

Online Methods

Pallares V*, Insabato A*, Sanjuan A, Mantini D, Kühn S, Deco G, Gilson M

October 23, 2017

1 Description of fMRI datasets

Three datasets acquired at different locations were used in this work:

- Dataset A was acquired for the Day2day project [1] at the Max Planck Institute for Human Development (Berlin, Germany) of 40-50 resting-state sessions recorded from the 6 subjects over 6 months. The uniqueness of this data lies in the capability to have statistically valid evaluation of the session-to-session variability for single subjects.
- Dataset B is publicly available and is part of the Consortium for Reliability and Reproducibility (CoRR) [16]. We used this dataset to generalize the results of the discriminability when increasing the number of subjects (up to 30 subjects).
- Dataset C [9] was recently analyzed using our model-brain dynamic model to extract effective connectivity [5]. We used it to perform a twofold classification with respect to both subjects and conditions (resting-state versus movie viewing).

In this section, we provide details about the acquisition of the blood-oxygen-level dependent (BOLD) signals for the three datasets. It is worth noting that the results of comparing the different measures and estimates of connectivity hold in spite of the fact that the data was collected in different scanners and preprocessed with different pipelines. This strongly support the general scope of our results.

1.1 Dataset A

This dataset has two parts. The first part (A1) is longitudinal and consists of resting-state fMRI sessions from 8 subjects (age 24-32, 6 female). 2 subjects (one male, one female) were not able to continue the study and were discarded. The other 6 subjects underwent scanning between 40 and 50 times along a period of 6 months. The second part of the dataset (A2) was acquired during the same period of time. A total 50 subjects (age 18-32, all female) were scanned during a single fMRI session each using the same MRI sequences.

Participants were free of psychiatric disorder according to a personal interview - mini-international neuropsychiatric interview [12] - and had never suffered from a mental disease. The study was approved by the local ethics committee (Charité University Clinic, Berlin). Participants were instructed to remain with their eyes closed and data acquisition had to be constrained to 5 min per scan due to experimental limitations.

Images were acquired on a 3 T Magnetom Trio MRI scanner system (Siemens Medical Systems, Erlangen, Germany) using a 12-channel radiofrequency head coil. Functional images were collected using a T2*-weighted echo planar imaging (EPI) sequence sensitive to BOLD contrast (TR = 2000 ms, TE = 30 ms, image matrix = 64×64 , FOV = $216 \times 216 \times 129$ mm³, flip angle = 80°, bandwidth=2042 Hz/pixel, voxel size= $3 \times 3 \times 3$ mm³, 36 axial slices using GRAPPA acceleration factor. Structural images were obtained using a three-dimensional T1-weighted magnetization-prepared gradient-echo sequence (MPRAGE) based on the ADNI protocol (www.adni-info.org): repetition time (TR) = 2500 ms; echo time (TE) = 4.77 ms; TI = 1100 ms, acquisition matrix = $256 \times 256 \times 192$ mm³, flip angle = 7°; bandwidth=140 Hz/pixel, $1 \times 1 \times 1$ mm³ voxel size.

1.1.1 Pre-processing

The data was preprocessed using SPM5 (Wellcome Department of Cognitive Neurology, London, UK) and DPARSF/DPABI [15] after discarding the first 10 volumes of each session. It included: slice timing and head motion correction (6 parameters spatial transformation), spatial normalization to the Montreal Neurological Institute (MNI) template, and spatial filtering of 4 mm FWHM. Linear trends were removed from the fMRI time courses before band-pass filtering (0.01-0.08 Hz). The data was parcellated using the automated anatomical labeling (AAL) atlas [14] into 116 regions of interest (ROIs), which includes the whole cortex and the cerebellum.

1.2 Dataset B

This dataset consists of 10 fMRI resting-state sessions acquired from 30 healthy participants every three days for one month [16]. Each session lasted 10 minutes. To minimize head movement, straps and foam pads were used to fix the head snugly during each scan. The participants were instructed to relax and remain still with their eyes open, not to fall asleep, and not to think about anything in particular. The screen presented a fixation point and after the scans, all the participants were interviewed, and none of them reported to have fallen asleep in the scanner. The time of day of MRI acquisition was controlled within participants.

Recording sessions were performed using a GE MR750 3.0 Tesla scanner (GE Medical Systems, Waukesha, WI) at CCB, Hangzhou Normal University. T2-weighted echo-planar imaging (EPI) sequence was performed to obtain resting state fMRI images for 10 minutes using the following setup: TR = 2000 ms, TE = 30 ms, flip angle = 90°, field of view = 220×220 mm², matrix = 64×64 , voxel

size = $3.4 \times 3.4 \times 3.4 \text{ mm}^3$, 43 slices. A T1-weighted fast spoiled gradient echo (FSPGR) was used with the following protocol: TR = 8.1 ms, TE = 3.1 ms, TI = 450 ms, flip angle = 8° , field of view = $256 \times 256 \text{ mm}^2$, matrix = 256×256 , voxel size = $1 \times 1 \times 1 \text{ mm}^3$, 176 sagittal slices) was carried out to acquire a high-resolution anatomical image of the brain structure.

1.2.1 Pre-processing

Dataset B was preprocessed with SPM12 (Wellcome Trust Centre for Neuroimaging, London, UK) and DPARSF/DPABI [15]. The first 5 fMRI volumes were discarded in order to let the BOLD signal reach stability. The pre-processing pipeline included: slice-timing correction, realignment for motion correction, co-registration of the T1 anatomical image to the mean functional image, detrending, regression of 6 movement parameters, 5 principal component analysis (PCA) white matter and CSF Compcorr, and spatial normalization to MNI coordinates. Scrubbing with Power 0.5 and linear interpolation was applied. This data was also parcellated into 116 ROIs using the AAL parcellation [14] and time courses were band-pass filtered between 0.01 and 0.08 Hz, as done with Dataset A. No further global signal regression and spatial smoothing were applied.

1.3 Dataset C

We used a third dataset to study the discrimination between subjects and conditions. In this case, a total of 22 subjects (age 20-31, 15 females) were scanned during rest with eyes opened and natural viewing condition. Volunteers were informed about the experimental procedures, which were approved by the Ethics Committee of the Chieti University, and signed a written informed consent. In the resting state, participants fixated a red target with a diameter of 0.3 visual degrees on a black screen. In the natural viewing condition, subjects watched and listened to 30 minutes of the movie ‘The Good, the Bad and the Ugly’ in a window of 24×10.2 visual degrees. Visual stimuli were projected on a translucent screen using an LCD projector, and viewed by the participants through a mirror tilted by 45 degrees. Auditory stimuli were delivered using MR-compatible headphones. For each subject, 2 and 3 scanning runs of 10 minutes duration were acquired for resting state and natural viewing, respectively.

The BOLD signals were acquired with a 3T MR scanner (Achieva; Philips Medical Systems, Best, The Netherlands) at the Institute for Advanced Biomedical Technologies in Chieti, Italy. The functional images were acquired using T2*-weighted echo-planar images (EPI) with BOLD contrast using SENSE imaging. EPIs comprised of 32 axial slices acquired in ascending order and covering the entire brain with the following protocol: TR = 2000 ms, TE = 3.5 ms, flip angle = 90° , in-plane matrix = 230×230 , voxel size = $2.875 \times 2.875 \times 3.5 \text{ mm}^3$. For each subject, 2 scanning sessions of 10 minutes duration were acquired for resting state and 3 for natural viewing. Each run included 5 dummy volumes, allowing the MRI signal to reach steady state and the subsequent 300 functional

volumes were used for the analysis. Eye position was monitored during scanning using a pupil-corneal reflection system at 120 Hz (Iscan, Burlington, MA, USA). A three-dimensional high-resolution T1-weighted image was acquired for anatomical referencing using an MPRAGE sequence with $TR = 8.1$ ms, $TE = 3.7$ ms, voxel size= $0.938 \times 0.938 \times 1$ mm³ at the end of the scanning session.

1.3.1 Pre-processing

Data were preprocessed using SPM8 (Wellcome Department of Cognitive Neurology, London, UK), including slice-timing and head-motion correction, co-registration between anatomical and mean functional image, and spatial normalization to MNI stereotaxic space (Montreal Neurological Institute, MNI) with a voxel size of $3 \times 3 \times 3$ mm³. Mean BOLD time series were extracted from $N = 66$ regions of interest (ROIs) of the brain atlas used in [8] for each recording session.

The data are available at github.com/MatthieuGilson/EC_estimation. The discarded subjects in the present study are 1, 11 and 19, among the 22 subjects available online (numbered from 0 to 21). The same subjects were discarded in our recent study [5] because of abnormally high BOLD variance.

1.4 Structural connectivity (SC)

For all models, we used a generic matrix of structural connectivity to determine the skeleton of the effective connectivity (i.e., existing connections). For Dataset C, structural connectivity for Dataset C was estimated from diffusion spectrum imaging (DSI) data collected in five healthy right-handed male participants [8]. The gray matter was first parcellated into the $N = 66$ ROIs, using the same low-resolution atlas used for the FC analysis. For each subject, we performed white matter tractography between pairs of cortical areas to estimate a neuro-anatomical connectivity matrix. In our method, the DSI values are only used to determine the skeleton: a binary matrix of structural connectivity (SC) obtained by averaging the matrices over subjects and applying a threshold for the existence of connections. The strengths of individual intracortical connections do not come from DSI values, but are optimized as explained below. For Datasets A and B, the generic SC corresponded to the AAL parcellation with $N = 116$ ROIs [14]. A similar pipeline was used with diffusion tensor imaging.

It is known that both tractography and DTI underestimate inter-hemispheric connections [8]. Homotopic connections between mirrored left and right ROIs are important in order to model whole-cortex BOLD activity [10]. For both SC matrices, we added all possible homotopic connections, which are tuned during the optimization as other existing connections.

2 Connectivity measures and model estimates

Here we provide details about the calculation of the functional and effective connectivity measures introduced in Figure 1.

2.1 Empirical measures of functional connectivity

For each fMRI session, the BOLD time series is denoted by s_i^t for each region $1 \leq i \leq N$ with time indexed by $1 \leq t \leq T$ (time points separated by a TR=2 seconds). The time series were first centered by removing - for each individual ROI i - the session mean $\bar{s}_i = \frac{1}{T} \sum_t s_i^t$. Following [6], the spatiotemporal FC corresponds to covariances calculated as:

$$\begin{aligned}\hat{Q}_{ij}^0 &= \frac{1}{T-2} \sum_{1 \leq t \leq T-1} (s_i^t - \bar{s}_i)(s_j^t - \bar{s}_j) , \\ \hat{Q}_{ij}^1 &= \frac{1}{T-2} \sum_{1 \leq t \leq T-1} (s_i^t - \bar{s}_i)(s_j^{t+1} - \bar{s}_j)\end{aligned}\tag{1}$$

The classical BOLD correlations (corrFC in the main text) correspond to

$$\hat{K}_{ij}^0 = \frac{\hat{Q}_{ij}^0}{\sqrt{\hat{Q}_{ii}^0 \hat{Q}_{jj}^0}} .\tag{2}$$

2.2 Model of cortical dynamics

The model uses two sets of parameters to generate the spatiotemporal FC:

- The network effective connectivity (EC) between the ROIs (cf. Figure 2 in the main text) is denoted by the matrix C in the following equations. Its skeleton is determined by the SC matrix, but not its weight values: weights for absent connections are kept equal to 0 at all times, but weights for existing connections are estimated from FC matrices for each session.
- The local variability is described by the variances (1 per ROI) on the diagonal of the matrix Σ .

The model FC comes from the propagation of the local variability - inputed to every ROI - that propagates via EC, generating network feedback.

Formally, the network dynamics is described by a multivariate Ornstein-Uhlenbeck process, where the activity variable x_i of node i decays exponentially with time constant τ_x - estimated using Eq. (7) - and evolves depending on the activity of other populations:

$$dx_i = \left(\frac{-x_i}{\tau_x} + \sum_{j \neq i} C_{ij} x_j \right) dt + dB_i ,\tag{3}$$

where dB_i is equivalent to white noise with covariance matrix Σ (formally a Wiener process); note that only variances on the diagonal are non zero here.

The simplicity of the model allows for an analytic (feedforward) estimation of the covariances $Q_{ij}^0 = \langle x_i(t) x_j(t) \rangle$ and $Q_{ij}^1 = \langle x_i(t) x_j(t+1) \rangle$, which must reproduce the empirical \hat{Q}_{ij}^0 and \hat{Q}_{ij}^1 , respectively. In practice, we use the two

time shifts 0 and 1 TR, because this gives sufficient information to uniquely infer the network parameters (in the theory). Assuming known network parameters C and Σ , the matrix Q^0 can be calculated by solving the Lyapunov equation (for example using the Bartell-Stewart algorithm):

$$JQ^0 + Q^0J^\dagger + \Sigma = 0 . \quad (4)$$

For Q^1 , it is simply given by

$$Q^1 = Q^0 e^{J^\dagger} . \quad (5)$$

Here J is the Jacobian of the dynamical system and depends on the time constant τ_x and the network effective connectivity:

$$J_{ij} = \frac{-\delta_{ij}}{\tau_x} + C_{ij} , \quad (6)$$

where δ_{ij} is the Kronecker delta and the superscript \dagger denotes the matrix transpose; note also that e^{J^\dagger} is a matrix exponential.

2.3 Parameter estimation procedure

For each individual and session, we calculate the time constant τ_x associated with the exponential decay of the autocovariance averaged over all ROIs:

$$\tau_x = \frac{1}{N} \sum_{1 \leq i \leq N} \frac{1}{\log(\hat{Q}_{ii}^0) - \log(\hat{Q}_{ii}^1)} \quad (7)$$

This is used to “calibrate” the model, before its optimization.

In order to invert the model (i.e., for the model FC to reproduce the experimental FC), we iteratively tune the parameters to reduce the model error defined as

$$E = \frac{1}{2} \frac{\sum_{i,j} (\Delta Q_{ij}^0)^2}{\sum_{i,j} (\hat{Q}_{ij}^0)^2} + \frac{1}{2} \frac{\sum_{i,j} (\Delta Q_{ij}^1)^2}{\sum_{i,j} (\hat{Q}_{ij}^1)^2} . \quad (8)$$

Here each term - for FC0 and FC1 - is the matrix distance between the model and the data observables, normalized by the norm of the latter; for compactness, we have defined the difference matrices are $\Delta Q^0 = \hat{Q}^0 - Q^0$ and $\Delta Q^1 = \hat{Q}^1 - Q^1$.

The idea behind the tuning algorithm is to start from zero connectivity $C = 0$ and homogeneous Σ , then calculate the model Q^0 and Q^1 using the desired Jacobian update given by

$$\Delta J^\dagger = (Q^0)^{-1} [\Delta Q^0 + \Delta Q^1 e^{J^\dagger}] , \quad (9)$$

which decreases the model error E at each optimization step, similar to a gradient descent. The best fit corresponds to the minimum of E . Finally, the connectivity update is

$$\Delta C_{ij} = \eta_C \Delta J_{ij} \quad (10)$$

for existing connections only; other weights are forced at 0. We also impose non-negativity for the EC values during the optimization. To take properly the effect of cross-correlated inputs into account, we use the Σ as in [5]:

$$\Delta\Sigma = -\eta_{\Sigma}(J\Delta Q^0 + \Delta Q^0 J^{\dagger}) . \quad (11)$$

As with C for non-existing connections, off-diagonal elements of Σ are kept equal to 0 at all times.

In numerical simulations, we use $\eta_C = 0.0005$ and $\eta_{\Sigma} = 0.05$. Further details about the derivation of the optimization updates are provided in [6].

The optimization code is available with Dataset C at github.com/MatthieuGilson/EC_estimation.

2.4 Comparison of the model to state-of-the-art dynamic models to interpret fMRI data

Compared to dynamic causal modeling [3, 2], our model makes the simpler assumption of linearity for the local dynamics. Doing so, it ignores an explicit modeling of the mapping between the neuronal activity and the BOLD signals [13]. Moreover, it uses a simple model of local variability (Wiener process) related to Σ to generate FC than recent development of DCM for resting state [4]. In exchange for this simplicity, we obtain a very efficient estimation procedure for networks of about 100 ROIs with 30% density, yielding 3000 EC parameters. It is also worth noting that the objective function for our framework are the BOLD covariances, which are canonically related to the BOLD cross-spectrum used in the DCM for resting state [4].

3 Analysis and classification of vectorized EC and corrFC

In this section, we provide details about the analysis of the corrFC and estimated EC matrices that are compared across sessions for subject and condition identification. As illustrated in Figure 1C, the connectivity measure for each session k was transformed into a vector v^k by extracting the lower triangle for corrFC, and by applying the SC mask for EC. Following the literature in machine learning, we refer to a connectivity measure for a single session as a sample. The size of the samples v^k is $p = 6670$ for an FC session and $p = 4056$ for an EC session with Datasets A and B (corresponding to $N = 116$ ROIs). For Dataset C, we used only EC with $p = 1114$ vector elements (for $N = 66$). *Note that we use a slightly different indexing in this section compared to the previous one: in the following i refers to a link in the vector $v^k = (v_i^k)$, similarly to the pair (i, j) for a matrix element C_{ij} before.*

3.1 Similarity between sessions

We used Pearson correlation coefficient (PCC) as a measure of similarity, both within and between subjects. For a pair of sessions k and l as reported in Figure 2A and B, this is:

$$S^{kl} = \text{PCC}(v^k, v^l) \quad (12)$$

The distribution of within-subject similarity (WSS) in Figure 2B was obtained by using all pairs of vectors v^k and v^l with $k \neq l$ from the same subject, in both Dataset A1 and B. To compute the distribution of between-subject similarity (BSS), all possible combinations of vector pairs v^k and v^l from distinct subjects were used.

3.2 Dimensionality analysis

To study visually how the variability of the data is spread over in the space with high dimension p , we applied principal component analysis (PCA) to extract the main dimensions (principal components, or PCs) that capture the largest portion of the data variance. In Figure 2C that compares corrFC and EC, PCA was applied to the whole Dataset A1 (6 subjects, 40-50 sessions per subject) and the first 6 PCs were retained. Each panel corresponds to PC1 to PC3 on the one hand, and PC4 to PC6 on the other hand.

3.3 Silhouette coefficient

The silhouette coefficient [11] is defined for each vectors $v^{k,s}$ with indices k for the session and s for the subject. Here, each subject s is a cluster and the similarity in Eq. (12) is taken as the metric, but the indices k and l are replaced by doublets of the type (k, s) here. For a given sample k, s , we have the average similarity within his own cluster s defined as

$$a^{k,s} = \text{mean}_{l \neq k}(S^{(k,s)(l,s)}) , \quad (13)$$

and the maximum - over all other clusters - of the same average similarity, but with elements from another cluster:

$$b^{k,s} = \max_{s'}[\text{mean}_{l'}(S^{(k,s)(l',s')})] . \quad (14)$$

The silhouette is then given by the following contrast between the cohesion of the element within its cluster ($a^{k,s}$) and the separation from other clusters ($b^{k,s}$):

$$\sigma^{k,s} = \frac{b^{k,s} - a^{k,s}}{\max(a^{k,s}, b^{k,s})} . \quad (15)$$

Values of silhouette range to 1 for fully separated clusters to -1 for fully overlapping clusters. These values correspond to the violin plots without PCA in Figure 2E.

In Figure 2D and E, the silhouette coefficient were computed for each point in these clouds in a 6-dimensional PC space for Dataset A1. The reason for choosing the first 6 PCs is because the mean silhouette coefficient of EC data reaches a maximum, before decreasing (Figure S3, left panel). The same method was applied to the Dataset B (30 subjects, 10 sessions per subject), for which the retained maximum was 30 PCs (Figure S3, right panel).

3.4 Within-session z-scoring

For the classification in Figures 3 and 4, the values of the EC and corrFC links were z-scored within each session, using the mean and standard deviation of the corresponding vectorized connectivity measure v^k in Figure 1C:

$$\hat{v}_i^k = \frac{v_i^k - \text{mean}_i(v_i^k)}{\text{std}_i(v_i^k)}, \quad (16)$$

where the vector elements are indexed by i , corresponding to a link in the EC or corrFC matrix. The z-scored vectors \hat{v}^k are the inputs of the classifiers, which means that the classification relies on the ranking of the vectorized elements v^k rather than the absolute values of their elements. This is important to understand our claim in the Discussion about where the discriminative information is: the repartition of the weak/strong EC weights across the brain is different across subjects, which is picked up by the algorithms.

3.5 Classification of sessions to attribute them to subjects

Figure 3A shows the classification procedure applied to identify subjects using connectivity measures and estimates. First, a fixed number of sessions per subject were selected from each data set, corrFC and EC matrices. These matrices were vectorized (see Figure 1C) and individually z-scored using Eq. (16), using the mean and standard deviation of each v^k . Then, the corresponding classifier - 1-nearest-neighbor (1NN) or multilinear logistic regression (MLR) - was trained using two different approaches. In the main text, results are presented without applying PCA as a preprocessing step. This is motivated because PCA does not significantly improve the classification performance for the MLR. Results about classifier with PCA are discussed in Supplementary Material (Section S4).

In all cases, the accuracy of a classifier is evaluated by its prediction of samples in the test set, as a cross-validation. We used Dataset A1 to study the effect of increasing the samples in the train set, and Dataset B for increasing the number of subjects. The curves in Figure 3B and C were obtained after iterating over different sessions and subjects 100 times with this cross-validation (mean and standard deviation are plotted). The same method was also applied to Dataset C by training two MLR classifiers, one for subjects and one for conditions. Both classifiers are available in the scikit-learn package (<http://scikit-learn.org>, python language).

3.5.1 1NN classifier

A kNN classifier is a technique that assigns to a new sample the class to which belong the majority of its k closest neighbors. In our case, we use $k = 1$ with a single nearest neighbor. Moreover, we use the PCC-based similarity measure in Eq. (12) as the metric to evaluate the inverse distance between two samples (here sessions). Like with clustering algorithm in general, closest samples (i.e., most similar sessions) are grouped together. Because the PCC similarity is not linear, it can be considered as a non-linear classifier, in comparison to a linear classifier such as a perceptron or a MLR. In practice, the database has an equal number of sessions per subject - either vectorized and z-scored corrFC or EC - ranging from 1 to 40 for Dataset A1 in Figure 2B. The identity of the each target session k from the test set is predicted by the identity of the most similar session from the test set $(S^{k,1}, S^{k,2}, \dots, S^{k,D})$ with D the size of the database, as illustrated in Figure S4 for 1 session per subject as database (and corresponding to the results presented in Figure 3B and C).

3.5.2 MLR classifier

The MLR classifier is a classical tool in machine learning. The parameters (or regressors) of the model are adjusted in order to predict the probabilities of new samples of belonging to each category or class. It relies on the following logistic function that relates to the probability for the vectorized connectivity measure v^k (with elements indexed by i) for session k to be in the subject class s :

$$\Pr(v^k \in s) = \phi\left(\sum_i w_i^s v_i^k\right), \quad (17)$$

where ϕ is the sigmoid function (ranging from 0 to 1). The training is performed by a regression to find the classification weight w_i^s such that $\Pr(v^k \in s)$ discriminates the class s against the last subject s' . A refinement that does not appear in Eq. (17) is that there is an extra weight to correct for the possibly non-zero mean of the samples v^k . Note also that for there are $M - 1$ regressors for M subjects, such that the weights are well constrained. In practice, we used train sets with equal numbers of sessions per subject.

3.5.3 Preprocessing using PCA

PCA is a preprocessing step commonly used in machine learning to remove noise while keeping the dimensions that capture most of the variability of the data. This implies that the largest part of the data variances captures the relevant information for the classification. After applying PCA, the original high-dimensional z-scored vectors \hat{v}^k from Eq. (16) are projected into a space of lower dimension determined by a number of PCs. The performance of the classification with PCA thus increases with the number of PCs until saturation (Figure S6), which indicates the point when subsequent PCs contain redundant or irrelevant information for the classification.

3.6 Extraction of discriminative support networks

We examined which links strongly contributed to the classification, in a similar fashion to PCs as mentioned above. The motivation was that individual links might be mixed in PCs and appear redundantly in several PCs that significantly contributed to the classification (Figures S7 and S8). Therefore, we evaluated the contribution of individual links that supported the classification, forming networks of most discriminative elements in EC (as represented in Figures 3D and 4C). In order to extract these support networks, we employed a commonly used method in machine learning, recursive feature elimination (RFE), to rank the links - taken as features - according to their relevance for the classification [7].

For each application of the RFE algorithm, the train set was composed of 90% randomly chosen samples (to capture the full variability of the data) and test set of the 10% remaining samples. After fitting the MLR classifier, RFE removes the link with the smallest classification weight w_i^s in the MLR formula in Eq. (17), which measures the contribution of the link to the classification. The removing procedure is repeated recursively on the shrinking subset of links until only one is left. This gives a ranking for the links according to their relevance for the classification. We then evaluated the accuracy of MLR on the test set when increasing the number of features following the order given by the RFE ranking. This training and testing procedure was repeated 100 times with different train and test sets each time. We selected the number of features for which the mean test set accuracy was maximum. In order to find the maximum we chose the number of features for which the numerical derivative of the mean was less than 10^{-6} . In order to reduce the impact of fluctuations due to the random selection of samples, we smoothed the curve of means with a rolling average of width 2 features. Since the accuracy is expected to increase initially as a function of the number of features and then either saturate or decrease, this method allows for finding the number of features for which the classifier performance is maximum or adding more features has no practical benefit.

In contrast, kNN cannot be easily used for RFE since it does not estimate weights associated to links. Therefore, for kNN to be used with RFE, one needs to put the model in a wrapper to compare the effect of removing each combination of links on the performance. However, given the high amount of features of our setting, wrappers cannot be evaluated on all subsets of features ($\sim 10^{300}$ tests would be required for 1000 features). Wrappers may rely on approximate greedy algorithms, for example, eliminating the feature that scores worst. It is well known that greedy algorithms might produce inappropriate solutions if the problem does not have optimal substructure. In addition the computation time almost scales as p^2 with p the number of links, while for RFE it is linear in p .

3.7 Software tools

The computer code for the model optimization are available online at github.com/MatthieuGilson/EC_estimation; it relies on the numpy and scipy libraries in python.

Classification and feature extraction analyses were performed in python using the scikit-learn library for machine-learning routines (<http://scikit-learn.org>). The code for the classification will be made available on github.com/MatthieuGilson/EC_estimation.

Connectivity measures and estimates, as well as similarity analyses were performed in MATLAB 2016 (TM). Network plots in Figure 5 of the main text were done in Gephi 0.9.1 (<http://gephi.org>).

References

- [1] E. Filevich, N. Lisofsky, M. Becker, O. Butler, M. Lochstet, J. Martensson, E. Wenger, U. Lindenberger, and S. Kühn. Day2day: investigating daily variability of magnetic resonance imaging measures over half a year. *BMC Neurosci*, 18:65, 2017.
- [2] K. Friston. Functional and effective connectivity: A review. *Brain Connect*, 1:8, 2011.
- [3] K. J. Friston, L. Harrison, and W. Penny. Dynamic causal modelling. *Neuroimage*, 19:1273–1302, 2003.
- [4] K. J. Friston, J. Kahan, B. Biswal, and A. Razi. A DCM for resting state fMRI. *Neuroimage*, 94:396–407, 2014.
- [5] M. Gilson, G. Deco, K. Friston, P. Hagmann, D. Mantini, V. Betti, G. L. Romani, and M. Corbetta. Effective connectivity inferred from fmri transition dynamics during movie viewing points to a balanced reconfiguration of cortical interactions. *NeuroImage*, 2017.
- [6] M. Gilson, R. Moreno-Bote, A. Ponce-Alvarez, P. Ritter, and G. Deco. Estimation of directed effective connectivity from fMRI functional connectivity hints at asymmetries of cortical connectome. *PLoS Comput Biol*, 12:e1004762, 2016.
- [7] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik. Gene selection for cancer classification using support vector machines. *Machine Learning*, 46(1):389–422, 2002.
- [8] P. Hagmann, L. Cammoun, X. Gigandet, R. Meuli, C. J. Honey, V. J. Wedeen, and O. Sporns. Mapping the structural core of human cerebral cortex. *PLoS Biol*, 6:e159, 2008.

- [9] D. Mantini, U. Hasson, V. Betti, M. G. Perrucci, G. L. Romani, M. Corbetta, G. A. Orban, and W. Vanduffel. Interspecies activity correlations reveal functional correspondence between monkey and human brain areas. *Nat Methods*, 9:277–282, 2012.
- [10] A. Messé, D. Rudrauf, H. Benali, and G. Marrelec. Relating structure and function in the human brain: relative contributions of anatomy, stationary dynamics, and non-stationarities. *PLoS Comput Biol*, 10:e1003530, 2014.
- [11] P. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65, 1987.
- [12] D. V. Sheehan, Y. Lecrubier, K. H. Sheehan, P. Amorim, J. Janavs, E. Weiller, T. Hergueta, R. Baker, and G. C. Dunbar. The mini-international neuropsychiatric interview (m.i.n.i.): the development and validation of a structured diagnostic psychiatric interview for dsm-iv and icd-10. *J Clin Psychiatry*, 59:22–33, 1998.
- [13] K. Stephan, L. Harrison, W. Penny, and K. Friston. Biophysical models of fMRI responses. *Curr Opin Neurol*, 14:629–635, 2004.
- [14] N. Tzourio-Mazoyer, B. Landeau, D. Papathanassiou, F. Crivello, O. Etard, N. Delcroix, B. Mazoyer, and M. Joliot. Automated anatomical labeling of activations in spm using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage*, 15:273–289, 2002.
- [15] C.-G. Yan, X.-D. Wang, X.-N. Zuo, and Y.-F. Zang. Dpabi: Data processing & analysis for (resting-state) brain imaging. *Neuroinformatics*, 14:339–351, 2016.
- [16] X.-N. Zuo, J. S. Anderson, P. Bellec, R. M. Birn, B. B. Biswal, J. Blautzik, J. C. S. Breitner, R. L. Buckner, V. D. Calhoun, F. X. Castellanos, A. Chen, B. Chen, J. Chen, X. Chen, S. J. Colcombe, W. Courtney, R. C. Craddock, A. Di Martino, H.-M. Dong, X. Fu, Q. Gong, K. J. Gorgolewski, Y. Han, Y. He, Y. He, E. Ho, A. Holmes, X.-H. Hou, J. Huckins, T. Jiang, Y. Jiang, W. Kelley, C. Kelly, M. King, S. M. LaConte, J. E. Lainhart, X. Lei, H.-J. Li, K. Li, K. Li, Q. Lin, D. Liu, J. Liu, X. Liu, Y. Liu, G. Lu, J. Lu, B. Luna, J. Luo, D. Lurie, Y. Mao, D. S. Margulies, A. R. Mayer, T. Meindl, M. E. Meyerand, W. Nan, J. A. Nielsen, D. O’Connor, D. Paulsen, V. Prabhakaran, Z. Qi, J. Qiu, C. Shao, Z. Shehzad, W. Tang, A. Villringer, H. Wang, K. Wang, D. Wei, G.-X. Wei, X.-C. Weng, X. Wu, T. Xu, N. Yang, Z. Yang, Y.-F. Zang, L. Zhang, Q. Zhang, Z. Zhang, Z. Zhang, K. Zhao, Z. Zhen, Y. Zhou, X.-T. Zhu, and M. P. Milham. An open science resource for establishing reliability and reproducibility in functional connectomics. *Sci Data*, 1:140049, 2014.