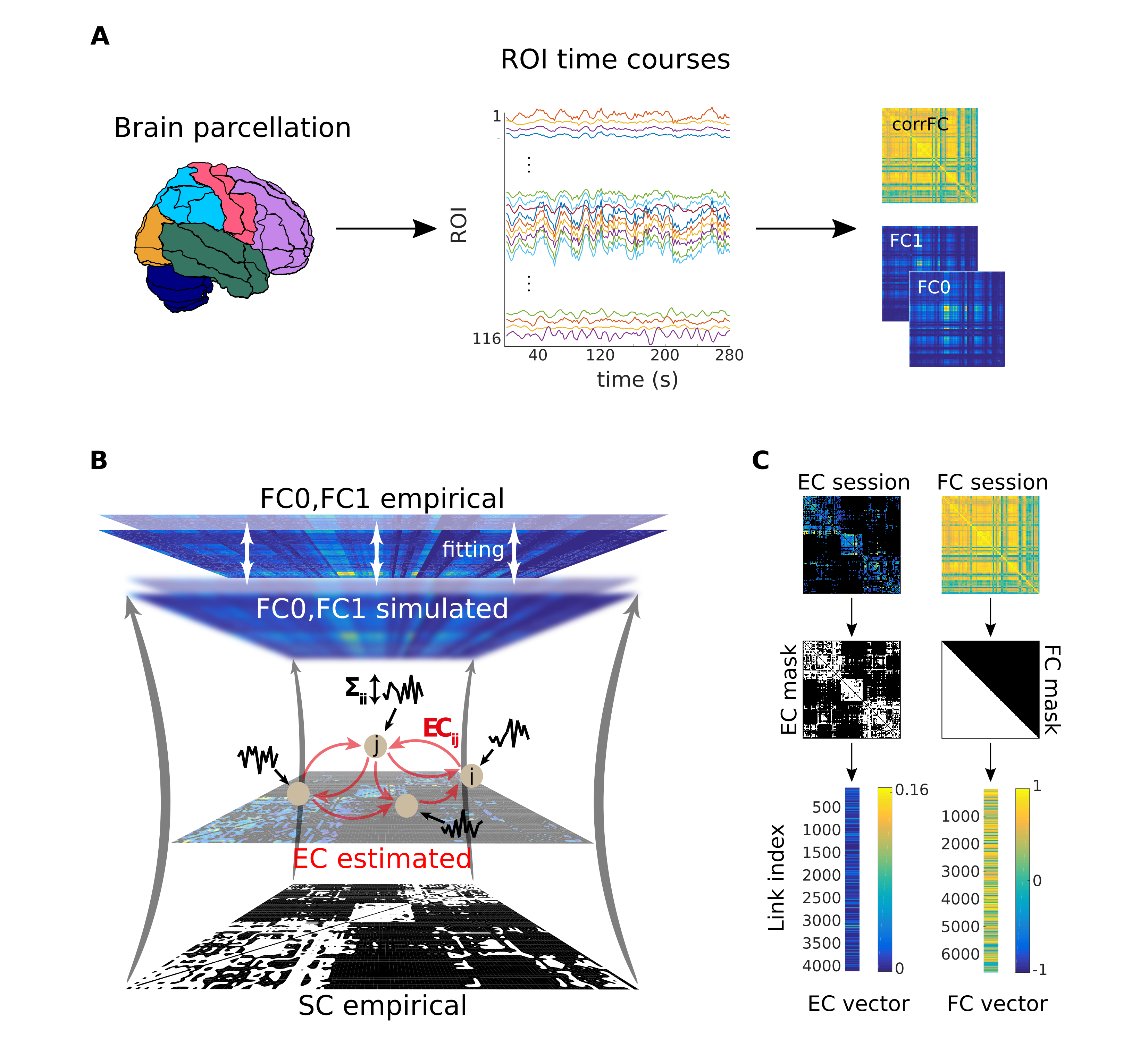
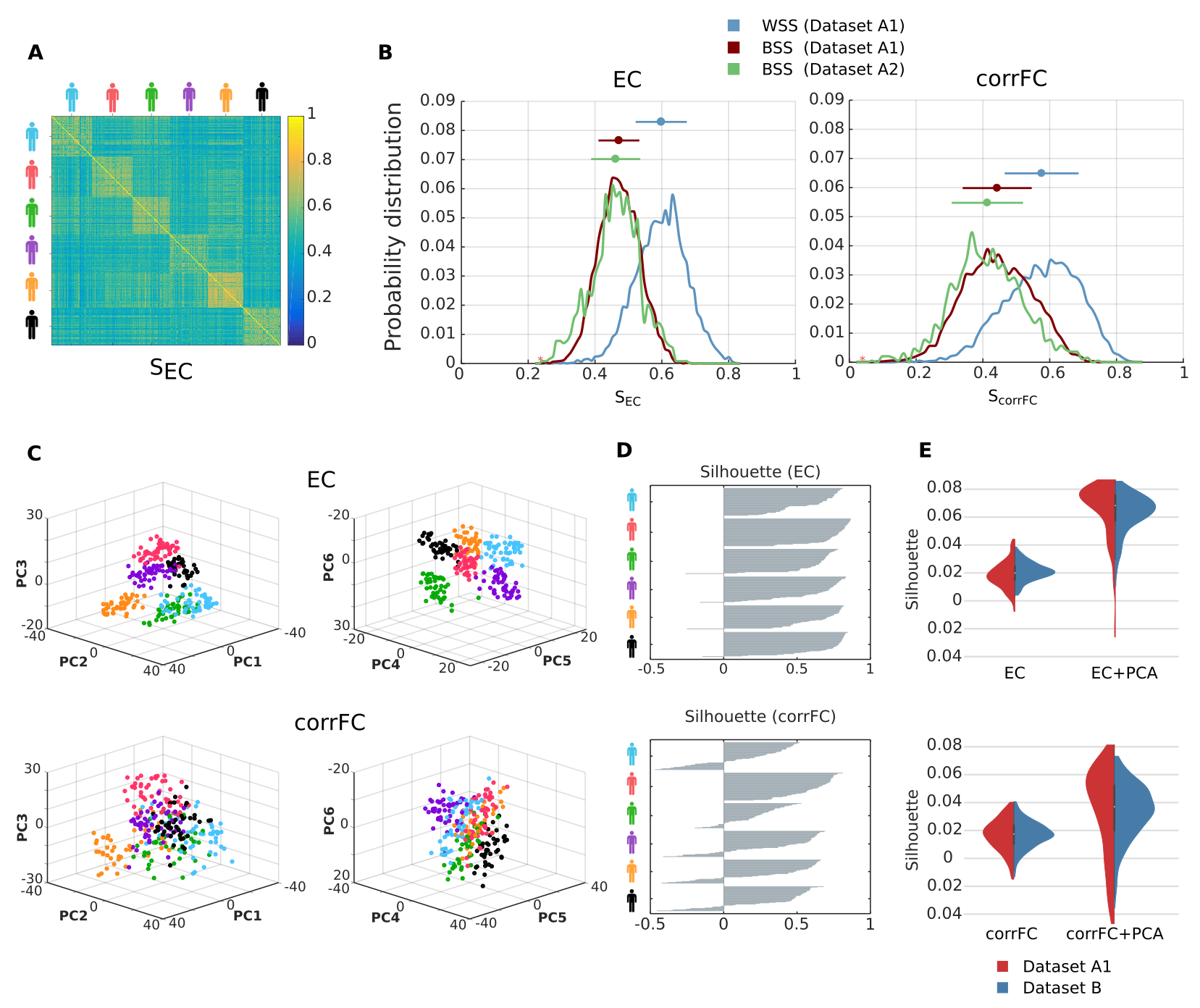
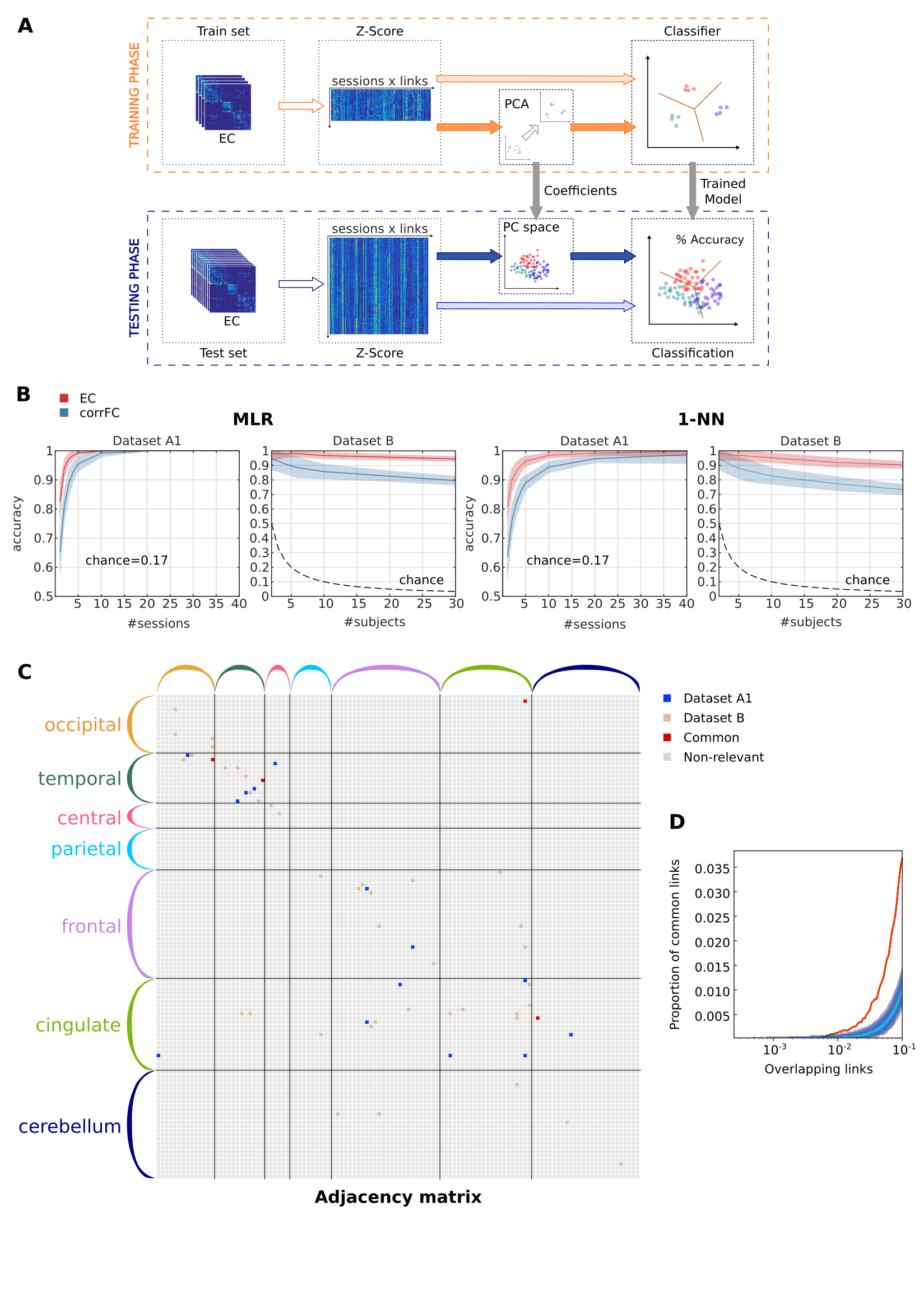
****

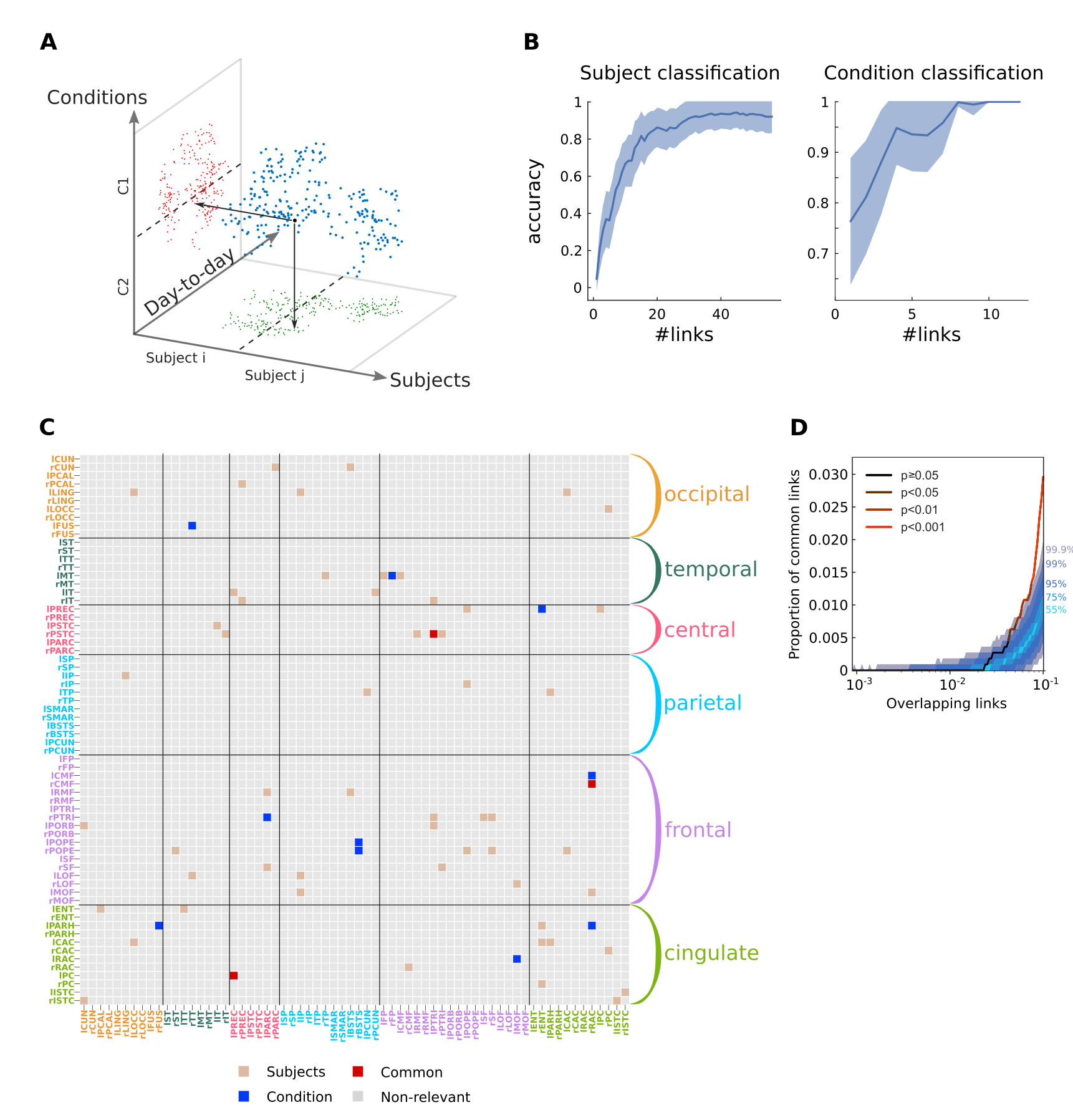
**Figure 1: Workflow for the calculation of the connectivity measures from fMRI measurements. A)** After a standard pre-processing pipeline, a parcellation covering the whole-brain is applied to extract BOLD time series: 116 ROIs for AAL, and 66 ROIs for the Hagmann parcellation; each color represents an anatomical subsystem of several ROIs. Here we consider two versions of functional connectivity: the classical corrFC corresponding to the Pearson correlation between pairs of time series; the matrices FC0 and FC1 correspond to the covariances without and with time shift, respectively (see Methods for the detail of the calculation). **B)** Whole-brain network model to interpret fMRI data. The local fluctuating activity (where Σii is the variance of the input to region i) propagates via the recurrent EC to generate the correlation patterns at the network level. Structural connectivity (SC, bottom) obtained using diffusion-weight imaging (DWI) determines the skeleton of EC. The estimation procedure iteratively tunes EC and Σ such that the model best reproduces the empirical FC0 and FC1. **C)** Each corrFC matrix has 116x116 elements, is symmetric and has all diagonal elements equal to 1, so only the 6670 independent links are retained for identification/classification (lower triangle). Likewise, the EC matrix has 4056 non-zero elements that are used in the classification (density of 30%).

****

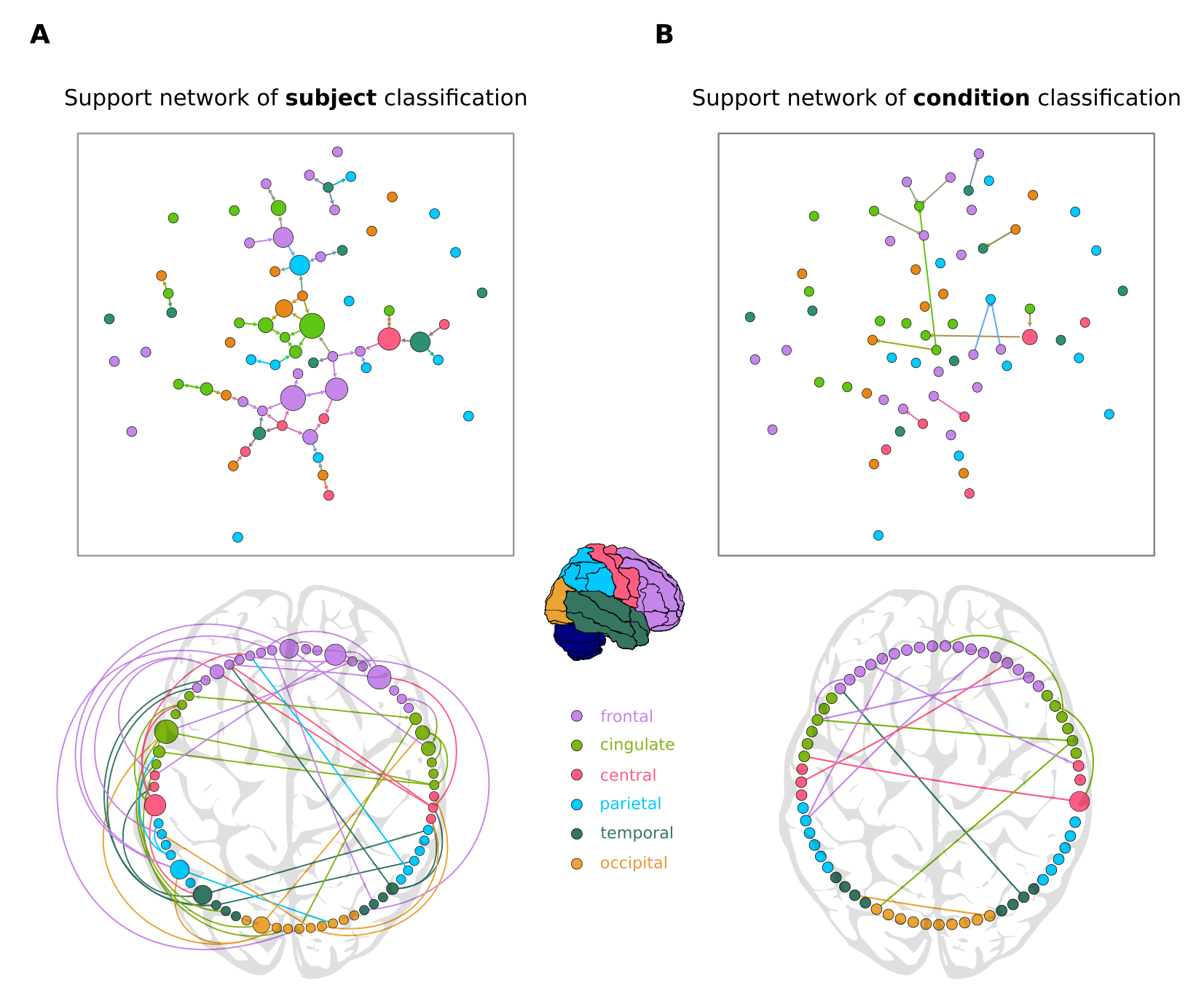
**Figure 2: Within- and between-subject similarity (WSS and BSS, respectively) for EC and corrFC. A)** Matrix of Pearson correlation coefficient (PCC) values for EC between all pairs of sessions from Dataset A1. ECs from two sessions are transformed into two vectors, from which similarity is calculated (SEC); the sessions are grouped by subjects, as indicated by the colored symbols. **B)** Distributions of WSS (blue) and BSS (red) values for Datasets A1 – corresponding to diagonal and off-diagonal in panel A, respectively – and of BSS (green) for Dataset A2. Comparison between EC (left panel) and corrFC (right panel). Dots and lines on the top represent means and standard deviations, respectively, indicating smaller overlap between WSS and BSS for EC. **C)** Visualization of the sessions of Dataset A1 in the space of the first 6 PCs obtained from PCA for EC (top row) and corrFC (bottom row). Each point corresponds to a scan session and each color to one of the 6 subjects, as in panel A. **D)** Silhouette coefficients of each session in panel C (see main text and Eq. 16 in Methods for further details). **E)** Distribution of the silhouette coefficients for EC (top panel) and corrFC (bottom panel): comparison between the original link space (left) and the space of the first 6 PCs (right). Both Datasets A1 (6 subjects, in red) and B (30 subjects, in blue) are represented by the violin plots. Note the larger silhouette coefficients for EC than corrFC.

****

**Figure 3: Classification performance of EC and FC depending on sessions and subjects. A)** Classification pipeline used to assess the performance on subject identification. The full set of connectivity measures (here EC) over all fMRI sessions was split into two groups: a train set and a test set. We applied a z-score standardization for the elements of each session matrix (see Eq. 12 in Methods). We trained the classifier – with or without applying PCA – and evaluated the classification accuracy on the test set. **B)** Performance of multinomial logistic regression (MLR, left panel) and 1-nearest-neighbor (1-NN, right panel) classifiers when increasing the number of sessions per subject used as training set with Dataset A1 (6 subjects with about 50 sessions each). The mean (solid curve) and standard deviation (colored area) were calculated for 100 repetitions with cross-validation. **C)** Same as B when varying the number of subjects using Dataset B, using a single training session per subject (leaving 9 sessions to classify for each subject). **D)** Extracted links that contribute to the classification with both datasets, obtained using feature recursive elimination (FRE). **E)** Overlap between the two signatures for Datasets A1 and B when increasing the number of selected links (indicated on the x-axis, in the order of RFE rankings). The curve represents the amount of common links in the data. The shaded area represents different quantiles of the surrogate distribution of common links under the null-hypothesis of random rankings. The red color of the curve indicates a p-value < 0.001 when considering more than 1% of the total links (namely 40 links).

****

**Figure 4: Two-fold discrimination between subjects and conditions using EC.   
A)** Desired scheme for the twofold classification where each session (blue dots) is “projected” on two subspaces, one for subjects (green) and one for conditions (red). On each subspace, detection can be performed efficiently. Depending on the subspace orthogonality, the two signatures have more or less overlap. **B)** Performance of the classification for 19 subjects and 2 conditions using Dataset C: the accuracy stabilizes around 90% around 40 links for the subjects and at 100% after 10 links for conditions. Note the distinct scales for the y-axis. **C)** Signatures of the most discriminative EC links for the twofold classification in B: 54 links for subject classification in brown, 10 for condition classification in blue, 3 common links in red. RFE was applied to find the most relevant EC connections between ROI for both subject and condition discriminations, until reaching maximum performance. **D)** Proportion of common links between the subject and condition signatures when varying the number of selected links (indicated on the x-axis, in the order of RFE rankings). The color coding is the same as in Figure 3E: the two signatures are significantly different, i.e., with a number of common links corresponding to the null hypothesis with p-value ≥ 0.05 (cf. legend) up to 4% of the total links.



**Figure 5: Support networks of subject and condition classification. A)** The top graph plot represents the 54 most discriminative EC links supporting the classification of subjects (same as in Figure 3C). The size of each node represents its betweenness centrality in the extracted network. The most central regions are located mainly in the frontal and cingulate cortices. The bottom circular plot shows the asymmetry and lateralization of the network, with more links located in the left hemisphere. Links that are inside the circle correspond to contralateral connections, while links outside the circle are ipsilateral connections. **B)** Similar graph and circular plots to those in A for the 14 links supporting the classification between the two conditions (resting versus movie). Fewer links are required to reach high accuracy in the condition discrimination: they form a network with many disjoint components and are mainly contralateral, in comparison to the subject support network.