Time scale separation of information processing between sensory and associative regions

Sorrentino P*[1,2,q], Rabuffo G*[1], Rucco R[2,3], Baselice F[4], Troisi Lopez E[3,5], Liparoti M[3,5], Quarantelli M[6], Sorrentino G[2,3,4]+, Bernard C[1]+, Jirsa V[1,q]+

1. Institut de Neurosciences des Systèmes, Aix-Marseille University, Marseille, France
2. Institute of Applied Sciences and Intelligent Systems, National Research Council, Pozzuoli, Italy
3. Department of Motor Sciences and Wellness, Parthenope University of Naples, Naples, Italy
4. Department of Engineering, Parthenope University of Naples, Naples, Italy
5. Institute for Diagnosis and Cure Hermitage Capodimonte, Naples, Italy
6. Biostructure and Bioimaging Institute, National Research Council, Naples, Italy

*** contributed equally**

q = corresponding authors

+ co-senior authors

Abstract

Stimulus perception is assumed to involve the (fast) detection of sensory inputs and their (slower) integration. The capacity of the brain to quickly adapt, at all times, to unexpected stimuli suggests that the interplay between the slow and fast processes happens at short timescales. We hypothesised that, even during resting-state, the flow of information across the brain regions should evolve quickly, but not homogeneously in time. Here we used high temporal-resolution Magnetoencephalography (MEG) signals to estimate the persistence of the information in functional links across the brain. We show that short- and long-lasting retention of the information, entailing different speeds in the update rate, naturally split the brain into two anatomically distinct subnetworks. The "fast updating network" (FUN) is localized in the regions that typically belong to the dorsal and ventral streams during perceptive tasks, while the "slow updating network" (SUN) hinges classically associative areas. Finally, we show that only a subset of the brain regions, which we name the multi-storage core (MSC), belongs to both subnetworks. The MSC is hypothesized to play a role in the communication between the (otherwise) segregated subnetworks.

Significance statement

The human brain constantly scans the environment in search of relevant incoming stimuli, and appropriately reconfigures its large-scale activation according to environmental requests. The functional organization substanding these bottom-up and top-down processes, however, is not understood. Studying the speed of information processing between brain regions during resting state, we show the existence of two spatially segregated subnetworks processing information at fast- and slow- rates. Notably, these networks involve the regions that typically belong to the perception stream and the associative regions, respectively. Therefore, we provide evidence that, regardless of the presence of a stimulus, the bottom-up and top-down perceptive pathways are inherent to the resting state dynamics.

Introduction

Efficient communication and long-range coherent activity among brain areas is a prerequisite for the correct functioning of the brain (1, 2). One way to synthetically represent the complex relationships among brain areas is by conceptualizing the brain as a network, whereby brain areas are represented as nodes, and structural connections or functional communication among them are represented as edges (3). This approach led to the identification of a number of invariant features in the large-scale organization of the brain. At the structural level, small-world (4) and scale-free (5) architectures seem to be found consistently in various human cohorts (6) . Theoretical (7–9) , computational (10), and empirical works (11) have shown that such features maximize the efficiency of the network, in the sense that they allow optimal communication while also complying to spatial and energetic constraints. From the functional point of view, large-scale patterns of interactions among areas are non-stationary, and reorganize in time (12–15). It has been hypothesized that the correct unfolding of the dynamics during resting state is tightly linked to the constant readiness to stimuli that vary across multiple time-scales (1, 2). Accordingly, the loss of efficient dynamics has been linked to brain pathology and clinical symptoms (16). To achieve this readiness, the brain has to quickly acquire information from the environment, integrate it and interpret it. However, the spatiotemporal fabric underpinning such capability is yet to be understood. While multiple mechanisms of communication have been proposed (among others, communication through resonance (17), coherence (18), oscillations (19), synfire chains (20), they all require fine-tuned timing. From a theoretical perspective, both the topological organization of long-range structural connections and their myelination (which reflects itself in the time-delays) could contribute to shape up communication

channels with intrinsic operational times, resulting in a spatio-temporally nested activity (21, 22). In this scenario, the temporal unfolding of pairwise interactions might not be homogeneous throughout the brain network. In fact, some tasks might require quick interactions that are constantly reupdated, while other functions might instead require integration of more information, and consequently a longer time to unfold. In this line of thinking, it has been shown in the Macaque that areas that are hierarchically lower in information processing interact at higher velocities as compared to associative areas (which integrate information) (23). Such differences in the integration process should reflect in the duration of effective interactions between areas. In this paper, we hypothesize that, if this is the case, then classifying the edges upon retention of information should naturally lead to topographic clustering separating low (retaining information only shortly) and high (retaining information longer - as to integrate it) hierarchical structures. Traditionally, time-resolved analyses have been performed using EEG. However, EEG is difficult to analyze and interpret because the electric fields are distorted and filtered between their sources and the recording sites. Hence, to test our hypothesis we have used source-reconstructed magnetoencephalographic (MEG) data from a cohort of 58 healthy subjects. MEG records the (weak) magnetic fields induced by post-synaptic electric activity. Given that the magnetic field is (almost) undistorted by the skull and meningeal sheets, MEG allows for an optimal trade-off between spatial and temporal resolution (24, 25). The temporal unfolding of the interactions between region signals was estimated based on the pairwise co-activations. This procedure defines a characteristic time series for each edge of the brain network (26). The Auto Mutual Information (AMI) is used to estimate the time up to which the functional information is retained within each edge. Finally, we further hypothesized that specific regions should mediate the integration of groups of edges operating at different

processing speeds. To test this, we built a multiplex network where each layer is filled by either fast or slow decaying edges, so as to pinpoint the set or regions that are relevant to both and, likely, mediate the temporal integration of the whole brain.

Results

Functional connections store information differently

Using source-reconstructed magnetoencephalographic (MEG) data from a cohort of 58 healthy young subjects, we analyzed the temporal unfold of the pairwise correlations between the MEG source signals reconstructed in any couple of brain regions. Namely, for each functional edge in the brain network we extracted a time series which keeps track of the co-fluctuations of the MEG signals at the extremal nodes (Fig.1.A) (See Materials and Methods) . Information theoretic tools were used in order to estimate the decay in time of the functional information within each edge. In particular, the Auto Mutual Information (AMI) was used to measure the amount of statistical dependence between any co-activation time series and its time-delayed copy (Mackay, 1995). Repeating this operation for several delays, a profile of functional information decay was drawn for each edge of the brain network (Fig.1.B). For short time delays, the high value of the AMI indicates little information loss. The AMI drops as we increase the time delay, suggesting a loss of functional information. A more rapid decay indicates faster information loss and, consequently, smaller information storage capability. The Functional Information Decay (FID) matrix in Figure 1.C shows the characteristic decay time for each couple of brain regions, indicating the average information storage capability within specific functional connections. From our analysis on the temporal

properties of the network interactions, an organized spatial structure emerges, with heterogeneous basins each with a characteristic duration of information storage. This may entail different update rates across the brain network, in line with our working hypothesis that information buffering is not homogeneous.

Functional information storage discerns two brain subnetworks

The distribution of the decay times of the functional information reflects a different capability of storing information across the brain network (Fig.2.A). Considering the edges populating the left and right tails of the distribution separately, we find two topographically segregated subnetworks characterized by short and long storage of the functional information, respectively. We expect that the different rates of information loss in the two subnetworks entail a different speed in the information re-updating. In the light of this interpretation, we refer to them as the fast-updating network (FUN) (Fig2.B) and the slow-updating network (SUN) (Fig2.C). On the one hand, the FUN network naturally appears as spanning regions that are related to the processing of stimuli, such as occipital, parietal and posterior temporal areas (Fig2.B). On the other hand, the SUN network mainly involves regions related to higher functions, typically involved in the integration of the information, such as prefrontal regions and the temporal poles.

The multi-storage core defines brain regions hinging communication across subnetworks

Following from the above results, we built a multiplex network, with the FUN and the SUN as the two layers (Fig3). Using the multiplex participation coefficient (MPC) we

characterized each node based upon its connections within the FUN and the SUN layers. We discover the presence of specialized nodes (with low MPC), whose edges project selectively into a single layer (FUN or SUN). Furthermore, there exist nodes sharing their edges equally with both SUN and FUN layers. We define the latter group as the multi-storage core (MSC). Notice that the above definitions of SUN and FUN, as well as that of the MSC, depend on the choice of a threshold density (i.e., the number of edges) (see Fig.2-3). A proper definition of the FUN, SUN and MSC, independent of the network density, is achieved by repeating the analysis for several values of threshold density and averaging over the results (Fig.4 A-C).

The MSC (Fig.4A) is hypothesized to play a role in the communication between the (otherwise) segregated subnetworks and, hence, to be able to integrate both fast and slow timescales. Intriguingly, the MSC naturally clusters around areas that are concerned with high-level processing of auditory stimuli, and speech in particular. Regions involved in the MSC are reported in Table 1.

Discussion

In this paper, we set out to test the hypothesis that the information retained in the link between any two regions has a characteristic lifetime. To test this hypothesis, we estimated the decay of the mutual information of each edge co-activation signal. Based on this measure, we discover two complementary subnetworks. These networks appear to be robust across individuals (Fig.S2). We then proceeded to estimate whether some regions are simultaneously relevant in both (fast and slow) networks. To do so, we conceptualized the two different temporal networks as two layers of a multiplex

network and estimated a nodal metric of concurrent participation to multiple layers, namely the multiplex participation coefficient (MPC) (27, 28). The distribution of the information decays naturally splits the functional network into two spatially segregated subnetworks: one with short storage capability, that we called fast updating network (FUN), and one with greater storage capability, that we named slow updating network (SUN) (Fig2.B-C). The FUN encompasses occipito-parietal and latero-frontal areas, while the SUN spans prefrontal areas and the temporal poles. Multiple models (29) and experimental evidence (30) show that the processing of external stimuli involves (bottom-up) perception and abstraction, as well as (top-down) interpretation according to expectations (priors), embodied in the internal brain state. In turn, such expectations are constantly updated according to new incoming environmental stimuli (31). Previous functional imaging data show that, on the one hand, the prefrontal cortex and areas related to memory storage are involved in top-down processing (30), perhaps by facilitating the entrainment of incoming stimuli into already-existing neuronal ensembles (29). On the other hand, parietal regions contribute to the bottom-up processing of visual tasks (30). Our findings suggest that the division that arises from the timescale separation distinguishes regions involved in bottom-up perception, that are related to a faster timescale, as they need to quickly update external stimuli (FUN), and regions typically involved in top-down processing, operating more slowly (SUN), to allow for higher order integration of information. It has then been posited that communication that answers to different evolutive requirements is underpinned by frequency-specific channels. For example, Bastos et al. reported, in the visual cortical areas of two macaque monkeys, that feedforward influences operate in the gamma and theta band, while feedback influences operate in the beta band (32). In humans, a recent paper applying phase transfer entropy to MEG data shows that two distinct,

frequency-specific directed flows exist in resting-state, with an alpha-band forward flow, and a front-to-back flow in the theta band (33). The two separated flows show some overlap with the FUN and the SUN networks, with the alpha flow partly overlapping the FUN, and the theta flow partly overlapping the SUN. While current evidence does not allow inference about such spatial relationships, one could speculate that the specific frequencies of oscillations might subserve the different velocities at which information needs to be processed across large-scale networks.

The FUN and SUN subnetworks are topographically distinct, as shown in Fig. 2. Fast-updating edges are often incident upon occipital regions, although not uniformly. In fact, it appears that the cunei and the occipital gyri are specifically involved in the FUN network. These regions roughly correspond to Brodmann areas 17, 18 and 19, and are known to be hierarchically related to the extraction of abstract shapes from the visual input. The cuneus directly receives inputs from the retina, that are subsequently processed up to the superior occipital gyrus, in the visual associative cortex, which is responsible of the perception of the properties of objects and plays a role, for instance, in face and object recognition (34, 35) . The occipital regions are connected bilaterally to the inferior parietal gyrus (e.g.. the angular and supramarginal gyri as well as the inferior parietal lobe), the posterior temporal lobes and frontal/prefrontal cortex. The posterior part of the parietal lobe has been included in the "dorsal stream" of the visual systems, which is mainly concerned with the analysis of spatial relationships of objects, and the information about the position of the body in space (36). With regard to the temporal lobe, the posterior part appears to be selectively involved in the FUN. From a structural standpoint, the inferior-longitudinal fasciculus (ILF), as well as an "indirect" stream of U-bundle fibers, connects the occipital and temporal areas (37). Functionally,

these regions are related to either the high-level analysis of images, such as the inferior temporal lobe, whose lesion compromises the ability to recognize facial expressions, or to the perception of- and attendance to auditory stimuli, as in the case of the superior temporal lobe (38). Finally, in the frontal lobes, the FUN appears to involve the ventrolateral and dorsolateral prefrontal cortices, involved in higher functions such as allocation of attention, and typically considered the end of the ventral and dorsal streams, respectively (39, 40). In short, all the regions that are known to be relevant in the process of sensory information emerge as a temporally homogeneous fast network that quickly updates the information that is being processed, from visual and acoustic areas, through layers of abstraction, up to areas related to conscious perception.

The opposite network, to which we refer to as slow-updating network (SUN), clusters topographically in frontal and temporal regions. The regions involved in the SUN network appear to be linked anatomically by the uncinate fasciculus. In fact, the SUN mainly connects the orbitofrontal cortex within the frontal cortex to the temporal poles in the temporal lobes. Patients who underwent brain surgery involving the removal of the uncinate fasciculus showed impairment of verbal memory, naming of objects, verbal fluency, and name anomia. These symptoms did not appear in patients whose surgery did not encompass the removal of the connections between the orbitofrontal cortex and the temporal lobes (41). Furthermore, the semantic variant of primary progressive aphasia and herpes encephalitis, both diseases that induce a damage to the temporal poles, cause disruptions in categorical discrimination, word comprehension and naming (42). Recently, Warren et al. (43), using fMRI showed that, during narrative speech comprehension, the left anterior basal temporal cortex shows high correlation with the left anterior basal frontal cortex, the left anterior inferior temporal gyrus as well as to the

corresponding homotopic temporal cortex contralaterally. From a neurophysiological perspective, the N400 response is an event-related potential (ERP) evoked ~400 ms after (potentially) meaningful material is presented. Using MEG, the N400 was localized in the superior temporal sulcus (44), and intracranial recordings showed it originates in the anteroventral temporal lobe.

Furthermore, the FUN entails more long-range inter-hemispheric connections as compared to the SUN, which might suggest that inter-hemispheric coordination is specifically achieved via fast interactions (perhaps specifically relying on fast, white-matter bundles such as the corpus callosum).

Finally, we used multiplex - network analysis to distinguish regions that are important in both networks from those that are segregated in one layer only. We hypothesize that regions that are equally important in both networks would likely be the ones that allow functional integration between the FUN and the SUN, and this is why we shall refer to this group of nodes as the multi storage core (MSC). Regions that rank very high in the MSC are the Heschl's gyri and the Rolandic opercula, bilaterally. Superior and inferior frontal gyri are also mainly involved in the MSC. These regions are primarily concerned with language comprehension and retrieval, as well as secondary somatic and motor functions (45–47). In conclusion, the MSC seems to localize mainly in regions that are high in the functional hierarchy mediating the cooperation of associative areas (interpretation) and lower areas (that convey and abstract the words). Being able to deal simultaneously with both fast and slow networks might underpin readiness to potentially meaningful stimuli. In fact, it has been proposed that acoustic stimuli must be analyzed at multiple timescales to correctly extract the meaning (48). Hence, the

MSC might represent a link between timescales that are proper of the environment (via bottom-up pathways) with intrinsic brain timescales (proper of top-down processes).

Together, our results are among the first ones to localize temporally coherent patterns of information degradation in segregated brain subnetworks. Importantly, such temporally coherent pathways involve either quickly evolving regions that are concerned with stimulus perception and abstraction, or slowly evolving regions dedicated to hierarchically higher processes. A specific network hinges both time-scales and localizes in regions that are generally involved in language comprehension and production. These subnetworks are found during rest, and might constitute the spatio-temporal scaffolding underpinning the constant readiness to new stimuli. Future work should describe in detail how such spatio-temporal structure supports efficient dynamics and state exploration.

Materials and Methods

Participants

Fifty-eight right-handed and native Italian speakers were considered for the analysis. To be included in this study, all participants had to satisfy the following criteria: a) to have no significant medical illnesses and not to abuse substances or use medication that could interfere with MEG/EEG signals; b) to show no other major systemic, psychiatric, or neurological illnesses; and c) to have no causes of focal or diffuse brain damage at routine MRI. The study protocol was approved by the local Ethics Committee. All participants gave written informed consent.

MRI acquisition

3D T1-weighted brain volumes were acquired at 1.5 Tesla (Signa, GE Healthcare) using a 3D Magnetization-Prepared Gradient-Echo BRAVO sequence (TR/TE/TI 8.2/3.1/450 ms, voxel 1 × 1 × 1 mm3, 50% partition overlap, 324 sagittal slices covering the whole brain).

MEG acquisition

Subjects underwent magnetoencephalographic examination in a 163 – magnetometers MEG system placed in a magnetically shielded room (AtB Biomag UG - Ulm - Germany). The preprocessing was done similarly as in (49). In short, the position of four coils and of four reference points (nasion, right and left pre-auricular point and

apex) were digitized before acquisition using Fastrak (Polhemus®). The brain activity was recorded for 7 minutes, with eyes closed, with a break at 3.5 minutes, so as to minimize the chances of drowsiness. The head position was recorded at the start of each segment. The data were sampled at 1024 Hz and a 4th order Butterworth band-pass filter was applied to select components between 0.5 and 48 Hz. During the acquisitions, electrocardiogram (ECG) and electrooculogram (EOG) were also recorded (50) . These steps were done using Matlab 2019a and the Fieldtrip toolbox 2014 (51).

Preprocessing

Principal component analysis (PCA) was performed to reduce the environmental noise (52, 53). Noisy channels were removed manually through visual inspection of the whole dataset by an experienced rater. Supervised independent component analysis (ICA) was performed to eliminate the ECG (and the EOG) component from the MEG signals (54). Trials that did not contain artefacts (either system related or physiological) or excessive environmental noise were selected.

Source reconstruction

The data were co–registered to the native MRI. A modified spherical conductor model was used as a forward model (55). The voxels in the MRI were labelled according to the Automated Anatomical Labelling (AAL) atlas (56, 57). We used the cortical regions for a total of 78 areas of interest. Subsequently, a linearly constrained minimum variance beamformer was used to compute 78 time series (one per area of interest) at a sampling frequency of 1024 Hz (58). Reconstructed sources were again visually inspected by an experienced rater. Of the initial 58 subjects, 44 had enough artefact-free acquisitions

and were selected for further analysis. The source reconstructed signals were downsampled to 256 Hz.

Edge-centric approach to MEG

In this work, we adopted an edge-centric approach that, rather than focusing on the local activity of the regions (nodes), represents the dynamics of the interactions between couples of brain regions (59). This allowed us to characterize the whole brain network activity in terms of dynamical non-local interactions, highlighting the relational properties of each node. Given any couple of nodes i and j and their respective source-reconstructed signals $X_i(t)$ and $X_j(t)$, we defined a characteristic time series E_{ij} for the edge ij as the product of the z-scored signals, i.e.,

$$E_{ij}(t) = \frac{X_i(t)-\bar{X}_i(t)}{\sigma(X_i(t))} \cdot \frac{X_j(t)-\bar{X}_j(t)}{\sigma(X_j(t))} ,$$

where $\bar{X}$ and $\sigma(X)$ denote the mean and variance of the signals, respectively. One can interpret the edge-coactivation time series as the unfold in time of the pairwise correlations. In fact, the average of the above expression over time corresponds to the Pearson correlation between the signals at nodes i and j. The edge time series were further analyzed by Information theoretic measures, aiming at characterizing the Information storage capability of each functional edge.

Estimation of information decay time through Mutual Information

Shannon Entropy, defined as

$$H(X) = -\sum_{x_i} P_X(x_i) \log(P_X(x_i))$$

quantifies the uncertainty over the possible outcomes x_i of a random variable X with probability distribution P_X. If the uncertainty over future outcomes of X decreases as we measure the outcome y_i of another random variable Y, we conclude that X and Y represent two processes that are not independent. The new resulting uncertainty over X is then defined by

$$H(X|Y = y_i) = -\sum_{x_i} P_{XY}(x_i \mid y_i) \log(P_{XY}(x_i \mid y_i)) = -\sum_{x_i} \frac{P_{XY}(x_i, y_i)}{P_Y(y_i)} \log(\frac{P_{XY}(x_i, y_i)}{P_Y(y_i)}),$$

with $P_{XY}(x_i, y_i)$ denoting the joint probability distribution of the pair (X, Y). The weighted sum over all possible outcomes of Y defines the Conditional Entropy, i.e. the uncertainty of X conditioned on Y

$$H(X|Y) = -\sum_{y_i} P_Y(y_i) H(X|Y = y_i) = -\sum_{x_i, y_i} P_{XY}(x_i, y_i) \log(\frac{P_{XY}(x_i, y_i)}{P_Y(y_i)})$$

$$= H(X, Y) - H(Y),$$

where

$$H(X, Y) \equiv \sum_{x_i, y_i} P_{XY}(x_i, y_i) \log(P_{XY}(x_i, y_i)) .$$

The reduction in uncertainty (or - equivalently - the increase of information) over X given by the knowledge of Y is measured by the Mutual Information (MI)

$$I(X, Y) = H(X) - H(X|Y) = H(X) + H(Y) - H(X, Y).$$

Unlike other measures, such as partial autocorrelation (PAC), Mutual Information (MI) evaluates statistical dependencies which take into account nonlinear interactions, which are ubiquitously found in brain data. In order to quantify the time span before the information in a signal $X(t)$ is lost, we rely on the Auto Mutual Information (AMI) i.e., the MI between the signal and its time delayed copy $Y = X(t-\tau)$. According to previous works on M/EEG (60, 61), a stable estimate of the probability distribution of any real-valued signal $X(t)$ is obtained by dividing the data into 64 bins. The joint probability distribution for the pair $(X(t),\ X(t-\tau))$, needed for the evaluation of the AMI, is given by a 64x64 contingency matrix which is the joint histogram of the two variables. The AMI decays as a function of the delay τ, from a maximal value at $\tau = 0$ to a relatively stable lower value as $\tau \to \infty$. The more gentle ("slower") the decay, the longer the information lasts in the signal.

Information Storage Capability of the functional edges

For each co-activation signal E_{ij}, we estimated the AMI profile (Fig1.B) and we evaluated the time delay τ corresponding to the AMI settles around a stable minimum i.e., when the original signal and its τ-delayed versions have become maximally independent. To do so, we first fit the last 80 points of every AMI profile to a straight line, so as to find the stable minimum. Then, we found the τ corresponding to the moment where the AMI decay falls within a threshold, defined as 1 standard deviation from the stable minimum. Examples of the estimate of the AMI minimum for different edges, at the single epoch level, are shown in Supplementary figure S1. An analysis for different thresholds (number of standard deviations around the stable minimum) is

found in the supplementary material (Fig.S3). Averaging across 10 time windows (epochs) of 10 seconds and across 44 subjects, we found the $N \times N$ Functional Information Decay (FID) matrix (where $N = 78$ is the number of brain regions), containing the decay times for each edge (Fig1.C). The AMI analysis of the co-activations shows that the decay times are different among edges, as revealed by the histogram in Figure 2.A. Selecting the edges from either the left or right tails of the distribution leads to the appearance of two topographically organized subnetworks (Fig.2.B-C).

Multiplex Network analysis

Starting from the FID matrix, i.e. the matrix containing the decay times of the functional information for each edge, we built a multiplex network (Fig3). First, we selected a generic threshold density $\rho = \frac{N_L}{N_E}$, where N_L denotes the number of edges in each layer, and $N_E = \frac{1}{2}N(N-1)$, the total number of edges. We then percolated the delay matrix twice, first by removing the edges starting from the slowest decaying ones, and then starting from the fastest ones. As a result, we obtained two equally dense matrices. In the former, only the fast-decaying edges are included, defining a fast updating subnetwork (FUN). The latter, containing the slow-decaying edges, identifies a slow updating subnetwork (SUN), which is complementary to the FUN. Finally, we defined a two-layered multiplex network, with the layers corresponding to the FUN and the SUN. Then, we evaluated the multiplex participation coefficient (MPC) to distinguish between nodes with edges mainly included in one of the two layers from

nodes with edges that are projecting into both layers. The MPC, for a two-layered network, is defined as:

$$MPC_i = 4\frac{k_i^{FUN} k_i^{SUN}}{(k_i^{FUN}+k_i^{SUN})^2}$$

where k_i^{FUN} and k_i^{SUN} denote the degrees of the i-th node in layer FUN and SUN, respectively. The MPC is calculated per each node and is bounded between 0 and 1, where 1 indicates that the edges that are incident upon a given node are equally distributed on both layers, and 0 denotes a node that is only connected in one of the two layers (27). The greater the MPC, the more a given node participates in both layers.

Parameter independent definitions

The above definitions of the FUN and SUN networks as well as of the MPC depend on the choice of a unique parameter, which is the density ρ of the edges that we include within each of the two layers (See Materials and Methods and Figure 3). In order to make our results independent of this choice, we repeat our analysis for several threshold densities, spanning 100 values within the density range 0.9-0.1 (each time keeping the density equal in both layers). For each density value, we evaluated the FUN and SUN binary matrices as well as the MPC. Averaging across all the density values, we obtained the MPC for each brain region (Fig.4.A) and the typical fast-updating and slow-updating subnetworks (Fig.4.B-C).

Null Model

To statistically validate the MPC results, we compared them to a null model. Our null model is based on the hypothesis that, if there is no organization in the time-delays, the nodes should have equal probabilities of connecting within each layer. This would result in a small distribution width of the MPCs, given that all regions will tend to have the MPC equal to one. However, if a subset of nodes projects selectively into a single layer, then the width of the distribution of the MPC would increase, indicating a different functional role of such regions within the network. Our null model was obtained by calculating the MPC 10000 times, as described above and in figure 3, each time starting from a random surrogate of the FID matrix (Fig.S5.A). The surrogates were obtained by randomly shuffling the FID matrix across rows and columns, keeping the matrix symmetric. Per each iteration, the width of the distribution of the MPC across nodes was stored. Finally, the 10000 widths derived from the permutations were compared to the empirically observed width (Fig.S5.B), showing a clear deviation from the null model. All analyses were made using Matlab, 2019a.

Acknowledgments

The authors thank Michele Allegra for insightful discussions. The work was supported by the University of Naples Parthenope "Ricerca locale" grant, by the grant ANR-17-CE37-0001 - CONNECTOME, by the European Union's Horizon 2020 research and innovation programme under grant agreement No. 945539 (SGA3) Human Brain Project and VirtualBrainCloud No. 826421.

Bibliography

1. K. J. Friston, G. Tononi, O. Sporns, G. M. Edelman, Characterising the complexity of neuronal interactions. *Hum. Brain Mapp.* **3**, 302–314 (1995).
2. A. R. McIntosh, N. Kovacevic, R. J. Itier, Increased Brain Signal Variability Accompanies Lower Behavioral Variability in Development. *PLoS Comput. Biol.* **4**, e1000106 (2008).
3. O. Sporns, The human connectome: Origins and challenges. *NeuroImage* **80**, 53–61 (2013).
4. D. Watts, S. Strogatz, Collective dynamics of 'small-world'networks. *Nature* **393**, 440–442 (1998).
5. Barabasi, Albert, Emergence of scaling in random networks. *Science* **286**, 509–12 (1999).
6. E. Bullmore, O. Sporns, Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**, 186–198 (2009).
7. J. Hahm, *et al.*, Gating of memory encoding of time-delayed cross-frequency MEG networks revealed by graph filtration based on persistent homology. *Sci. Rep.* **7**, 41592 (2017).
8. J. J. Hopfield, Neural networks and physical systems with emergent collective computational abilities. *Proc. Natl. Acad. Sci. U. S. A.* **79**, 2554–8 (1982).
9. O. Kinouchi, M. Copelli, Optimal dynamical range of excitable networks at criticality. *Nat. Phys.* **2**, 348–351 (2006).
10. J. García-Prieto, R. Bajo, E. Pereda, Efficient Computation of Functional Brain Networks: toward Real-Time Functional Connectivity. *Front. Neuroinformatics* **11**, 8 (2017).
11. S. Achard, E. Bullmore, Efficiency and cost of economical brain functional networks. *PLoS Comput. Biol.* **3**, e17 (2007).
12. D. S. Bassett, *et al.*, Dynamic reconfiguration of human brain networks during learning. *Proc. Natl. Acad. Sci.* **108**, 7641–7646 (2011).
13. G. Deco, V. K. Jirsa, A. R. McIntosh, *Emerging concepts for the dynamical organization of resting-state activity in the brain* (Nature Publishing Group, 2011).
14. M. Pedersen, A. Omidvarnia, J. M. Walz, A. Zalesky, G. D. Jackson, Spontaneous brain network activity: Analysis of its temporal complexity. *Netw. Neurosci.* **1**, 100–115 (2017).
15. A. Zalesky, A. Fornito, L. Cocchi, L. L. Gollo, M. Breakspear, Time-resolved resting-state brain networks. *Proc. Natl. Acad. Sci. U. S. A.* **111**, 10341–6 (2014).
16. P. Sorrentino, *et al.*, *Extensive functional repertoire underpins complex behaviours: insights from Parkinson's disease* (Cold Spring Harbor Laboratory, 2019) https:/doi.org/10.1101/823849.
17. G. Hahn, A. F. Bujan, Y. Frégnac, A. Aertsen, A. Kumar, Communication through Resonance in Spiking Neuronal Networks. *PLoS Comput. Biol.* **10**, 1003811 (2014).
18. P. Fries, Rhythms For Cognition: Communication Through Coherence HHS Public Access. *Neuron* **88**, 220–235 (2015).
19. G. Buzsaki, Neuronal Oscillations in Cortical Networks. *Science* **304**, 1926–1929 (2004).
20. A. Kumar, S. Rotter, A. Aertsen, Conditions for propagating synchronous spiking and asynchronous firing rates in a cortical network model. *J. Neurosci.* **28**, 5268–5280 (2008).
21. G. Deco, V. Jirsa, A. R. McIntosh, O. Sporns, R. Kötter, Key role of coupling, delay, and noise in resting brain fluctuations. *Proc. Natl. Acad. Sci. U. S. A.* **106**, 10302–10307 (2009).
22. A. S. Pillai, V. K. Jirsa, Symmetry Breaking in Space-Time Hierarchies Shapes Brain Dynamics and Behavior. *Neuron* **94**, 1010–1026 (2017).
23. J. D. Murray, *et al.*, A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* **17**, 1661–1663 (2014).
24. S. Baillet, Magnetoencephalography for brain electrophysiology and imaging. *Nat. Neurosci.* **20**, 327–339 (2017).
25. F. Lopes da Silva, EEG and MEG: Relevance to neuroscience. *Neuron* **80**, 1112–1128 (2013).
26. J. Faskowitz, F. Z. Esfahlani, Y. Jo, O. Sporns, R. F. Betzel, Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture. *Nat.*

Neurosci., 1–11 (2020).
27. F. Battiston, V. Nicosia, V. Latora, Structural measures for multiplex networks. *Phys. Rev. E - Stat. Nonlinear Soft Matter Phys.* **89**, 032804 (2014).
28. M. Yu, *et al.*, Selective impairment of hippocampus and posterior hub areas in Alzheimer's disease: an MEG-based multiplex network study. *Brain* **140**, 1466–1485 (2017).
29. A. K. Engel, P. Fries, W. Singer, Dynamic predictions: Oscillations and synchrony in top–down processing. *Nat. Rev. Neurosci.* **2**, 704–716 (2001).
30. T. J. Buschman, E. K. Miller, Top-Down Versus Bottom-Up Control of Attention in the Prefrontal and Posterior Parietal Cortices. *Science* **315**, 1860–1862 (2007).
31. K. Friston, The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138 (2010).
32. A. M. Bastos, *et al.*, Visual Areas Exert Feedforward and Feedback Influences through Distinct Frequency Channels. *Neuron* **85**, 390–401 (2015).
33. A. Hillebrand, *et al.*, Direction of information flow in large-scale resting-state networks is frequency-dependent. *Proc. Natl. Acad. Sci. U. S. A.* (2016) https:/doi.org/10.1073/pnas.1515657113.
34. E. Freud, J. C. Culham, D. C. Plaut, M. Behrmann, The large-scale organization of shape processing in the ventral and dorsal pathways. *eLife* **6** (2017).
35. L. A. Renier, *et al.*, Preserved functional specialization for spatial processing in the middle occipital gyrus of the early blind. *Neuron* **68**, 138–148 (2010).
36. M. A. Goodale, A. D. Milner, *Separate visual pathways for perception and action* (Trends Neurosci, 1992).
37. G. Herbet, I. Zemmoura, H. Duffau, *Functional Anatomy of the Inferior Longitudinal Fasciculus: From Historical Reports to Current Hypotheses* (Frontiers Media S.A., 2018).
38. M. Vander Ghinst, *et al.*, Left superior temporal gyrus is coupled to attended speech in a cocktail-party auditory scene. *J. Neurosci.* **36**, 1596–1606 (2016).
39. F. de Pasquale, *et al.*, A Cortical Core for Dynamic Integration of Functional Networks in the Resting Human Brain. *Neuron* **74**, 753–764 (2012).
40. P. S. Goldman, H. E. Rosvold, Localization of function within the dorsolateral prefrontal cortex of the rhesus monkey. *Exp. Neurol.* **27**, 291–304 (1970).
41. C. Papagno, *et al.*, What is the role of the uncinate fasciculus? Surgical removal and proper name retrieval. *Brain* **134**, 405–414 (2011).
42. U. Noppeney, *et al.*, Temporal lobe lesions and semantic impairment: A comparison of herpes simplex virus encephalitis and semantic dementia. *Brain* **130**, 1138–1147 (2007).
43. J. E. Warren, J. T. Crinion, M. A. Lambon Ralph, R. J. S. Wise, Anterior temporal lobe connectivity correlates with functional outcome after aphasic stroke. *Brain* **132**, 3428–3442 (2009).
44. K. Marinkovic, *et al.*, Spatiotemporal Dynamics of Modality-Specific and Supramodal Word Processing The importance of the left anterior temporal cortex for. *Neuron* **38**, 487–497 (2003).
45. C. Y. Chen, *et al.*, MR of the cerebral operculum: topographic identification and measurement of interopercular distances in healthy infants and children. *Am. J. Neuroradiol.* **16** (1995).
46. S. E. Starkstein, M. Berthier, R. Leiguarda, Bilateral opercular syndrome and crossed aphemia due to a right insular lesion: A clinicopathological study. *Brain Lang.* **34**, 253–261 (1988).
47. G. Villa, C. Caltagirone, Speech suppression without aphasia after bilateral perisylvian softenings (bilateral rolandic operculum damage). *Ital. J. Neurol. Sci.* **5**, 77–83 (1984).
48. X. Teng, X. Tian, D. Poeppel, Testing multi-scale processing in the auditory system. *Sci. Rep.* **6**, 1–13 (2016).
49. P. Sorrentino, *et al.*, Brain functional networks become more connected as amyotrophic lateral sclerosis progresses: a source level magnetoencephalographic study. *NeuroImage Clin.* **20**, 564–571 (2018).
50. J. Gross, *et al.*, Good practice for conducting and reporting MEG research. *NeuroImage* **65**, 349–63 (2013).
51. R. Oostenveld, *et al.*, FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data, FieldTrip: Open Source Software for Advanced Analysis of MEG, EEG, and Invasive Electrophysiological Data. *Comput. Intell. Neurosci.* (2011)

https:/doi.org/10.1155/2011/156869, 10.1155/2011/156869.
52. A. de Cheveigné, J. Z. Simon, Denoising based on time-shift PCA. *J. Neurosci. Methods* **165**, 297–305 (2007).
53. Sadasivan P.K., Narayana Dutt D., SVD based technique for noise reduction in electroencephalographic signals. *Signal Process.* **55**, 179–189 (1996).
54. G. Barbati, C. Porcaro, F. Zappasodi, P. M. Rossini, F. Tecchio, Optimization of an independent component analysis approach for artifact identification and removal in magnetoencephalographic signals. *Clin. Neurophysiol.* **115**, 1220–1232 (2004).
55. G. Nolte, The magnetic lead field theorem in the quasi-static approximation and its use for magnetoencephalography forward calculation in realistic volume conductors. *Phys. Med. Biol.* **48**, 3637–3652 (2003).
56. G. Gong, *et al.*, Mapping anatomical connectivity patterns of human cerebral cortex using in vivo diffusion tensor imaging tractography. *Cereb. Cortex N. Y. N 1991* **19**, 524–36 (2009).
57. N. Tzourio-Mazoyer, *et al.*, Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *NeuroImage* **15**, 273–89 (2002).
58. B. D. Van Veen, W. Van Drongelen, M. Yuchtman, A. Suzuki, Localization of Brain Electrical Activity via Linearly Constrained Minimum Variance Spatial Filtering. *IEEE Trans. Biomed. Eng.* **44** (1997).
59. J. Faskowitz, F. Z. Esfahlani, Y. Jo, O. Sporns, R. F. Betzel, Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture. *bioRxiv*, 799924 (2019).
60. C. Gómez, R. Hornero, D. Abásolo, A. Fernández, J. Escudero, Analysis of the magnetoencephalogram background activity in Alzheimer's disease patients with auto-mutual information. *Comput. Methods Programs Biomed.* **87**, 239–247 (2007).
61. J. Jeong, J. C. Gore, B. S. Peterson, Mutual information analysis of the EEG in patients with Alzheimer's disease. *Clin. Neurophysiol.* **112**, 827–835 (2001).

Figures and figure captions

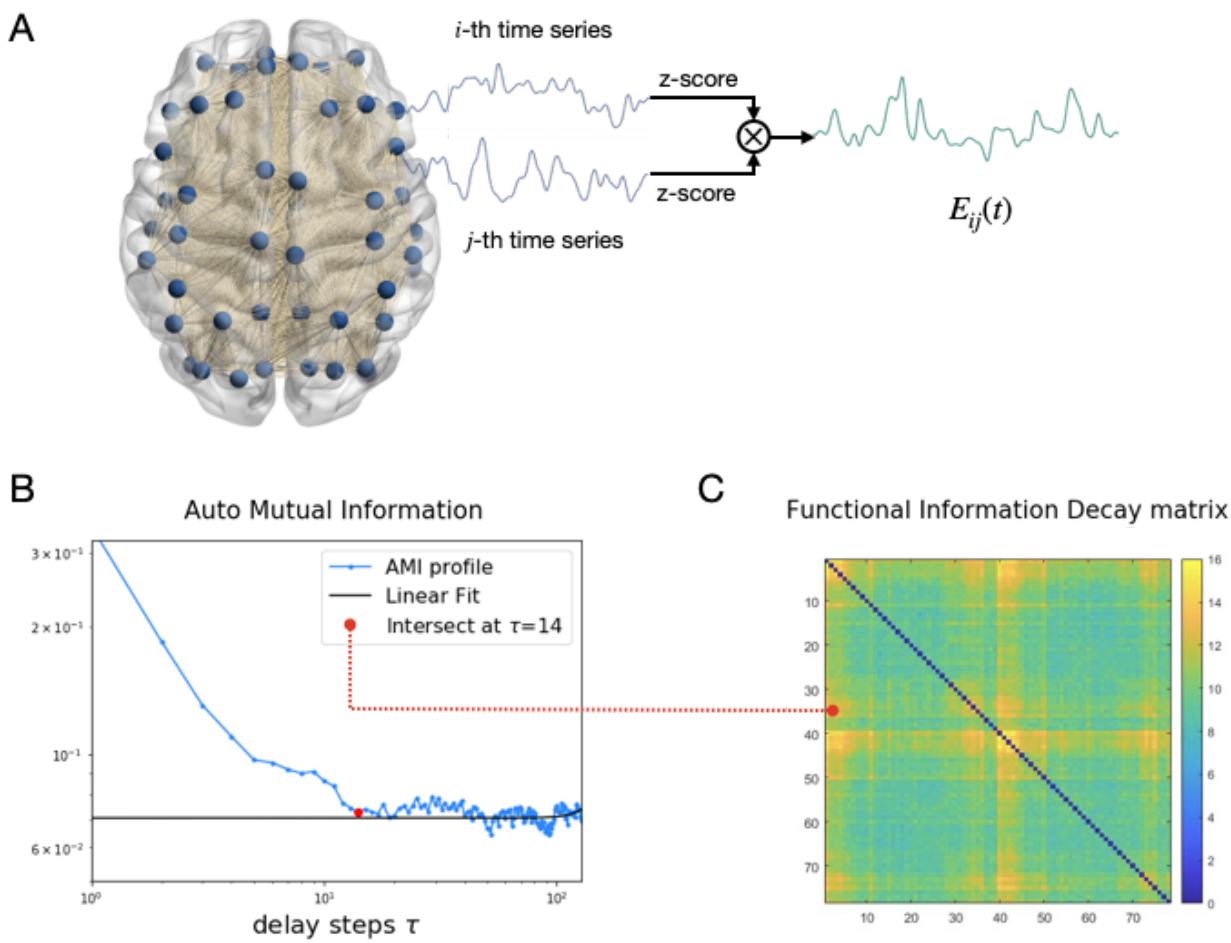

Fig. 1. Edge-centric analysis of MEG activity. (A) The preprocessed magnetoencephalography (MEG) dataset consists of $N = 78$ source-reconstructed signals, one for each brain region. A co-activation time series $E_{ij}(t)$ is defined for every edge as the product of the z-scored signals at the extremal nodes i and j, and it describes the temporal unfold of their correlation. (B) The Auto Mutual Information (AMI) between an edge time series and its τ-delayed copy measures the non-linear

statistical dependencies of the two signals. The AMI rapidly decays as we increase the time lag τ . In order to estimate the time after which the functional information across nodes i and j is fully lost, we fit the tail of the AMI profile with a linear function (black line) and we find the first point where the AMI profile approaches the tail (red dot), indicating the information decay-time τ_{ij} for the edge ij . Notice that each time step τ corresponds to 3,9 milliseconds, given the sample frequency of 256 Hz (C) The $N \times N$ Functional Information Decay (FID) matrix is defined in each entry by the value $< \tau_{ij} >$ (averaged across 44 healthy subjects and across ten different time windows of 10 seconds acquisition). For each couple of brain regions this matrix describes the typical time that the functional information is preserved along the edge connecting them.

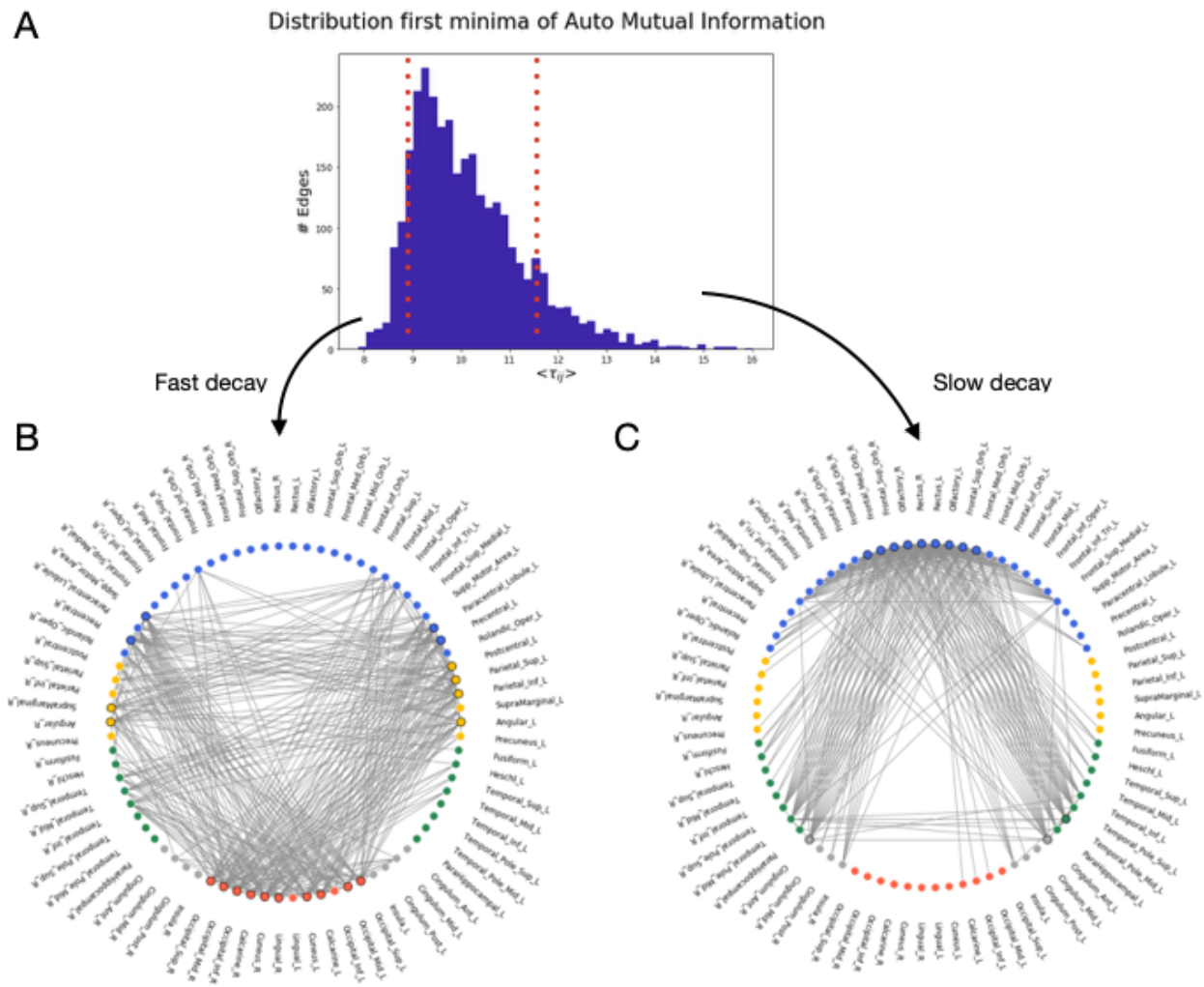

Fig. 2. Long- and short-storage subnetworks. (A) Distribution of the average decay times of the AMI for all the edges. (B-C) We use the fastest and slowest decaying edges to define two network layers with the same density ρ. Here, we show the networks resulting from $\rho = 0.1$. A characteristic topographic distribution of the layers emerges. (B) The fastest decaying edges define a fast-updating network (FUN) in which the rapid loss of information requires a quick update of the information. (C) The slowest decaying edges form a complementary slow-updating network (SUN) with lower buffering speed.

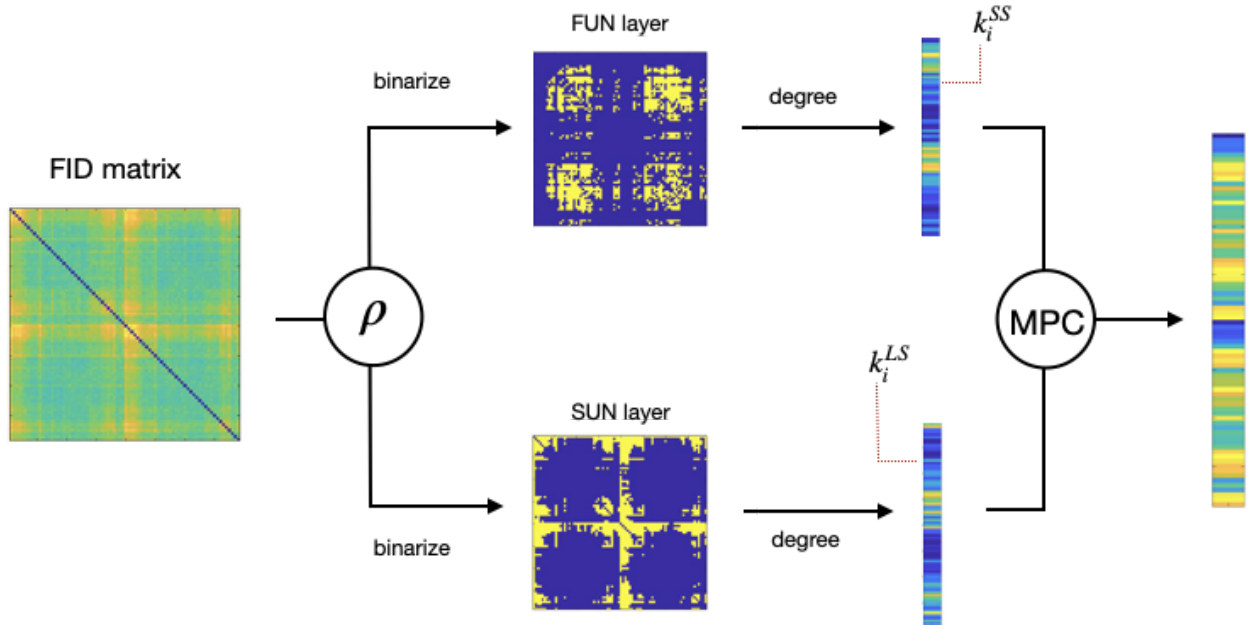

Fig. 3. Multiplex network and regional analysis. Given a generic threshold density ρ, links are added starting from the lowest (highest) entries until density ρ is reached, obtaining a binary network containing the fastest (slowest) decays of the AMI, named FUN (SUN). Summing over all the columns of the binarized matrix we obtained the degree k_i^{FUN} (k_i^{SUN}) for node i, which provides a measure of participation within the FUN (SUN) layer. The degree of participation of a brain region into the two-layered multiplex network (made by the FUN and the SUN) is measured by the Multiplex Participation Coefficient (MPC). The MPC for node i is obtained as described in the methods section. If the MPC is zero, the associated node projects its edges to a single (either FUN or SUN) layer. A unitary value of the MPC is associated with those nodes whose edges equally connect into both network layers.

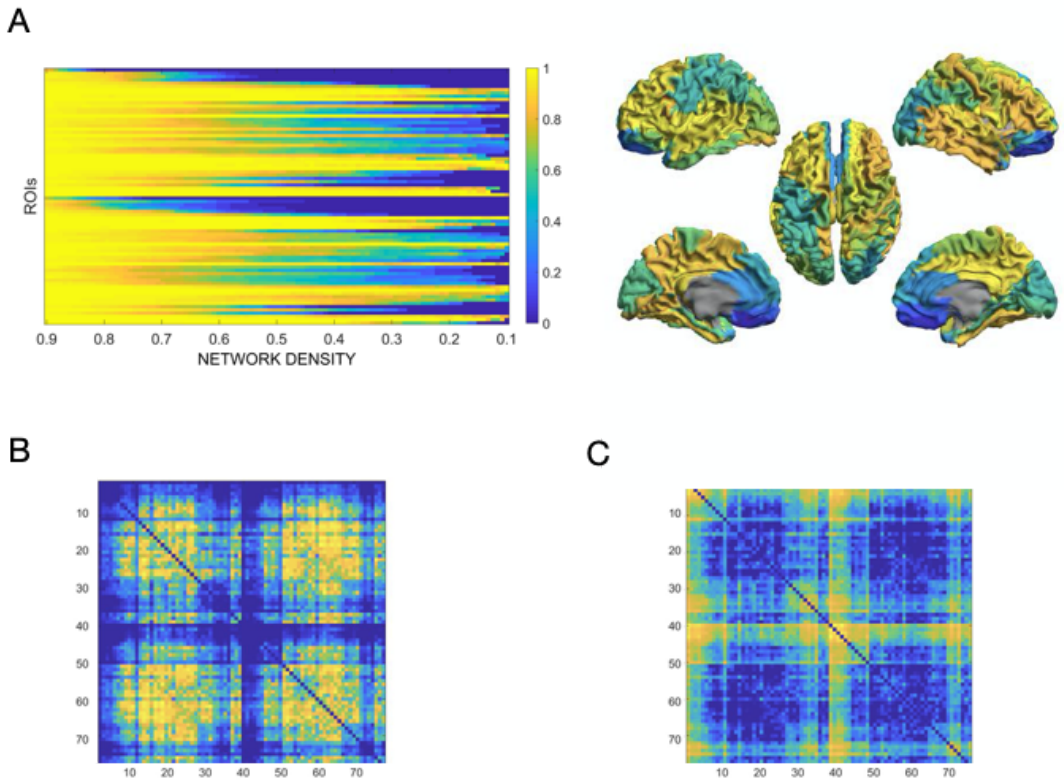

Fig. 4. Definition of the multi-storage core (MSC). (A) We evaluated the MPC for different values of the threshold density ρ. For high values of graph density the FUN and SUN layers largely overlap. Therefore each node shares its edges with both layers and the associated MPC is everywhere approximately 1. As density decreases, the layers become more segregated revealing a relict structure. A specialized subset of brain areas (in blue) operate within either the short- or the long-storage subnetwork. Another topographically organized group of nodes (in yellow) hinge the fast and slow storage subnetworks defining a multi-storage core (MSC). The MSC emerges without the a priori setting of any parameter and is hypothesized to play a role in the integration between the two layers. (B-C) The FUN and SUN layers averaged across the whole range of densities.

Table 1. Regions belonging to the MSC, ranked.

Rolandic Operculum (R)
Insula (L)
Heschl gyrus (R)
Rolandic Operculum (L)
Superior Frontal Gyrus (R)
Heschl gyrus (L)
Inferior frontal gyrus, triangular part (L)
Posterior cingulate gyrus (R)
Superior temporal gyrus (L)
Middle temporal gyrus (L)
Middle frontal gyrus (L)
Frontal frontal gyrus (L)
Posterior cingulate gyrus (L)
Median cingulate and paracingulate gyrus (R)
Precuneus (R)
Inferior frontal gyrus, opercular part (R)
Superior temporal gyrus (R)
Fusiform gyrus (L)
Median cingulate and paracingulate gyrus (L)
Fusiform gyrus(R)

Supplementary Material

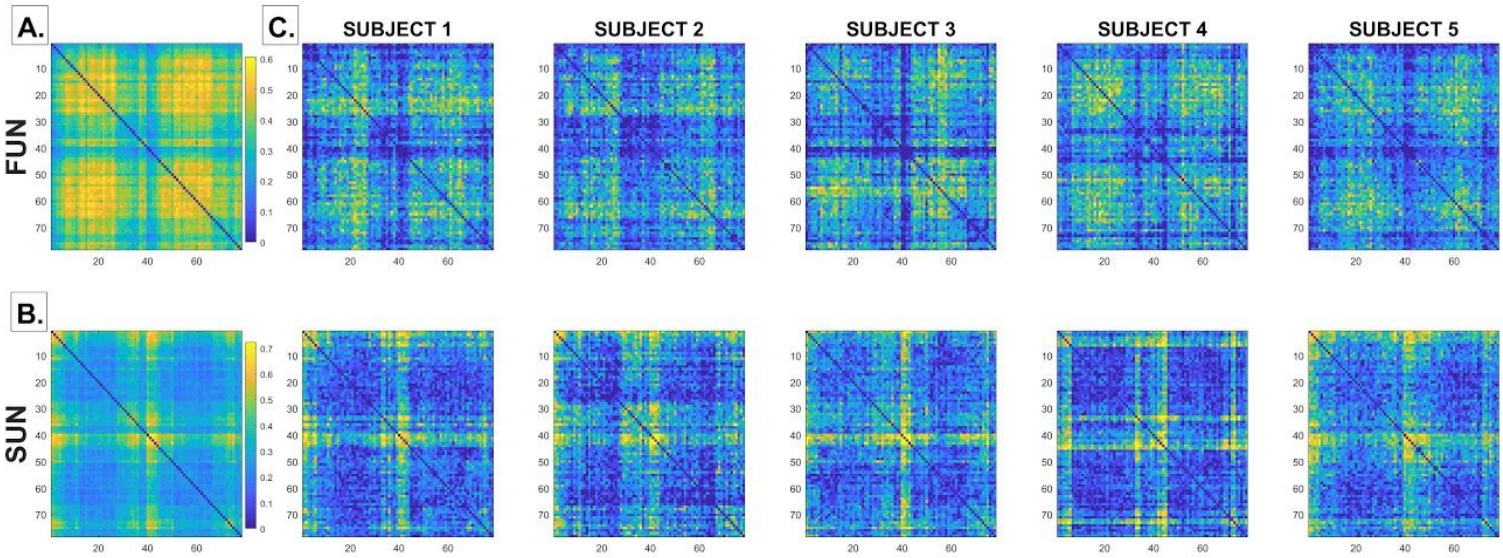

Fig. S1: To test the validity of the results at the single subject level we have estimated, in each epoch and for every subject, the membership of each edge to the FUN (SUN). Namely, for each trial we found the AMI minima of each edge and set it to 1 if it belonged to the first (last) quartile of the minima distribution, to 0 otherwise. We averaged across epochs and subjects obtaining an estimate of how reliably each edge belongs to the FUN (SUN), as shown in panel A (B). The results at the individual level are shown for five randomly selected subjects in panel C.

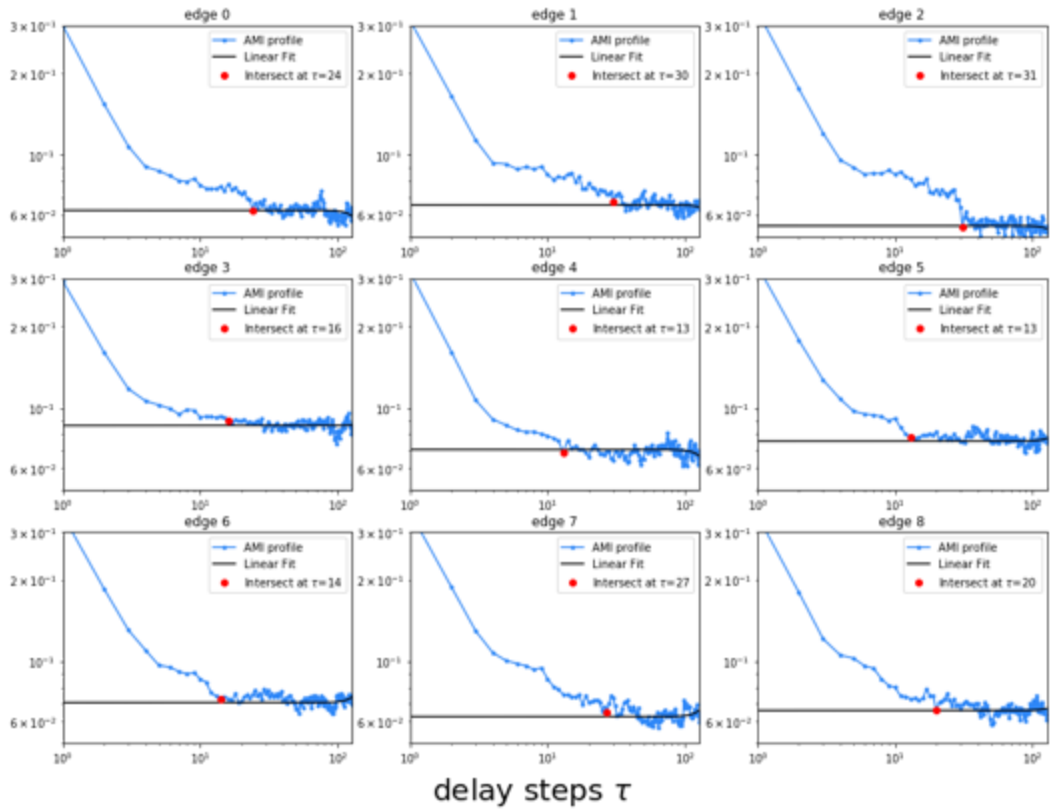

Fig. S2: Examples of the AMI decays of 9 randomly selected edges in a single trial (i.e., form the same 10 seconds-long epoch from a randomly selected subject) (log-log scale). In order to estimate the time after which the co-activation time series and its delayed version become maximally independent, we fit the last 80 points of the AMI profile (blue) to a straight line (black) and we find the first point (red) where the difference between the AMI decay and the fitted line is below a threshold. The results reported in the main text refer to the threshold set to 1 standard deviation from the AMI tail.

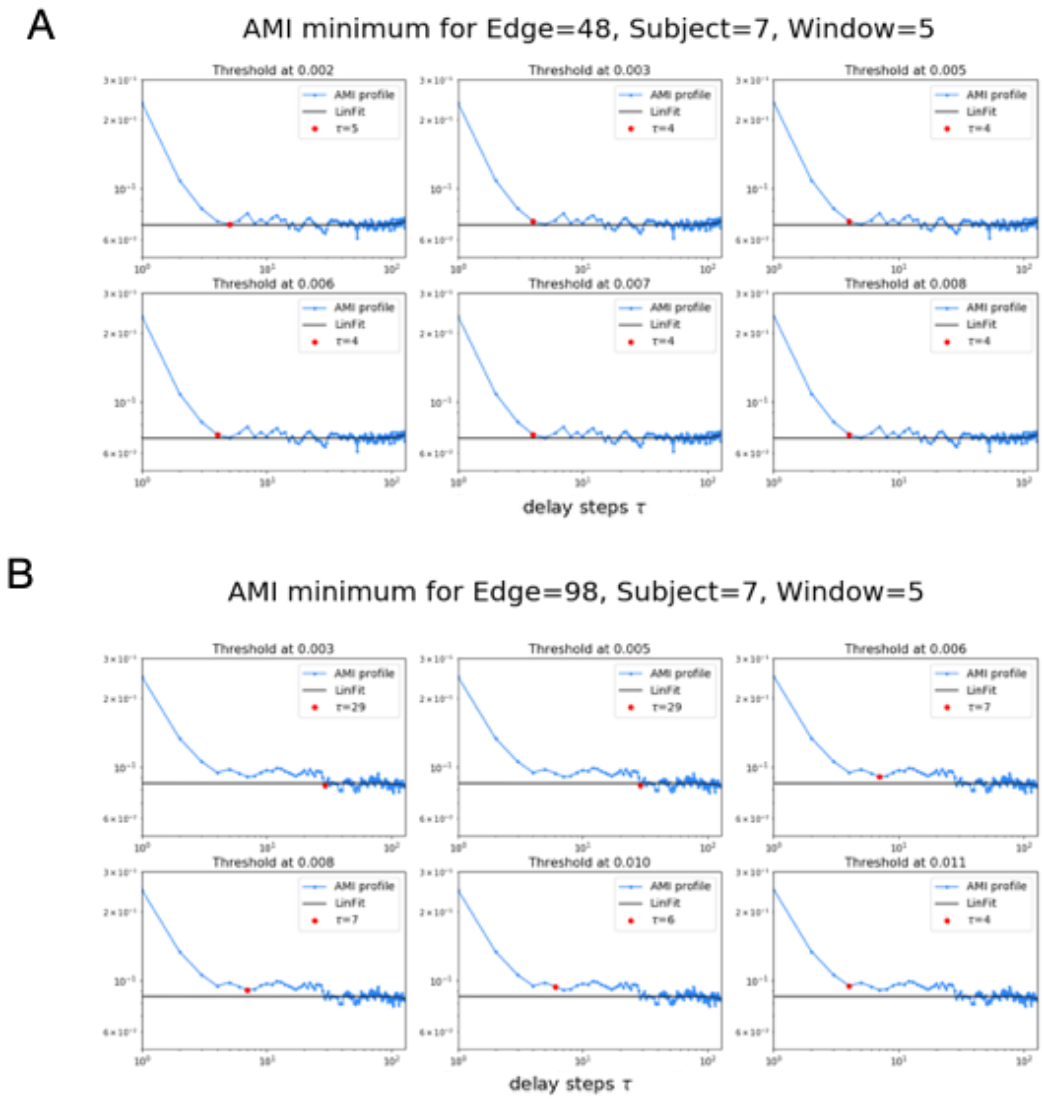

Fig. S3: The blue lines show the AMI decay profile for a single epoch. The red dots represent the estimate of the decay-times. The decay-time is measured as the first time that the AMI profile approaches the null-line (fitted to the tail of the AMI distribution, in black). The distance from the null-line, considered as a threshold, is defined as multiples of the standard deviation ([1, 1.5, 2., 2.5, 3., 3.5]) of the AMI tail. The six panels show the corresponding estimate of the decay time as the threshold is increased, from top-left to bottom right. (A) An example is drawn from a group of edges where the AMI decay profile approaches the minimum directly. The corresponding estimate

of the decay time is stable across thresholds. (B) An example is drawn from a subset of edges where the decay profile does not approach the minimum directly. In such cases, the AMI decay profile shows a fat tail before reaching a stable minimum. Since we are interested in the time required for the information to be completely extinguished in an edge, we considered the smallest threshold for the estimation of all the minima. This provides a stable estimate of the minimum over all the observed decay profiles.

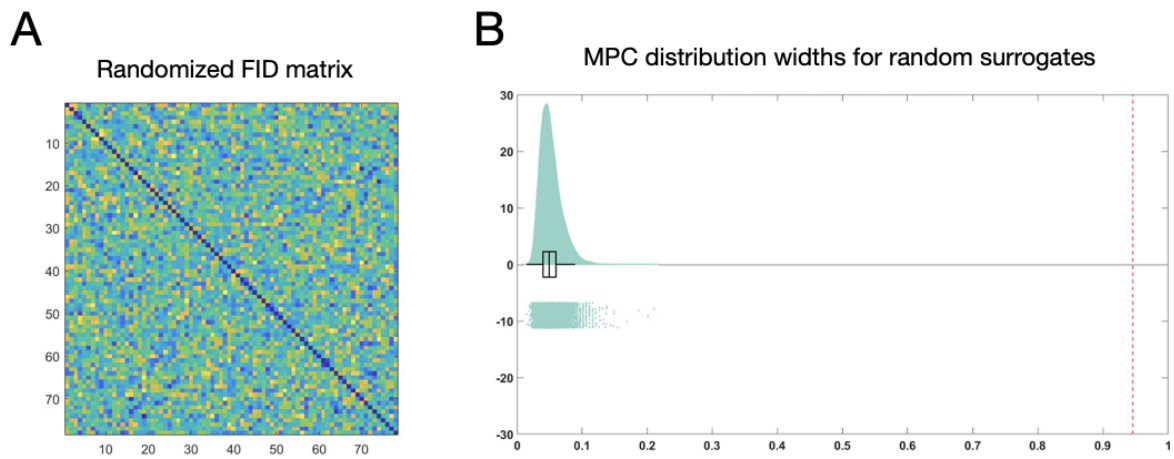

Fig. S4. Deviation from a random null model. (A) To check the statistical significance of the results, the MPC was calculated on 10000 randomizations of the FID matrix. (B) In the null models, the distribution width across regions of the MPC is very low, meaning that, at random, all the regions equally belong to both layers. That is, they have the same probability of sharing an edge with either the FUN and the SUN subnetworks. Notably, the distribution width across regions in the test case is significantly larger (red line), meaning that some regions specifically operate within only one of the two layers.