51, when compared to wakefulness at 0.24 ($p<0.001$, permutation test). We validated these results using the ARI (see Supplementary Section S.7), with similar results ($p<0.001$, permutation test). However, wakefulness was found to be more correlated in terms of the local clustering and degree centrality statistics ($p<0.001$, Mann-Whitney U, two-sided). This suggests that while structural modules are more similar to anaesthesia functional modules than wakeful functional modules, this similarity does not hold across other local connectivity measures.

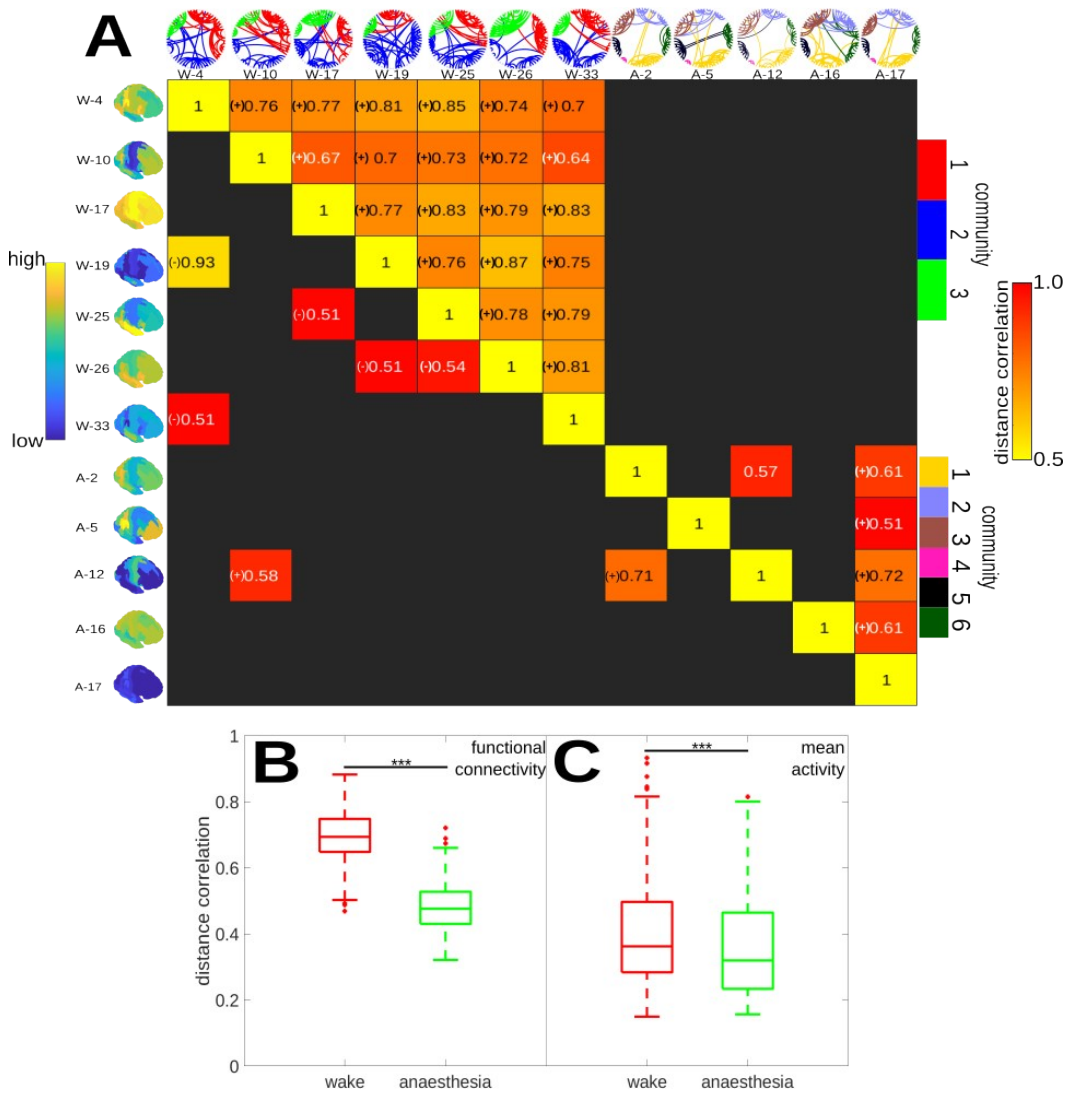


Figure 7: This figure illustrates the functional mean activity and connectivity across sink states in anaesthesia and wakefulness. ***(A)*** *This figure is a heatmap of the pairwise signed distance correlation between sink states functional activity (bottom left) and sink state connectivity (top right). Brain maps (left) are the same as those in Fig. 3. Community interconnectivity (top) is shown as a circular graph, depicting the level of intercommunity connectivity in each state.* ***(B)*** *This boxplot shows the signed distance correlation between all states functional connectivity in wakefulness and anaesthesia (Mann-Whitney U, p<0.001, two sided). Red dots denote outlying states.* ***(C)*** *Similarly, signed distance correlation is shown to be higher for mean activity amongst wakeful than anaesthesia states (Mann-Whitney U, p<0.001, two-sided).*

2.8 Sink State Occupancy Correlates with Age and Susceptibility

Multivariate linear regression was performed to determine which factors most closely predict the time subjects spent in dominant sink state A-17 (FO). Increasing age was found to be the single variable most associated with positive increases with A-17 ($p=0.005<0.01$, t-test, two-sided) when considered alone. However, time at loss of responsiveness (T_{LOBR}) in combination with age was found to be the best performing model of A-17 occurrence with respect to both the BIC and adjusted R^2 criteria (see Supplementary Table 2) with significant improvements in goodness of fit over the age only model ($p=0.040<0.05$, Chi-squared test). The model shows that age ($p=0.004<0.01$, t-test, two-sided) is a significant predictor, while T_{LOBR} is marginally negatively associated ($p=0.079$, t-test, two-sided) with

A-17 occurrence. Given the very high negative correlation of A-17 and A-12 occurrence this can be interpreted as T_{LOBR} being at least partially predictive of A-17 dominance over other sink states such as A-12. The final model showed no significant deviation from normality of its residuals ($p=0.981$, Kolmogorov-Smirnov test) or heteroskedacity in its errors ($p=0.131$, Breusch–Pagan-Kroenker test). Supplementary Figure 6 shows the alternative model which includes time T_{LOBR} conditioned on age. Surprisingly, grey matter density (see Supplementary Section S.8) did not add any significant model improvement (see Supplementary Tables 3, 4, 5 and 6).

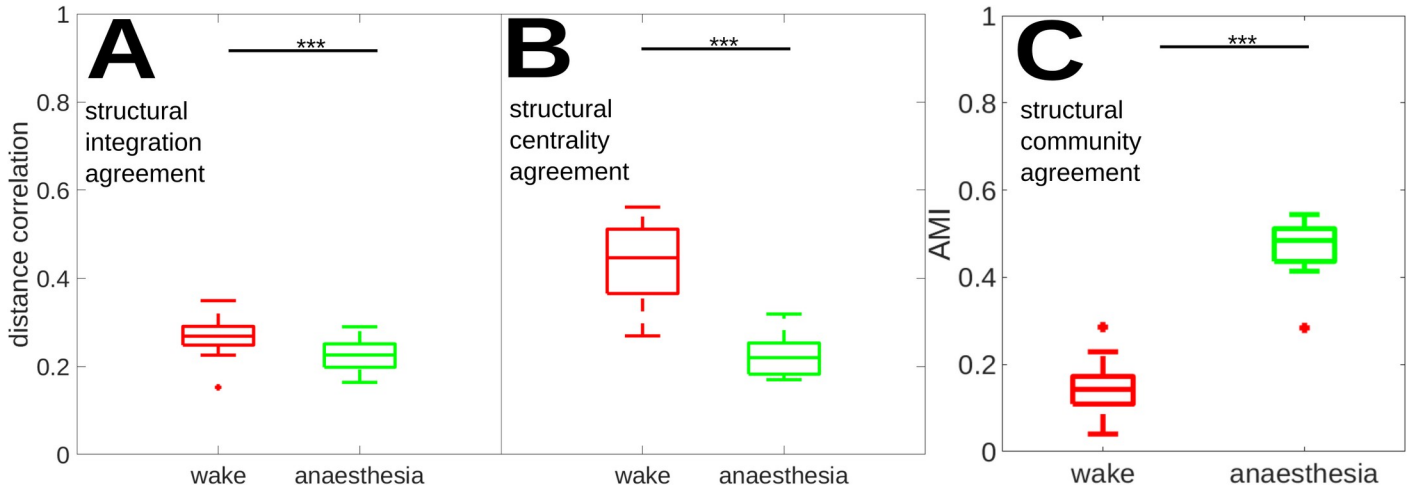


*Figure 8: This figure summarises the relationship of brain state functional connectivity in both conditions to structural connectivity. (**A**) shows the distance correlation between the local clustering coefficient for each state functional connectivity with the normative adult structural connectivity model (p<0.001, Mann-Whitney U, two-sided). (**B**) shows the distance correlation between the degree centrality measure in each state with the structural graph model (p<0.001, Mann-Whitney U, two-sided). (**C**) shows the AMI, a measure of modular similarity, when comparing the structural model to functional modules in wakeful and anaesthesia states respectively, suggesting that anaesthesia functional modules are more similar to the normitive structural modularisation of the brain (p<0.001, Mann-Whitney U, two-sided). Outlying samples are shown in red.*

3 Discussion

In this paper, we examined the differences in resting brain state dynamics between wakefulness and deep anaesthesia-induced unconsciousness using a graph model framework first proposed in [26]. Our results show that anaesthesia dynamics are dominated by sink states with a highly reduced capacity for complex information integration and decreased cortical and subcortical excitability when compared to wakefulness. Deep anaesthesia is temporally less complex with a reduced diversity of accessible brain states, suggesting a decrease in flexibility when compared to wakefulness. Furthermore, our results show that brain state functional connectivity in anaesthesia does not strongly resemble the dynamic re-wiring of regional networks seen in wakefulness but instead is stratified into structurally confined islands of interacting regions with little re-configuration occurring between states.

While a dose-dependent reduction in the number of unique brain states has been observed in subjects undergoing propofol anaesthesia [27], a characterisation of these states in terms of the brain state dynamics that might maintain this condition (dynamic attraction towards a small number low complexity, low activity sink states) has not been shown before. We show that functional and temporal information capacity is reduced and propose principled, information theoretic approaches to measure these along multiple axes including a novel measure of information rate. We demonstrate a link

between sink state dominance and the subject's anaesthetic susceptibility specifically their age, and to a lesser extent, time at loss of responsiveness. The findings suggest that subjects with varying susceptibility may arrive at different anaesthesia end-points with differing capacities for information integration and excitability. This trajectory change persists even after subjects are held at the same high anaesthesia dosage for nearly an hour, which suggests significant subject-specific inertia. Age has been demonstrated to be an important predictor of anaesthesia susceptibility with continuous decreases in EEG alpha power and coherence from young adulthood. However these findings would benefit from validation in a larger cohort covering a broader age-range [61].

We observe fragmentation of functional modules under anaesthesia into small, static subnetworks with little thalamocortical or even cortical-subcortical re-organisation from state to state. In particular, thalamic community associations mostly occurred between neighbouring subcortical and entorhinal regions. Notably, some interchange is observed between prefrontal, parietal and nucleus accumbens, suggesting some sporadic default mode-like activity. Deteriorated but persistent default mode network activity has been observed in other studies of anaesthesia functional connectivity, with the anterior portions being most affected [62]. This thalamocortical isolation has also been shown from static measures of functional connectivity [12]. In our model, anaesthesia is contrasted with wakefulness in which larger communities dynamically exchange members between states to facilitate more complex task performance required for conscious awareness and for responsiveness to the external and internal environment [28].

Our study was limited in its structure-function comparisons by the availability of patient-specific structural (tractography) data. Sample size and small age range may also explain the apparent weak association between grey matter density and age. Finding a parcellation that was computationally tractable and relevant in a functional context was prioritised partially due to the relatively small cohort size for the ultra-slow anaesthethesia study. Due to this choice to prioritise coverage over spatial resolution some cortical areas lacked comparable left and right regions, preventing us from observing possible lateralisation of activity in some areas.

The strong attractor-like dynamics seen in anaesthesia towards low complexity, low excitability anaesthesia states have implications for the imperturbability of subjects from anaesthesia. These results also have possible implications DOCs which may show similar patterns of imperturbability from low complexity unconscious states [63][64]. Our findings have implications for targeting of specific anaesthesia depth during surgery and give novel evidence for early effects of age on a patients' anaesthesia trajectory. Patient susceptibility to one or another sink state may also have important implications for traditional ECG and EEG-based depth of anaesthesia models such as the Bispectral Index, or novel slow wave-based approaches [65][66][67]. More sophisticated measures of complexity and emergence for dynamical brain state activity are now being developed [68], but tractable methods that work with current models are needed. In particular methods are needed that take into account both functional and temporal complexity; taking into account the brains dynamic capacity for integrating the interdependent streams of information that support conscious awareness [69] [18] [70].

Acknowledgements

The authors are particularly grateful to Marco Fabus for his assistance in identifying periods of burst suppression from subject EEG. Minh Tran for his helpful advice in model design and testing.

Funding

J.W. acknowledges support from the Commonwealth Scholarship Commission, Oppenheimer Trust, Human Brain Project: Special Grant Agreement 3, The Alan Turing Institute UKRI Block grant, and EPSRC grant BB/Y003020/1. G.R. acknowledges support from EPSRC grants EP/T018445/1, EP/W037211/1, EP/V056883/1, and EP/R018472/1.

References

1 Urban B, Bleckwenn M. Concepts and correlations relevant to general anaesthesia *Brit J Anaesth*. 2002;89(1):3-16.

2 Errando C, Sigl J, Robles M, Calabuig E, Garcia J, Arocas F, Higueras R, Del Rosario E, Lopez D, Peiro C, others. Awareness with recall during general anaesthesia: a prospective observational evaluation of 4001 patients *Brit J Anaesth*. 2008;101(2):178-185.

3 Schwender D, Kunze-Kronawitter H, Dietrich P, Klasing S, Forst H, Madler C. Conscious awareness during general anaesthesia: patients' perceptions, emotions, cognition and reactions *Brit J Anaesth*. 1998;80(2):133-139.

4 Trabold B, Metterlein T. Postoperative delirium: risk factors, prevention, and treatment *J Cardiothor Vasc An*. 2014;28(5):1352-1360.

5 Wu L, Zhao H, Weng H, Ma D. Lasting effects of general anesthetics on the brain in the young and elderly:“mixed picture” of neurotoxicity, neuroprotection and cognitive impairment *Journal of Anesthesia*. 2019;33:321-335.

6 Bonhomme V, Staquet C, Montupil J, Defresne A, Kirsch M, Martial C, Vanhaudenhuyse A, Chatelle C, Larroque SK, Raimondo F, others. General anesthesia: a probe to explore consciousness *Front. Syst. Neurosci.*. 2019;13:36.

7 Stevner A, Vidaurre D, Cabral J, Rapuano K, Nielsen SFV, Tagliazucchi E, Laufs H, Vuust P, Deco G, Woolrich M, others. Discovery of key whole-brain transitions and dynamics during human wakefulness and non-REM sleep *Nat. Commun.*. 2019;10(1):1035.

8 Bai Y, He J, Xia X, Wang Y, Yang Y, Di H, Li X, Ziemann U. Spontaneous transient brain states in EEG source space in disorders of consciousness *NeuroImage*. 2021;240:118407.

9 Lopez-Gonzalez A, Panda R, Ponce-Alvarez A, Zamora-Lopez G, Escrichs A, Martial C, Thibaut A, Gosseries O, Kringelbach ML, Annen J, others. Loss of consciousness reduces the stability of brain hubs and the heterogeneity of brain dynamics *Commun. Biol*. 2021;4(1):1037.

10 Picchioni D, Pixa ML, Fukunaga M, Carr WS, Horovitz SG, Braun AR, Duyn JH. Decreased connectivity between the thalamus and the neocortex during human nonrapid eye movement sleep *Sleep*. 2014;37(2):387-397.

11 Heine L, Soddu A, Gomez F, Vanhaudenhuyse A, Tshibanda L, Thonnard M, Charland-Verville V, Kirsch M, Laureys S, Demertzi A. Resting state networks and consciousness *Front. Psychol.*. 2012;3:295.

12 Mhuircheartaigh RN, Warnaby C, Rogers R, Jbabdi S, Tracey I. Slow-wave activity saturation and thalamocortical isolation during propofol anesthesia in humans *Sci Transl Med*. 2013;5(208):208ra148-208ra148.

13 Biggs D, Boncompte G, Pedemonte JC, Fuentes C, Cortinez LI. The effect of age on electroencephalogram measures of anesthesia hypnosis: A comparison of BIS, Alpha Power, Lempel-Ziv complexity and permutation entropy during propofol induction *Front. Aging Neurosci.*. 2022;:876.

14 Jia L, Hou J, Zheng H, Sun L, Fan Y, Wang X, Hao M, Li Y, Yang T. Study of the rational dose of propofol in elderly patients under bispectral index monitoring during total intravenous anesthesia: A PRISMA-compliant systematic review *Medicine*. 2020;99(5):

15 Dasgupta M, Dumbrell AC. Preoperative risk assessment for delirium after noncardiac surgery: a systematic review *J. Am. Geriatr. Soc.*. 2006;54(10):1578-1589.

16 Drummond JC, Weiskopf RB. Monitoring depth of anesthesia: with emphasis on the application of the bispectral index and the middle latency auditory evoked response to the prevention of recall *The Journal of the American Society of Anesthesiologists*. 2000;93(3):876-882.

17 Kolmogorov AN. On tables of random numbers *Sankhya: The Indian Journal of Statistics, Series A*. 1963;:369-376.

18 Ruffini G. An algorithmic information theory of consciousness *Neurosci of Consciousness*. 2017;3(1):

19 Solms M. The hard problem of consciousness and the free energy principle *Front. Psychol.*. 2019;9:2714.

20 Fan S-Z, Yeh J-R, Chen B-C, Shieh J-S, others. Comparison of EEG approximate entropy and complexity measures of depth of anaesthesia during inhalational general anaesthesia *J Med Biol Eng*. 2011;31(5):359-366.

21 Marshall W, Gomez-Ramirez J, Tononi G. Integrated information and state differentiation *Front. Psychol.*. 2016;7:926.

22 Boly M, Sasai S, Gosseries O, Oizumi M, Casali A, Massimini M, Tononi G. Stimulus set meaningfulness and neurophysiological differentiation: a functional magnetic resonance imaging study *PloS One*. 2015;10(5):e0125337.

23 Sevenius Nilsen A, Juel BE, Marshall W. Evaluating approximations and heuristic measures of integrated information *Entropy*. 2019;21(5):525.

24 Galatolo S, Hoyrup M, Rojas C. Effective symbolic dynamics, random points, statistical behavior, complexity and entropy *Inform. and Comput.*. 2010;208(1):23-41.

25 Mediano PA, Rosas FE, Luppi AI, Noreika V, Seth AK, Carhart-Harris RL, Barnett L, Bor D. Spectrally and temporally resolved estimation of neural signal diversity *bioRxiv*. 2023;:2023-03.

26 Wilsenach JB, Warnaby CE, Deane CM, Reinert GD. Ranking of communities in multiplex spatiotemporal models of brain dynamics *Appl. Net Sci*. 2022;7(1):1-22.

27 Hudetz AG, Liu X, Pillay S. Dynamic repertoire of intrinsic brain states is reduced in propofol-induced unconsciousness *Brain Connect.*. 2015;5(1):10-22.

28 Barttfeld P, Uhrig L, Sitt JD, Sigman M, Jarraya B, Dehaene S. Signature of consciousness in the dynamics of resting-state brain activity *Proc Natl Acad Sci.*. 2015;112(3):887-892.

29 Hutchison RM, Womelsdorf T, Gati JS, Everling S, Menon RS. Resting-state networks show dynamic functional connectivity in awake humans and anesthetized macaques *Hum Brain Mapp*. 2013;34(9):2154-2177.

30 Shine JM, Bissett PG, Bell PT, Koyejo O, Balsters JH, Gorgolewski KJ, Moodie CA, Poldrack RA. The dynamics of functional brain networks: integrated network states during cognitive task performance *Neuron*. 2016;92(2):544-554.

31 Fortunato S, Barthelemy M. Resolution limit in community detection *Proceedings of the Natl Acad of Sci*. 2007;104(1):36-41.

32 Wilsenach J. Computational network models for molecular, neuronal and brain data in the presence of long range dependence *DPhil Thesis*. 2022;:

33 Forrest F, Tooley M, Saunders P, Prys-Roberts C. Propofol infusion and the suppression of consciousness: the EEG and dose requirements *BJA: Brit J Anaes*. 1994;72(1):35-41.

34 Jenkinson M, Beckmann CF, Behrens TE, Woolrich MW, Smith SM. FMRIB Software Library (FSL) *NeuroImage*. 2012;62(2):782-790.

35 Mazziotta J, Toga A, Evans A, Fox P, Lancaster J, Zilles K, Woods R, Paus T, Simpson G, Pike B, others. A four-dimensional probabilistic atlas of the human brain *J Am Med Inform Assoc.*. 2001;8(5):401-430.

36 Fritz BA, Kalarickal PL, Maybrier HR, Muench MR, Dearth D, Chen Y, Escallier KE, Abdallah AB, Lin N, Avidan MS. Intraoperative electroencephalogram suppression predicts postoperative delirium *Anesth Analg*. 2016;122(1):234.

37 Zhang Z, Cai D-C, Wang Z, Zeljic K, Wang Z, Wang Y. Isoflurane-induced burst suppression increases intrinsic functional connectivity of the monkey brain *Front Neurosci*. 2019;13:296.

38 Horn A, Blankenburg F. Toward a standardized structural--functional group connectome in MNI space *NeuroImage*. 2016;124:310-322.

39 Reisert M, Mader I, Anastasopoulos C, Weigel M, Schnell S, Kiselev V. Global fiber reconstruction becomes practical *NeuroImage*. 2011;54(2):955-962.

40 Horn JL. A rationale and test for the number of factors in factor analysis *Psychometrika*. 1965;30(2):179-185.

41 Wang R, Benner T, Sorensen AG, Wedeen VJ. Diffusion toolkit: a software package for diffusion imaging data processing and tractography 2007; **15**: 1.

42 Yeh F-C, Verstynen TD, Wang Y, Fernandez-Miranda JC, Tseng W-YI. Deterministic diffusion fiber tracking improved by quantitative anisotropy *PloS One*. 2013;8(11):e80713.

43 Yeh F-C, Panesar S, Fernandes D, Meola A, Yoshino M, Fernandez-Miranda JC, Vettel JM, Verstynen T. Population-averaged atlas of the macroscale human structural connectome and its network topology *NeuroImage*. 2018;178:57-68.

44 Lancichinetti A, Fortunato S. Consensus clustering in complex networks *Sci. Rep.*. 2012;2(1):1-7.

45 Mucha PJ, Richardson T, Macon K, Porter MA, Onnela J-P. Community structure in time-dependent, multiscale, and multiplex networks *Science*. 2010;328(5980):876-878.

46 Vidaurre D, Quinn AJ, Baker AP, Dupret D, Tejero-Cantero A, Woolrich MW. Spectrally resolved fast transient brain states in electrophysiological data *NeuroImage*. 2016;126:81-95.

47 Desikan RS, Segonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, others. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest *Neuroimage*. 2006;31(3):968-980.

48 Pattern recognition and machine learning : Springer; 2006.

49 Salvador R, Suckling J, Coleman MR, Pickard JD, Menon D, Bullmore E. Neurophysiological architecture of functional magnetic resonance images of human brain *Cerebral cortex*. 2005;15(9):1332-1342.

50 Sen B, Chu S-H, Parhi KK. Ranking regions, edges and classifying tasks in functional brain graphs by sub-graph entropy *Sci Rep*. 2019;9(1):7628.

51 Newman ME. Modularity and community structure in networks *Proc. Natl. Acad. Sci.*. 2006;103(23):8577-8582.

52 Brin S, Page L. The anatomy of a large-scale hypertextual web search engine *Computer Networks and ISDN Systems*. 1998;30(1-7):107-117.

53 Elements of information theory : John Wiley & Sons; 1999.

54 Geerligs L, Henson RN, others. Functional connectivity and structural covariance between regions of interest can be measured more accurately using multivariate distance correlation *NeuroImage*. 2016;135:16-31.

55 Pardo-Diaz J, Bozhilova LV, Beguerisse-D}az M, Poole PS, Deane CM, Reinert G. Robust gene coexpression networks using signed distance correlation *Bioinformatics*. 2021;37(14):1982-1989.

56 Girvan M, Newman ME. Community structure in social and biological networks *Proc Natl Acad Sci*. 2002;99(12):7821-7826.

57 Ward Jr JH. Hierarchical grouping to optimize an objective function *J. Amer. Statist Assoc.*. 1963;58(301):236-244.

58 Romano S, Vinh NX, Bailey J, Verspoor K. Adjusting for chance clustering comparison measures *J Mach Learn Res*. 2016;17(1):4635-4666.

59 Nunes CS, Alonso H, Castro A, Amorim P, Mendonca T. Towards the control of depth of anaesthesia: Identification of patient variability 2007;: 3109-3115.

60 Alitto HJ, Usrey WM. Corticothalamic feedback and sensory processing *Curr. Opin. Neurobiol.*. 2003;13(4):440-445.

61 Purdon PL, Pavone K, Akeju O, Smith A, Sampson A, Lee J, Zhou D, Solt K, Brown E. The ageing brain: age-dependent changes in the electroencephalogram during propofol and sevoflurane general anaesthesia *Brit J Anaesth*. 2015;115(suppl_1):i46-i57.

62 Luppi AI, Craig MM, Pappas I, Finoia P, Williams GB, Allanson J, Pickard JD, Owen AM, Naci L, Menon DK, others. Consciousness-specific dynamic interactions of brain integration and functional diversity *Nat Comms*. 2019;10(1):4616.

63 Pryor K, Root J, Mehta M, Stern E, Pan H, Veselis R, Silbersweig D. Effect of propofol on the medial temporal lobe emotional memory system: a functional magnetic resonance imaging study in human subjects *Brit J Anaesth*. 2015;115(suppl_1):i104-i113.

64 Billard V, Gambus PL, Chamoun N, Stanski DR, Shafer SL. A comparison of spectral edge, delta power, and bispectral index as EEG measures of alfentanil, propofol, and midazolam drug effect *Clin Pharma Therap*. 1997;61(1):45-58.

65 Sigl JC, Chamoun NG. An introduction to bispectral analysis for the electroencephalogram *J Clin Monitor*. 1994;10:392-404.

66 Warnaby CE, Sleigh JW, Hight D, Jbabdi S, Tracey I. Investigation of slow-wave activity saturation during surgical anesthesia reveals a signature of neural inertia in humans *Anesthesiology: J Am Soc Anesth*. 2017;127(4):645-657.

67 Rosas FE, Mediano PA, Jensen HJ, Seth AK, Barrett AB, Carhart-Harris RL, Bor D. Reconciling emergences: An information-theoretic approach to identify causal emergence in multivariate data *PLoS Comput. Biol*. 2020;16(12):e1008289.

68 Safron A. An integrated world modeling theory (IWMT) of consciousness: combining integrated information and global neuronal workspace theories with the free energy principle and active inference framework; toward solving the hard problem and characterizing agentic causation *Frontiers in Artificial Intelligence*. 2020;3:30.

69 Makeig S, Debener S, Onton J, Delorme A. Mining event-related brain dynamics *Trends Cogn Sci*. 2004;8(5):204-210.

70 Information theory, inference and learning algorithms : Cambridge university press; 2003.

71 Markov chains: from theory to implementation and experimentation : John Wiley & Sons; 2017.

72 Baum LE, Petrie T. Statistical inference for probabilistic functions of finite state Markov chains *Ann Math Stats*. 1966;37(6):1554-1563.

73 Jaynes ET. Information theory and statistical mechanics *Phys Rev*. 1957;106(4):620.

74 Hanteer O, Magnani M. Unspoken assumptions in multi-layer modularity maximization *Sci Rep.* 2020;10(1):1-15.

75 Traag VA, Waltman L, Van Eck NJ. From Louvain to Leiden: guaranteeing well-connected communities *Sci Rep.* 2019;9(1):1-12.

76 Hubert L, Arabie P. Comparing partitions *J. Classification*. 1985;2(1):193-218.

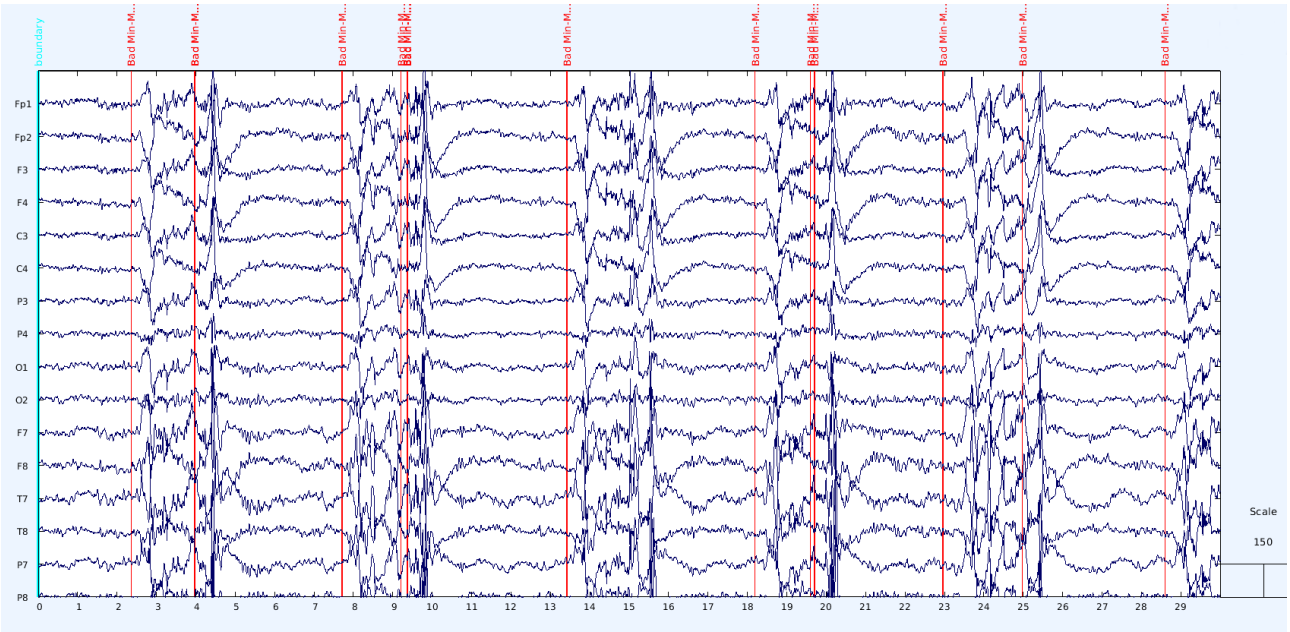
77 Zhang Y, Brady M, Smith S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm *IEEE Trans Med Imag*. 2001;20(1):45-57.

Supplementary Information

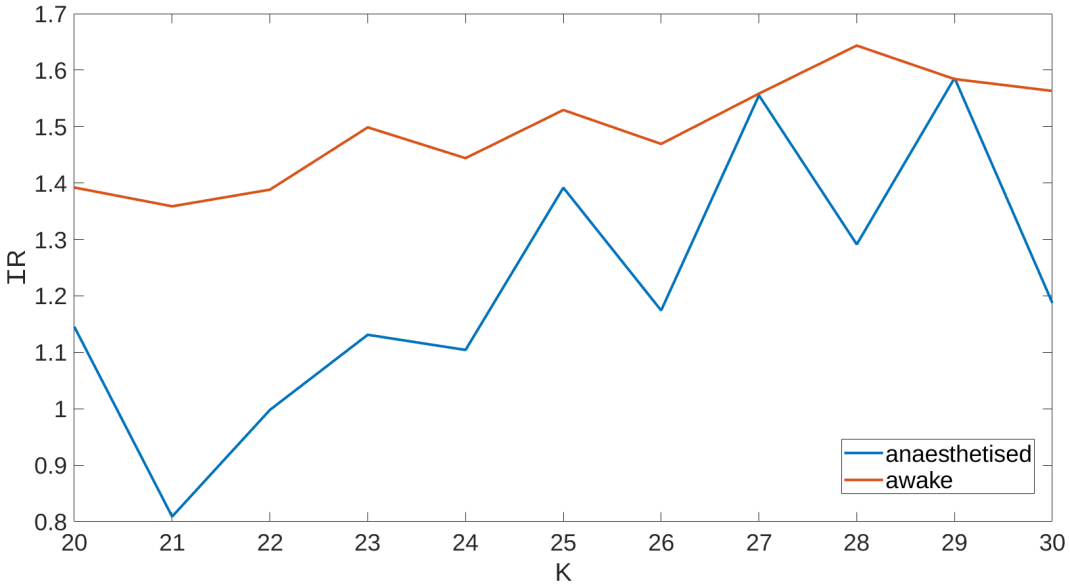
This section details supplementary tables and information for the study, organised into sections for each of the relevant methods used in the paper.

subject ID	1	2	3	4	5	6	7	8	9	10	11	12	13
age	21	24	21	23	30	29	29	33	19	25	32	25	34

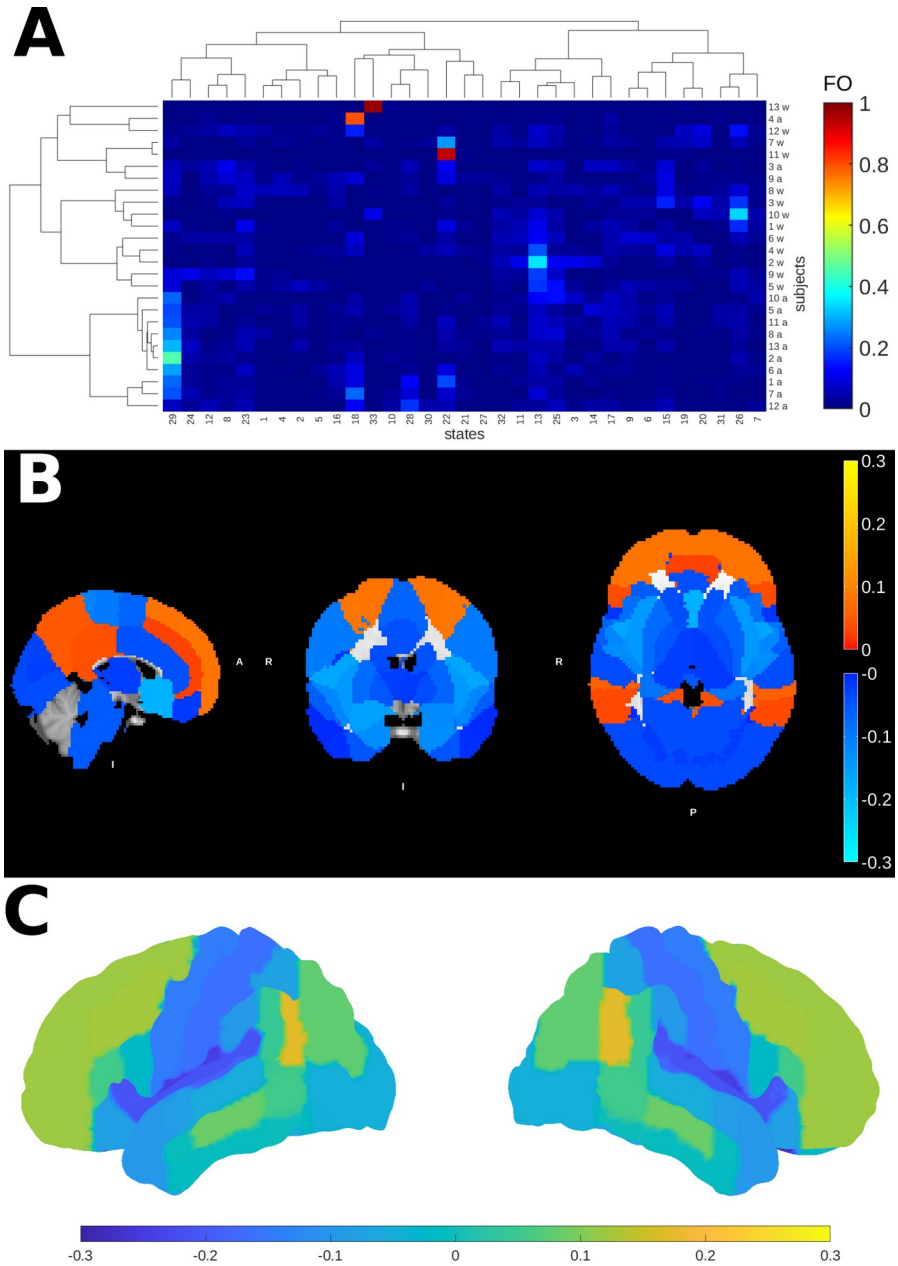
Supplementary Table 1:This table shows the anonymised subject ID and age of each participant used in the final analysis (after exclusion criteria were applied). The mean age of participants included in the study was 26.539±4.943, the median was 25, the minimum was 19 and maximum was 34.



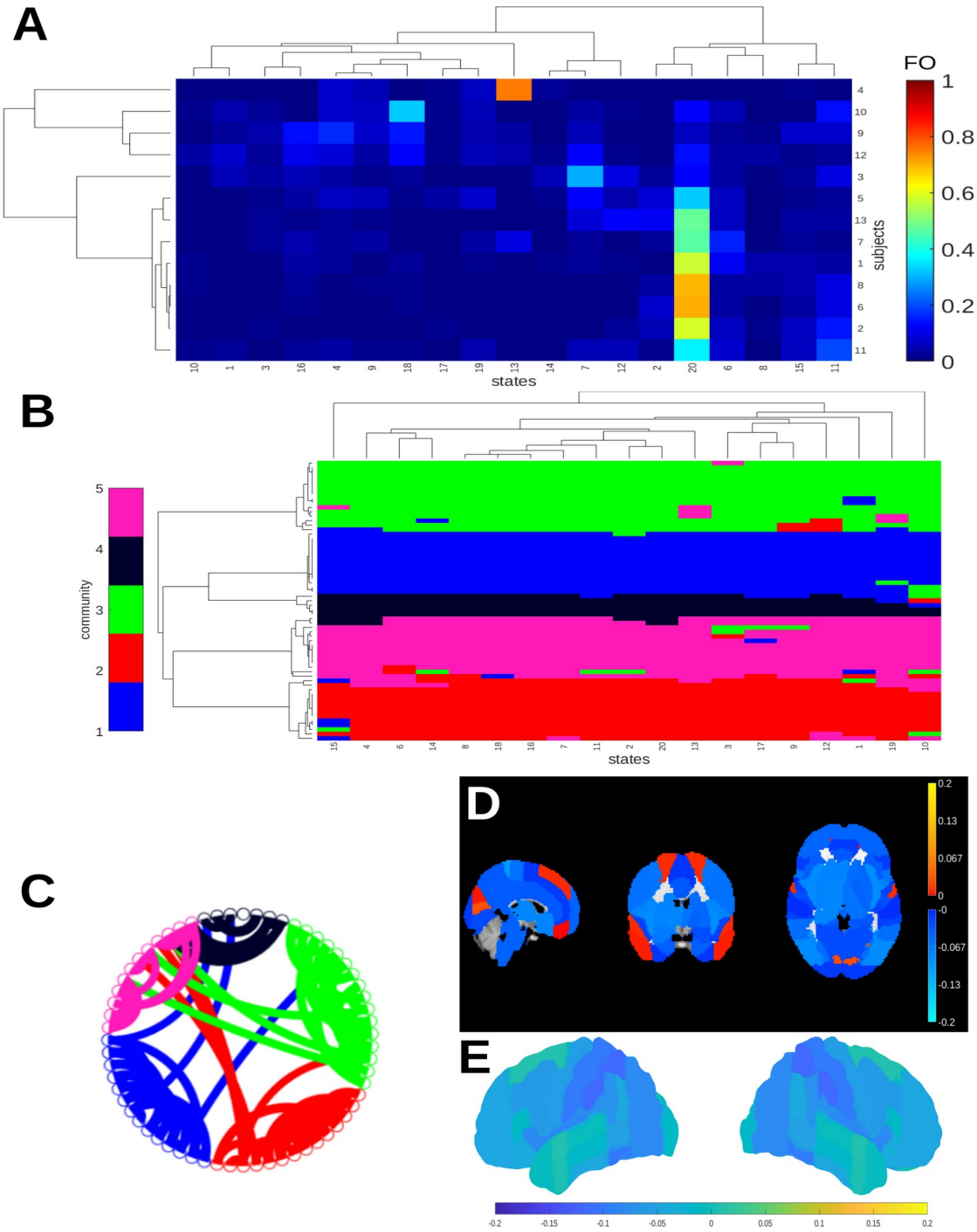
Supplementary Figure 1: This figure shows the burst suppressed activity in a subset of electrodes recorded for one of the subjects that was excluded. The red Min-Max markers show the start location of local peaks in activity which are common in the burst regime of burst suppression. This figure was generated in EEGLAB [69].



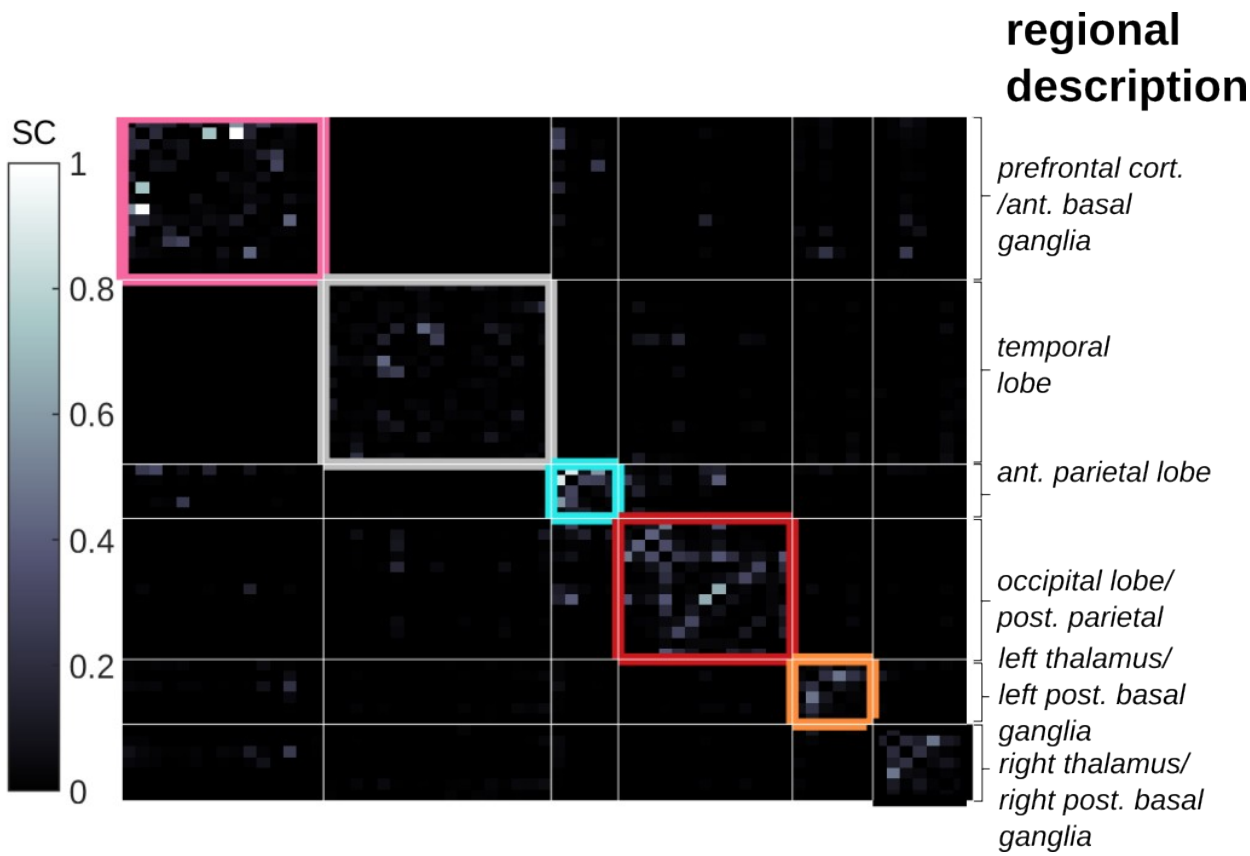
Supplementary Figure 2: This figure shows the information rate (IR) in anaesthesia and wakefulness as a function of the number of states K. The information rate in anaesthesia is always at or below the information rate for wakefulness for all models having the same number of states.



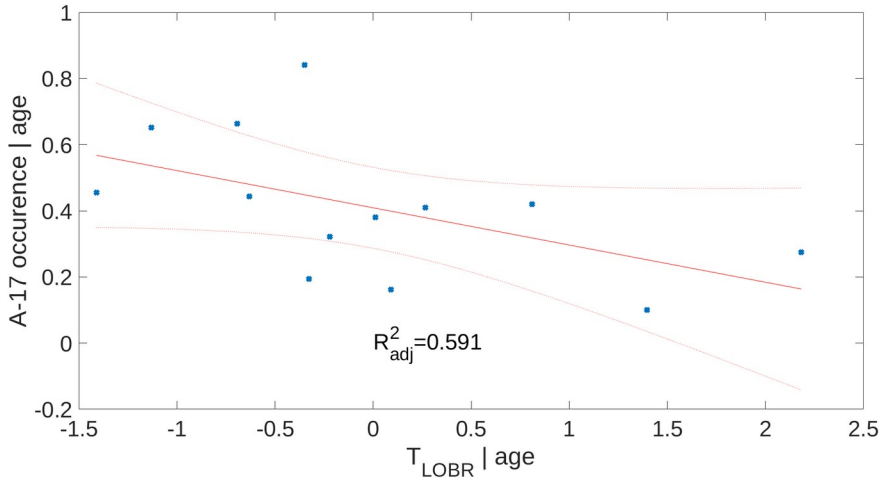
Supplementary Figure 3: This figure summarises the results of training the HMGM on the combined anaesthesia-wakefulness trial data with $K_{comb} = 33$ *states (the same as the wakeful model).* ***(A)*** *This figure shows a clustergram of Fractional Occupancy (FO) over subject-trial with "w" referring to wakeful trials and "a" for deep anaesthesia trials. As in the original model, anaesthesia trials are dominated by a single state, state 29 in this case. As in the single condition anaesthesia model with* $K_{anaes} = 18$*, dominant sink state occurrence is positively correlated with age (* $p = 0.005 < 0.01$*, t-test, two-sided).* ***(B)*** *Cross-sectional activity maps of state 29 also show generally low absolute mean functional activity but exhibit clear increases in activity above baseline in the precuneous and prefrontal cortices. It is notable that the baseline here includes wakeful activity as well as deep anaesthesia activity and so cannot be readily compared to anaesthesia only model.* ***(C)*** *The surface plot of state 29 shows the extent of prefrontal activation in the dorsal prefrontal cortex relative to baseline.*



Supplementary Figure 4: This figure summarises the results of fitting the anaesthesia model with one additional initial state, after state exclusion the resulting model has $K_{anaes} = 19$*. states. (A) This figure shows a clustergram of Fractional Occupancy (FO) over subjects for the model. As in the* $K_{anaes} = 18$ *anaesthesia model, subjects are dominated by a sink state, state 20. (B) The consensus community detection shows that stratification of activity in deep anaesthesia is conserved in the larger model, with the notable difference that most subcortical regions have combined into a single consensus community 1. (C) Sink state 20 connectivity is again relatively low between communities with low functional activity across the brain ((D) and (E)). These findings corroborate those for the simpler model with* $K_{anaes} = 18$*, and show that similar dominant sink states and dynamics emerge.*



Supplementary Figure 5: This figure summarises the structural connectivity and community composition of the structural graph model. Communities are shown as blocks with distinct colour borders and a general regional description is given to each community. These communities are the result of consensus community detection applied at the same resolution range as for the HMGM, where the range of values used is $\Gamma = \{1,2,3\}$*. Structural communities, like HMGM communities have a range of sizes and tend to be more lateralised than in the functional case due to the low levels of interhemispheric connectivity.*



Supplementary Figure 6: This figure shows the added variable plot for the multivariate linear regression of occurrence (in terms of FO) of sink state A-17, conditioned on age, against time at Loss of Behavioural Responsive (T_{LOBR}), also conditioned on age. The plot shows the residual data after fitting of the age only model (blue dots), the line of best fit (red line) has the slope equal to the effect size of T_{LOBR} *with 95% confidence interval (red dotted lines) shown (*$p = 0.005$*, F-test). A significant positive relationship between A-17 occurrence and age persists (*$p = 0.004$*, t-test, two-sided) and a marginally significant negative relationship between occurrence of A-17 and TLOBR is evident (*$p = 0.079$*, t-test, two-sided).*

variables	age	age+T_{LOBR}	age+GM_{LOBR}	age+T_{LOBR}+GM_{LOBR}
BIC	0.733	-0.917	3.250	1.633

Supplementary Table 2: This table shows the Bayesian Information Criterion (BIC), a Bayesian model selection criterion for the four models, shown in Supplementary Figures 2 and 3. For linear models, lower BIC is an indicator that a model better predicts the data given the number of dependent variables. Higher numbers of dependent variables are penalised by increasing the BIC. A Chi-squared was also performed to compare the likelihood of both models. The first model including only age was rejected in favour of the second model (age+T_{LOBR}),at the $p=0.040$ *significance level (Chi-Squared likelihood ratio test).*

variable	estimate	std.	t-statistic	p-value
intercept	-0.801	0.350	-2.289	0.043 *
age	0.046	0.013	3.513	0.005 **

Supplementary Table 3: The results of fitting a multivariate linear model of A-17 occurrence (in terms of FO) against age. The table shows the estimated variable names, the estimated coefficients, the standard deviation of these coefficients, the t-statistics of each coefficient and the associated two-sided t-test p-value. The model explains the variance significantly better than the constant model at the $p=0.005<0.01$ *significance level (F-test).*

variable	estimate	std.	t-statistic	p-value
intercept	-0.34141	0.390	-0.875	0.402
age	0.045574	0.013	3.713	0.004 **
T_{LOBR}	3.365×10^{-4}	1.719×10^{-4}	-1.957	0.079 (*)

Supplementary Table 4: The results of fitting a multivariate linear model of A-17 occurrence (in terms of FO) against age. The table shows the estimated variable names, the estimated coefficients, the standard deviation (std.) of these coefficients, the t-statistic of each coefficient and the associated two-sided p-value. The model explains the variance significantly better than the constant model at the $p=0.017$ *significance level (F-test).*

variable	estimate	std.	t-statistic	p-value
intercept	-0.302	2.617	-0.116	0.910
age	0.045	0.014	3.256	0.009 **
$GM_{cort.}$	-0.883	4.591	-0.192	0.851

Supplementary Table 5: The results of fitting a multivariate linear model of A-17 occurrence (in terms of FO) against age and cortical grey matter density ($GM_{cort.}$). The table shows the estimated variable names, the estimated coefficients, the standard deviation (std.) of these coefficients, the t-statistic of each coefficient and the associated two-sided p-value. The model explains the variance significantly better than the constant model at the $p=0.023$ *significance level (F-test).*

variable	estimate	std.	t-statistic	p-value
intercept	-0.107	2.351	-0.045	0.96475
age	0.043	0.012	3.444	0.007 **
T_{LOBR}	-3.353×10^{-4}	1.814×10^{-4}	-1.848	0.098 (*)
$GM_{cort.}$	-0.418	4.127	-0.101	0.921

Supplementary Table 6: The results of fitting a multivariate linear model of A-17 occurrence (in terms of FO) against age. The table shows the estimated variable names, the estimated coefficients, the standard deviation (std.) of these coefficients, the t-statistic of each coefficient and the associated two-sided p-value. The model explains the variance significantly better than the constant model at the $p=0.017$ *significance level (F-test).*

S.1 Markov Models and Hidden Markov Models

The basic form of a finite state Markov model is as a dynamic model in which, for our purposes, a subject is said to occupy a state s', in a finite set S of states, with some probability determined by the previous state s. In the case that the model describes a stationary process, these probabilities can be summarised in terms of a matrix P, with entries

$$P_{s,s'} = Pr(s \rightarrow s' | s),$$

where each entry describes the conditional probability of moving from state s to s' at some time given the current state s.

A Hidden Markov Model (HMM) is one in which the state, s is not directly observed but has to be inferred from the data. In our case we model each observation (time point) of the fMRI signal as arising from a multivariate Gaussian $O(s) \sim N(\mu(s), \Sigma(s))$ with mean and covariance dependent on the underlying state s. Such a model is more difficult to fit than a model with fixed, state independent mean anc covariance, but established approximate fitting procedures exist. Here, we use the minimal variational free energy approach applied in the HMMMAR package, a MATLAB package for learning of HMM models [47]. The model is trained using multivariate time series data X, with vector entries X_t for $1 \leq t \leq T$ which is derived from the BOLD signal up to the final time T. As part of the fitting procedure, data was dimensionally reduced following the Parallel Analysis procedure used in [26] which uses randomised surrogate data to determine the optimal number of Principal Components to use in the reduction. First the eigenmatrix U is calculated by Singular Value Decomposition (SVD) from the original dataset X^*, which has dimension $D \times T$, the reduced dimension $d < D$ is then calculated by Parallel Analysis (PA) and the final dataset X is projected out using the reduced form of $U_{d \times T}$, with d columns and used as input to train the HMM. The process is summarised below as

$$X^* \underset{SVD}{\rightarrow} U \underset{PA}{\rightarrow} U_{d \times T} X^* = X.$$

After training, the number of states in each model is selected by maximising the entropy over the cross-validated maximum entropy of the Fractional Occupancy (FO), following the procedure in [26]. Fractional Occupancy (FO) is the fraction of time a subject n is expected to spend in a state s over a trial or recording, it is calculated from the a postereori probabilities of a state occurring across all time points in the recording. It can be represented in matrix form, with entries

$$FO_{n,s} = \sum_{1 \leq t \leq T} \frac{Pr(S_t = s | X)}{T},$$

where the nth row can be interpreted as a distribution representing the probability of observing subject n across states in S over the course of the recording. For more details on Markov models and the related HMMs see [33], [71][72]. The motivation for this procedure is based on the statistical principle of maximum entropy which states that under conditions of only partial observations of the system (e.g. indirect observations of brain activity via the BOLD signal), the best model of the system is the model which retains maximal uncertainty about the unobserved components, while under the constraints of the model assumptions and available data [73].

S.2 Consensus Louvain Community Detection

Structural community detection is performed over the state transition graph $G=(V,A)$ (with regions in $x,y\in V$ and weighted, symmetric adjacency A). This method improves on standard Louvain modularity-based community detection [46], which uses the standard modularity score

$$Q(\boldsymbol{C})=\sum_{x,y\in V}\left[A^{x,y}-\gamma\frac{d^x d^y}{\sum_{u,v\in V}A^{u,v}}\right]\delta_{\boldsymbol{C}}(x,y),$$

where d^x is the weighted degree of $x\in V$, $\gamma>0$ is the resolution parameter and

$$\delta_{\boldsymbol{C}}(x,y)=\begin{cases}1 & \text{if } \exists C\in\boldsymbol{C} \text{ such that } x,y\in C\\ 0 & \text{otherwise,}\end{cases},$$

is the Dirac delta for membership in the partition $\boldsymbol{C}$. The Louvain algorithm includes some randomness in partitioning and hence different runs may give different results. The robust method improves on directly optimising $Q(\boldsymbol{C})$ by considering a plausible range of resolution parameters $\{\gamma_1,\gamma_2,\ldots,\gamma_g\}=\Gamma$ and integrating the results of multiple realisations of the stochastic Louvain algorithm.

In consensus Louvain modularity optimisation the same modularity function for weighted, symmetric graphs, is optimised $R=1000$ times resulting in a partition of V, denoted $\boldsymbol{C}_{\gamma,r}$ for each repetition $1\le r\le R$ and $\gamma\in\Gamma$. The consensus matrix

$$F_{x,y}=\frac{1}{R|\Gamma|}\sum_{\gamma\in\Gamma}\sum_{r=1}^{R}\delta_{\boldsymbol{C}_{\gamma,r}}(x,y).$$

is then formed, so that $F_{x,y}\in[0,1]$. The final consensus community is obtained by performing Louvain community detection on the consensus matrix F with resolution $\gamma=1$. The resulting partition is generally robust to repeated applications of the Louvain algorithm. We select the values from $F_{s,s'}$ in increments of 0.1 in the interval $\Gamma=[0.1,1.5]$. This range was chosen as for $\gamma>1.5$, only the trivial partition (consisting of singletons) was produced for all transition graphs tested. This method is based on a simple method for consensus clustering found in [31].

S.3 Stationary Distributions and The Information Rate

The stationary distribution π_s of a transition probability matrix, P of a regular Markov model (also known as a Markov chain), is the limiting long run probability of the occurrence of any one $s\in S$. It is defined as a vector over S satisfying

$$\pi P=\pi.$$

A Markov model is said to be regular if the following positivity condition holds for some positive integer n, that is $P^{(n)}_{s,s'}>0$ for all $s,s'\in S$, where $P^{(n)}$ is the result of taking P to the power of n. In the case that the Markov model is regular, then there exists a unique stationary distribution with

$$IR(P)=\sum_{s,s'\in S}\pi_s P_{s,s'}\log(P_{s,s'}),$$

the information rate of P [53]. Here, we confirmed regularity for all transition matrices used in our models by validating the positivity condition for $n=100$.

S.4 Multiplex Graph Model and Spatiotemporal Consensus Community Detection

We introduce a novel approach to community detection that respects both spatial and temporal connectivity for dynamic state space models. Our new multi-resolution consensus community detection method is overall less vulnerable to the resolution limits of standard community detection methods and more robust to noise than conventional community detection approaches. The HMGM is a multiplex graph model with intralayer edges weighted by the transition probabilities in P and interlayer edges determined by the functional connectivity $W(s)$ (analogous to A in the structural case) between brain regions in state s, [24]. The full spatiotemporal graph model is then $G(M)=(S\times V, A(W,P))$ with full spatiotemporal adjacency matrix

$$A(W,P)_{(s,x),(s',y)}=\begin{pmatrix} P_{s,s'} & \text{if } x=y, s\neq s' \\ W(s)_{x,y} & \text{if } s=s' \\ 0 & \text{otherwise}. \end{pmatrix}.$$

For a partition C of $S\times V$, the multiplex consensus community detection framework repeatedly optimizes the symmetric version of the multiplex modularity function

$$Q^*(\boldsymbol{C})=\sum_{x,y\in V}\sum_{s,s'\in S}\left[W(s)_{x,y}-\gamma\frac{d(s)^x d(s)^y}{\sum_{u,v\in V}W(s)_{u,v}}+\omega P'_{s,s'}\right]\delta_{\boldsymbol{C}}((x,s),(y,s'))$$

where

$$\delta_{\boldsymbol{C}_{\gamma,\omega,r}}((s,x),(s',y))=\begin{cases}1 & \text{if } \exists C\in \boldsymbol{C}_{\gamma,\omega,r} \text{ such that } (s,x),(s',y)\in C \\ 0 & \text{otherwise}\end{cases},$$

is the Dirac function, $P'=(P+P^T)/2$, and $d(s)_x$, $d(s)_y$ are the degrees of x and y in $W(s)$, and ω is the temporal resolution which encourages shared community membership between regions in different states with shared modular organisation.

This definition of Q^* is a modified form of the standard multiplex modularity, which generally cannot take into account weighted or directed intralayer relationships. Models with constant temporal resolution values (ω) have been found not to be able to recover many simple multiplex patterns of

community organisation [74]. The addition of the non-constant inter-state transition probabilities $P'_{s,s'}$ allows for more complex interlayer relationships and thus potentially discovery of more complex spatiotemporal community structures. So far as we are aware, this is the first instance of leveraging the multiplex, temporal information encoded in multivariate HMMs to learn the repertoire of functional associations directly from recordings of functional brain activity. We used Louvain over the alternative Leiden optimisation algorithm due to the availability of the generalised algorithm for Louvain optimisation, which allows for such alternative formulations [75].

For the sake of robustness, as in the purely temporal case we perform consensus community detection. The relatively large multiplex community consensus matrix F^* combines community detection results from across temporal and spatial resolutions. Given $\Gamma^*=\{\gamma_1,..,\gamma_{g^*}\}$, $\Omega=\{\omega_1,..,\omega_g\}$ and $R^*=1000$ replicates, F^* is a $DK\times DK$ matrix with entries given by

$$F^*_{(s,x),(s',y)}=\frac{1}{R^*|\Gamma^*||\Omega|}\sum_{\gamma\in\Gamma^*}\sum_{\omega\in\Omega}\sum_{r=1}^{R^*}\delta_{C_{\gamma,\omega,r}}((s,x),(s',y))$$

The resulting multiplex spatiotemporal consensus communities are then given by applying the standard Louvain algorithm to F^* with resolution $\gamma^*=1$.

We chose the temporal resolution parameters $\Omega=\{0.01,0.1,1\}$ to maximise coverage of the temporal resolution space and spatial resolution $\Gamma^*=\{1,2,3\}$ (motivated by a previous study [26]), with $R=1000$.

S.5 Fractional Occupancy and Permutation Testing

Permutation testing was performed by permuting the $N\times K$ FO matrix within each row, in order to produce surrogate FO distributions for each subject. The new matrices have the same size as the original matrix. This process is repeated $R=10^5$ times to produce surrogate FO matrices. These new surrogate matrices remove any persistent across-subject structure of the FO matrix. We used these surrogates in order to test whether the observed correlation between sink state FO is significant. The two states with the highest mean FO across subjects were chosen for each subject FO matrices and the Pearson correlation between subject FO across states was calculated. The empirical p-value was then calculated from the percentage of surrogate correlations that exceed the actual value.

S.6 Partition Similarity Measures

Two common comparison methods for partitions of a vertex set V, $\boldsymbol{C}_1$ and $\boldsymbol{C}_2$ (possible partitions of V), are the Adjusted Mutual Information (AMI) and Adjusted Rand Index (ARI) [58]. The two measures differ in their approach to determine partition similarity. The AMI is a measure of statistical dependence between the partitions, conceived of as a probability distribution over V. It is calculated as

$$AMI(\boldsymbol{C}_1,\boldsymbol{C}_2)=\frac{MI(\boldsymbol{C}_1,\boldsymbol{C}_2)-E_R[MI(\boldsymbol{C}_1^R,\boldsymbol{C}_2^R)]}{\max_{i\in\{1,2\}}\{H(\boldsymbol{C}_i)\}-E_R[MI(\boldsymbol{C}_1^R,\boldsymbol{C}_2^R)]},$$

where $H(\boldsymbol{C}_1)$ is the entropy of the partion, $MI(\boldsymbol{C}_1, \boldsymbol{C}_2)$ is the mutual information between partitions, $E_R[\cdot]$ is the expectation assuming membership in both partitions were assigned uniformly at random and $\boldsymbol{C}_i^{\boldsymbol{R}}$ is a re-assignment of labels (repartitioning) of $\boldsymbol{C}_i$.

On the other hand, the ARI quantifies the degree of shared membership between the partitions by counting the pairs of shared and unshared members across the partitions [76]. It adjusts for randomness according to some baseline model which is used to determine the expected amount of matching. Most commonly (and in our case), this is the permutation model which assumes that labels were randomly assigned. The ARI follows a similar form to the AMI by subtracting the expected value and dividing by the maximum

$$ARI(\boldsymbol{C}_1, \boldsymbol{C}_2) = \frac{RI(\boldsymbol{C}_1, \boldsymbol{C}_2) - E_R[RI(\boldsymbol{C}_1^{\boldsymbol{R}}, \boldsymbol{C}_2^{\boldsymbol{R}})]}{\max_R\{RI(\boldsymbol{C}_1^{\boldsymbol{R}}, \boldsymbol{C}_2^{\boldsymbol{R}})\} - E_P[MI(\boldsymbol{C}_1^{\boldsymbol{R}}, \boldsymbol{C}_2^{\boldsymbol{R}})]}.$$

S.7 Permutation Testing for Difference in AMI of Functional and Structural Communities

Permutation testing of which condition has the higher AMI when comparing communities to the structural models was carried out by permuting the columns of the two $S \times V$ community membership matrices M_{wake} and M_{anaes} of integer community labels. In this way, the distribution of communities across each state was preserved but the specific membership profile of each region is altered, resulting in surrogate membership matrices M^*_{wake} and M^*_{anaes}. Each row (state) is then compared to the community structure of the structural model, producing surrogate $AMI^*_{wake,i}$ and $AMI^*_{anaes,j}$ for rows i and j respectively. The difference of the median AMI scores for both conditions surrogate AMI scores were then calculated. This was repeated $R = 10^6$ times and the sample was compared to the observed difference in median AMI between models to produce a p-value based on the percentage of values that exceeded the observed score.

S.8 Cortical Grey Matter Density Calculation

Cortical grey matter density was calculated from subject specific T1-weighted MRI images using the FSL-Fast package. FSL-Fast is a method which can be used to estimate fractional levels of occurrence of each of three tissue types in a T1 or high-resolution T2-weighted image: cerebral spinal fluid, white matter and grey matter. FSL-Fast uses a Hidden Markov random field approach to estimate voxel-wise fractional occurrences for each type and outputs the resulting brain maps for each type [77]. To calculate a subject-wise cortical grey matter density we first transformed the whole brain grey matter map to MNI152 standard space and then used a mask consisting of cortical only areas of the Harvard-Oxford cortical-subcortical parcellation to calculate mean grey matter density across the masked region.