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[2,3,5]+, 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. Applying edge-wise connectivity to high temporal-resolution Magnetoencephalography (MEG) signals, we estimate the persistence of the information in functional interactions across the brain. The characterization of brain network edges according to their information storage capability naturally partitions the brain into two anatomically distinct networks. The "short storage network" (SSN) is localized in the regions that typically belong to the dorsal and ventral streams during perceptive tasks, while the "long storage network" (LSN) hinges on classically associative areas. Finally, we show that only a subset of the brain regions, which we name the multi-storage core (MSC), belong to both networks. The MSC is hypothesized to play a role in the communication between the (otherwise) segregated subnetworks.

Keywords: resting state dynamics, multiple time-scales, edge co-activations, magnetoencephalography, multilayer networks.

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. Characterizing the life-time of the information across interacting brain regions, we show the existence of two spatially and temporally segregated networks with distinct capabilities of retaining information. 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. From the functional point of view, large-scale patterns of interactions among areas are non-stationary, and reorganize over time (4–7). 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 (8). To achieve this readiness, the brain has to quickly acquire information from the environment, integrate it and interpret it. However, the spatiotemporal organization underpinning such capability is yet to be understood. While multiple mechanisms of communication have been proposed (among others, communication through resonance (9), coherence (10), oscillations (11), synfire chains (12)), 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 shaping communication channels with intrinsic operational times, resulting in a spatio-temporally nested activity (13, 14). In this line of thinking, it has been shown in both human and macaque that areas that are hierarchically lower in

information processing operate at higher speed as compared to associative areas (which integrate information) (15–19).

How this reflects itself on the interactions among regions is not clear. Here, we hypothesize that some tasks might involve quick interactions, while other functions might instead require the integration of more information and, presumably, a longer time to unfold. Such differences in the integration process should reflect in the retention of information across interacting regions. In this paper, we represent the dynamic interaction among pairs of regions as their edge co-activation time-series, a time-resolved measure of Pearson correlation (20). Then, classifying the edges upon retention of information of the co-activation time series should naturally lead to the 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 electroencephalography (EEG). However, EEG is difficult to analyze and interpret because the electric fields are distorted 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 (21, 22). The Auto Mutual Information (AMI) was used to estimate the time up to which the co-activation information is retained within each edge. Finally, we further hypothesized that specific regions should mediate the integration of groups of edges with different information storage duration. To test this, we built a multiplex network (23–25) where

each layer is filled by edges with either short and long storage capability, 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 unfolding 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 co-fluctuation 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 (26). 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 further loss of information. A more rapid decay indicates faster information loss and, consequently, smaller information storage capability.

We show that the AMI decay times of the edges in the brain network are organized according to a characteristic spatio-temporal pattern which is consistent across trials and subjects (Fig.1.c, top right). The trial-specific distributions of the decay times can be multi-modal, showing that the brain can dynamically split into sub-networks operating at different time-scales (examples are shown in Fig.1.c, bottom right). At the aggregate level (i.e. averaging across trials), we define the Functional Information Decay (FID) matrix (Fig.1.d), which reveals a temporal hierarchy of the edges. In conclusion, edge-wise analysis reveals an organized large-scale spatio-temporal structure, with heterogeneous basins each with a characteristic duration of information storage.

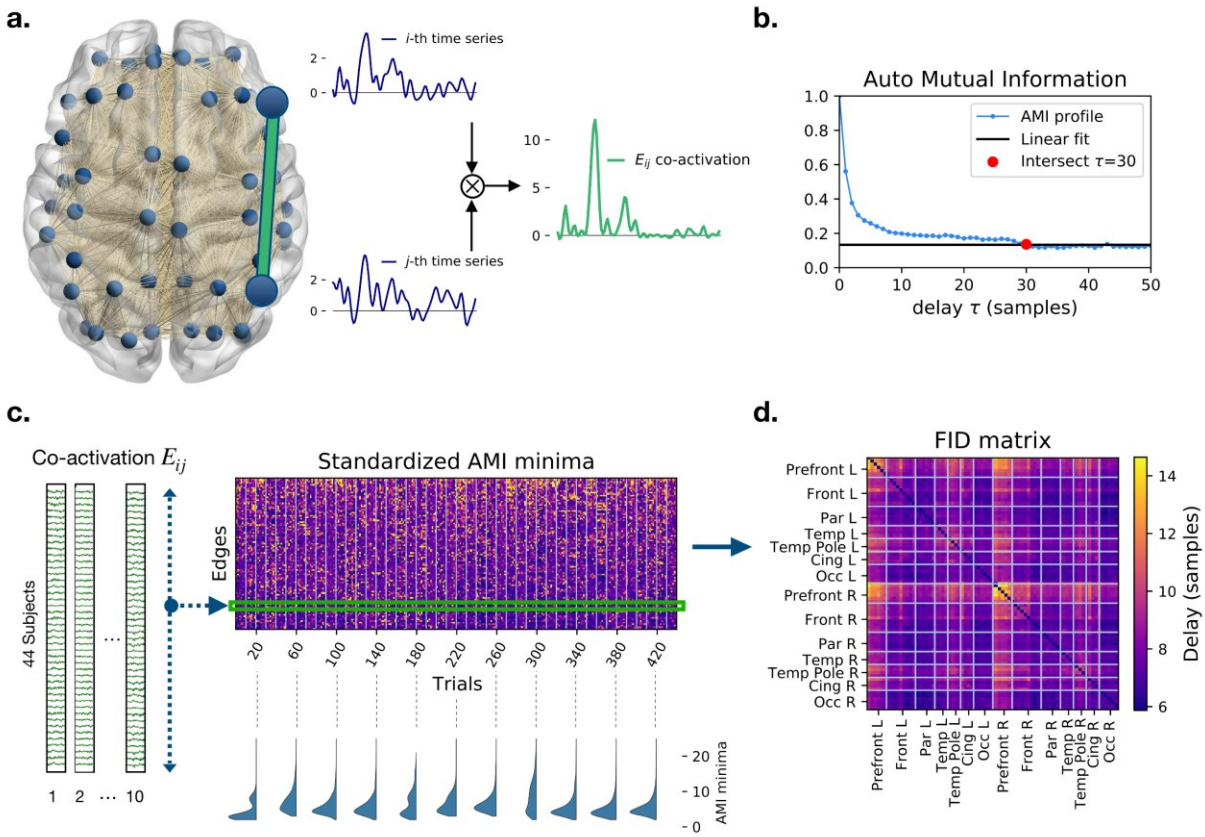

Fig. 1. Auto mutual information analysis of the edge time-series in MEG data. a. The preprocessed magnetoencephalography (MEG) dataset consists of $N = 78$ source-reconstructed signals, one for each brain region. We define the edge time series $E_{ij}(t)$ (green) as the element-wise product of nodal z-scored signals at i and j (blue). **b.** Each point in the plot represents the mutual information between the co-activation time series and its τ-delayed copy. The red dot represents the delay at which the AMI reaches a minimal level (defined as a linear fit of the AMI tail). Notice that each time step τ corresponds to 3,9 milliseconds, given the sample frequency of 256 Hz. **c.** For each of the 44 subjects, the time series are split into 10, 10 second-long segments (epochs). For each trial, the τcorresponding to the AMI minima is computed. In the carpet plot, the minima have been standardized across edges for each trial separately and sorted to allow comparison (sorting based on the averages across trials). The strong tendency of high decay to distribute in the top part of the carpet plot shows that the

average estimate conveys trial-specific information. The example minima distributions (bottom) in randomly selected epochs show that multimodality can emerge naturally **d.** Averaging over the standardized AMI minima, we obtain the $N \times N$ Functional Information Decay (FID) matrix, where rows and columns are regions, and the matrix elements are the average minima across trials. For each couple of brain regions this matrix describes the typical time that the functional information is preserved along the edge connecting them.

Functional information storage identifies two brain networks

Considering the edges with the lowest and highest retention of information in the FID matrix, we found two topographically segregated sub-networks. We refer to them as the short-term storage network (SSN) (Fig2.a, left) and the long-term storage network (LSN) (Fig.2.a, right). On the one hand, the SSN spans regions that are related to the processing of stimuli, such as the occipital, parietal and posterior temporal areas. On the other hand, the LSN network mainly involves regions related to higher cognitive functions, such as the prefrontal regions and the temporal poles (list of regions in Table 1).

To test whether the aggregate topography can be inferred from local properties alone, we generated Gaussian (G-) surrogates, where the observed nodal autocorrelations (i.e., the power spectrum) are imposed on otherwise independent processes. In this case, the distribution of the decay times is visibly shrunk (Fig.2.b, green and blue), and long decay times are absent. However, the topography is preserved (i.e., perceptive regions belong to the SSN and associative regions to the LSN), as the G-surrogate FID matrix correlates with the original one (Spearman's $r_s = 0.89$). In order to exclude the possibility that the long AMI decay times are the product of the static correlation across regions we

generated, starting from the original data, phase-randomized (P-) surrogates preserving both the regional autocorrelations and the pairwise static correlations (i.e., the cross-spectrum). In the P-surrogates, the heavy-tailed distribution of the decay times is also shrunk as compared to the empirical data (Fig.2.b, orange and blue, respectively), and the topography is still preserved ($r_s = 0.94$). Note that the pairwise Pearson correlation preserved by the P-surrogates is equivalent by definition to the time average of the corresponding edge co-activation, which is a static property. Thus, our results suggest that the long decay times of edge-specific information depend upon the dynamics of the interactions.

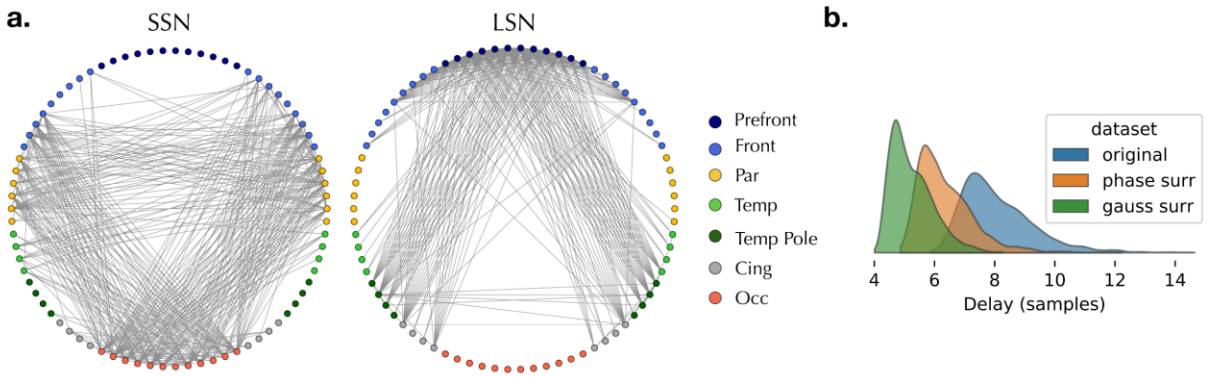

Fig. 2. Topography of delays and surrogate analysis. a. The edges with the fastest (left) and slowest (right) decay times naturally split into a short-storage and a long-storage network (SSN and LSN). **b.** Distributions of the average edge AMI decays in the original (blue), phase (orange) and Gaussian (green) surrogates shows that the long decay times of information depend on the dynamics of the edges and are not explained by nodal spectral features or by static correlations.

A multi-storage core links the short- and long-storage networks

Finally, we sought to understand whether specific regions mediate the interaction between the short-storage and long-storage sub-networks. To this end, we built a multiplex network (23) where each layer is filled by either fast or slow decaying edges extracted from the aggregate FID matrix (up to a fixed network density ρ ; Fig.3.a, left). Using a modified version of the multiplex participation coefficient (MPC) we characterized each node according to its connections within the SSN and the LSN layers (Fig.3.a, right and Methods). If the MPC of a region is zero, then its edges are equally split between both layers. If the MPC is -1 (+1), then the node projects solely to the SSN (LSN) layer. Threshold-independent definitions are achieved by repeating the analysis for several values of network density (Fig.3.b, bottom) and averaging over the results (Fig.3.b, bottom right). We discover specialized nodes defining short- and long-storage cores (SSC with MPC$<< 0$ and LSC with MPC$>> 0$, respectively), whose edges project prominently into a single layer (SSN or LSN). Also, we discover a subset of areas (MPC$\simeq 0$) sharing their edges equally with both the SSN and LSN layers, that we named multi-storage core (MSC). We repeated the same analysis at the single trial level showing that the results are consistent across subjects and epochs (Fig.3.c).

A structured MSC is observed also averaging across trials (Fig.3. c, right and brain plots in d) as opposed to starting from the aggregate FID matrix. The MSC is hypothesized to mediate the communication between both the 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. The regions encompassing the MSC are reported in Table 1).

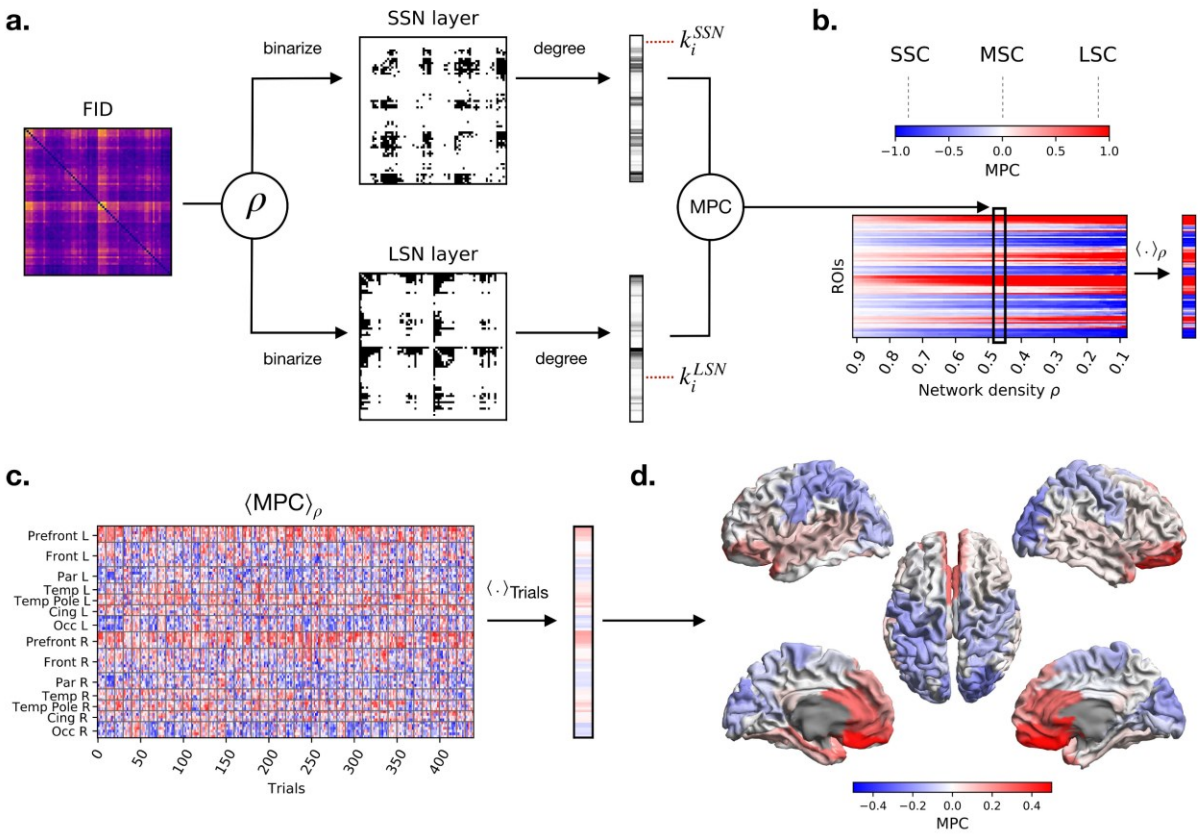

Fig. 3. Multiplex Network Analysis. a. Given a generic threshold density ρ, edges are added starting from the lowest (highest) entries until density $\rho$is reached, obtaining a binary network containing the fastest (slowest) AMI decays, named SSN (LSN). Summing over all the columns of the binarized matrix we obtained the degree ${k_i}^{SSN}({k_i}^{LSN})$ for node i, which provides a measure of participation within the SSN (LSN) layer. The degree of participation of a region into the two-layered multiplex network (made by the SSN and the LSN) is measured by the Multiplex Participation

Coefficient (MPC). The MPC for node i is obtained as described in the methods section. **b.** The procedure described in panels (a) is repeated across a range of densities (carpet plot, bottom), and the MPC is calculated, per each node, as the average MPC across all threshold densities (bottom, right). If the MPC is -1 (+1), the associated node projects all its edges into the SSN (LSN) layer. A zero value of the MPC is associated with those nodes whose edges equally connect into both network layers. **c.** The procedure in panel (a-b) is repeated at the single trial layer. The carpet plot (left) shows the trial-specific MPC's. The corresponding MPC's averaged across trials (right) are depicted in the glass brain in **d**.

Discussion

In this paper, we set out to test the hypothesis that the information retained in the co-fluctuation of any two regions has a characteristic lifetime. To test this hypothesis, we estimated the decay of the auto mutual information of each edge co-activation signal. Based on this measure, we discover two complementary subnetworks consistent across individuals and trials (Fig.1.c). 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) (23, 27). The distribution of the information decays naturally splits the functional network into two spatially segregated subnetworks: one with short storage capability, that we called short storage network (SSN), and one with longer storage capability, that we named long storage network (LSN) (Fig2.B-C). The SSN encompasses occipito-parietal and latero-frontal areas, while the LSN spans prefrontal areas and the temporal poles. Multiple models (28) and experimental evidence (29) 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 (30). 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 (29), perhaps by facilitating the entrainment of incoming stimuli into already-existing neuronal ensembles (28). On the other hand, parietal regions contribute to the bottom-up processing of visual tasks (29). Our findings suggest that the division arising from the timescale separation distinguishes regions involved in bottom-up perception, that are related to a faster timescale (SSN), and regions typically involved in top-down processing, operating more slowly (LSN), allowing 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 (31). In humans, a recent study applying phase transfer entropy to MEG data shows that two distinct, frequency-specific directed flows exist in the resting-state, with an alpha-band forward flow, and a front-to-back flow in the theta band (32). The two separated flows show some overlap with the SSN and the LSN networks, with the alpha flow partly overlapping the SSN, and the theta flow partly overlapping the LSN. In fact, one could speculate that the specific frequencies of oscillations might subserve the different speed at which information needs to be processed across regions. However, one should keep in mind that our approach might be also sensitive to scale-free dynamics, and hence comparing this evidence is non-trivial. For instance, parts of the frontal cortex canonically regarded as associative belong to the SSN,

despite displaying slower rhythms as compared to more posterior ones. A gradient of time-scales was also recently observed in (16) at the nodal level. Our approach is complementary since it focuses on higher-order interactions across regions (i.e., edges). At the single trial level the distributions of the decay times do not appear to have a constant shape during resting-state, with multimodality dynamically segregating time scales (Fig.1.c, bottom right). At the aggregate level, we show that the time-scales gradually morph, spanning between the SSN and the LSN.

From a technical standpoint, this paper uses the per-sample product of two z-scored time series to compute the instantaneous strength of functional connections (Fig.1.a). Such an approach allows a dynamical estimate of the connectivity while avoiding possible artifacts induced by more traditional sliding window approaches (33). It is worth noticing that the correlations among signals might be dominated by certain frequency bands. Hence, further analysis should disentangle the relative contribution of specific frequencies to the overall spatio-temporal unfolding of the brain activity. Using surrogate analysis we showed that the observed range of the decay times is not explained by nodal properties alone (Fig.2.b), since long-decays are absent when preserving auto- and cross-correlation of the nodes (power spectrum and static functional connectivity), but disrupting the dynamics. Importantly, the observed topography is not a spurious product of the source reconstruction. We verified this by generating random Gaussian processes, projecting them to the sensors using the subject-specific mixing matrices, and then projecting the sensors signals (with added white noise, correlated as 1/sensors distance, with an SNR = 12) back to the source-level (see leakage analysis in Methods). In this case, the pattern observed in the FID matrix (Fig.1.d) does not appear (Fig.S1.a and b).

The SSN and LSN subnetworks are topographically distinct, as shown in Fig.2.a. Edges with short information storage 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 SSN. 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 system, 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 SSN. 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 SSN appears to involve the ventrolateral and dorsolateral prefrontal cortices, which participate to higher functions such as allocation of attention, and are 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 stimuli emerge as a temporally homogeneous fast network carrying information 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 the long storage network (LSN), clusters topographically in frontal and temporal regions. The regions involved in the LSN appear to be linked anatomically by the uncinate fasciculus. In fact, the LSN 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). This evidence might highlight the functional significance of the LSN for the integration of information across these anatomically linked regions. 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 with 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 SSN entails more long-range inter-hemispheric connections as compared to the LSN, 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 hypothesized that regions that are equally important in both networks would likely be the ones that allow functional integration between the SSN and the LSN, 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 (i.e., with MPC values close to zero; see Table 1) are the Heschl's gyri, the Rolandic opercula and the superior and inferior frontal gyri, bilaterally. These regions are primarily concerned with language comprehension and retrieval, as well as secondary somatic and motor functions (45–47). Furthermore, the posterior cingulate cortex also belongs bilaterally to the MSC. This structure is part of the default mode network and has been linked to cognition and conscious awareness (48). 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 external information). 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 (52). The MSC stands out as a potential network linking timescales that are proper of the environment (via bottom-up pathways) with intrinsic brain timescales

(proper of top-down processes). It should be stressed that this work is based on resting-state data and, as a consequence, the functional meaning of the results is speculative. While we interpret the occurrence of co-activations as communication between regions, this is an inference based on the statistical properties of the signals, and further modelling and animal study should confirm the functional interpretation.

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 interactions across regions concerned with stimulus perception, or across regions dedicated to hierarchically higher processes. A specific network hinges both time-scales. These subnetworks are found during rest, and might constitute the spatio-temporal scaffolding underpinning the constant readiness to new stimuli.

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 evidence 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 (50). In short, the position of four coils and

399 of four reference points (nasion, right and left pre-auricular point and apex) were
400 digitized before acquisition using Fastrak (Polhemus®). The brain activity was recorded
401 for 7 minutes, with eyes closed, with a break at 3.5 minutes, so as to minimize the chances
402 of drowsiness. The head position was recorded at the start of each segment. The data were
403 sampled at 1024 Hz and a 4th order Butterworth band-pass filter was applied to select
404 components between 0.5 and 48 Hz. During the acquisitions, electrocardiogram (ECG)
405 and electrooculogram (EOG) were also recorded (51) . These steps were done using
406 Matlab 2019a and the Fieldtrip toolbox 2014 (52).

407 **Preprocessing**

408 Principal component analysis (PCA) was performed to reduce the environmental noise
409 (53, 54). Noisy channels were removed manually through visual inspection of the whole
410 dataset by an experienced rater. Supervised independent component analysis (ICA) was
411 performed to eliminate the ECG (and the EOG) component from the MEG signals (55).
412 Trials that did not contain artefacts (either system related or physiological) or excessive
413 environmental noise were selected.

414 **Source reconstruction**

415 The data were co–registered to the native MRI. A modified spherical conductor model
416 was used as a forward model (56). The voxels in the MRI were labelled according to the
417 Automated Anatomical Labelling (AAL) atlas (57, 58). We used the cortical regions for
418 a total of 78 areas of interest. Subsequently, a linearly constrained minimum variance
419 beamformer was used to compute 78 time series (one per area of interest) at a sampling
420 frequency of 1024 Hz (59). 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 (20). This allowed us to characterize the whole brain network activity in terms of dynamical non-local interactions, highlighting the relational properties of each couple of nodes. 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.,

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

where $\bar{X}$ and $\sigma(X)$ denote the mean and variance of the signals, respectively. One can interpret the edge co-activation 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) statistical dependencies which take into account nonlinear interactions, which are ubiquitously found in brain data (60, 61)P. 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 (62, 63), 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. It should be mentioned that there exists no unique estimator for information storage. We chose AMI since we were interested in an interval estimate, rather than a point estimate (see for example (64)).

Information Storage Capability of the functional edges

For each co-activation signal E_{ij}, we estimated the AMI profile (Fig.1.b) and we evaluated the time delay τcorresponding to the AMI reaching 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 S2. 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 (Fig.1.d). The AMI analysis of the co-activations shows that the decay times are different among edges, as revealed by the histogram in Figure 2.b (blue). Selecting the edges from either the left or right tails of the distribution leads to the appearance of two topographically organized subnetworks (Fig.2.a).

Surrogate analysis

Leakage analysis

We designed surrogate analysis to exclude that linear mixing alone might spuriously generate the patterns observed in the FID matrix. To this end, we generated for each subject n Gaussian time-series, with n = number of regions. Then, the subject-specific leadfield matrix was used to reconstruct the sensor signals for each subject. White noise correlated as 1/*sensor distance*, was added to the sensor time series with SNR = 12. Following this step, the sensor time series were inverted and new surrogate source-level time series were generated. On these source-level surrogates, we have computed the edge time-series and the FID matrix as described previously.

Surrogate data analysis

Firstly, we sought to investigate if the observed decay times might be derived by the spectral nodal properties alone. To this end, we generated n gaussian processes, with n = number of regions, we fourier-transformed them, and we multiplied the resulting power spectra by the power spectra of the observed time-series. Finally, we anti transformed and obtained surrogate time-series that are independent gaussian processes endowed with the same spectral power as the original data (G- surrogates). Secondly, we sought to investigate if the time decays convey a dynamical feature of the time-series or, alternatively, if they can be explained by static correlations. Hence, starting from the original data, we generated surrogates preserving not only the nodal spectral properties, but also the static functional correlation. To this end, we shifted by a random phase (extracted from a unimodal distribution) each frequency of the Fourier transformed nodal signals. The same shift was uniformly applied to each region. Hence, we obtained new surrogates that preserve the functional connectivity while not showing the dynamic of the original data.

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 (Fig.3.a). 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 binary matrices.

In the former, only the fast-decaying edges are included, defining a short storage network (SSN). The latter, containing the slow-decaying edges, identifies a long storage network (LSN), which is complementary to the SSN. Finally, we defined a two-layered multiplex network, with the layers corresponding to the SSN and the LSN. 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 either layer. The MPC, for a two-layered network, is defined as:

$$MPC_i = \frac{k_i^{LSN} - k_i^{SSN}}{k_i^{LSN} + k_i^{SSN}}$$

where k_i^{SSN} and k_i^{LSN} denote the degrees of the i-th node in layer SSN and LSN, respectively. The MPC is calculated per each node and is bounded between -1 and 1, where $-1(+1)$ indicates that all the edges that are incident upon a given node belong to the SSN (LSN), and 0 denotes a node whose edges are equally shared with the two layers (Fig.3.b, top). The above definition is a modified version of the MPC as described in (23).

Parameter independent definitions

The above definitions of the SSN and LSN 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 (Fig.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). We evaluated the MPC for each density value (Fig.3.b, bottom). Averaging across all the density values, we obtained the MPC for each brain region (Fig.3.b, bottom right). The same pipeline was applied

starting from the AMI minima at the single trial level, showing comparable results (Fig.3.c).

Null Model

To statistically validate the MPC results extracted from the aggregate FID matrix, 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, every node should have equal probability of connecting within each layer. To test this hypothesis, we based the following analysis on the original definition of the MPC provided in (23), whereby regions with MPC of 0 have all edges within a single layer (either SSN or LSN), while an MPC equal to 1 means that edges are equally split across layers. Hence, a random organization of the time-delays 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.a, each time starting from a random surrogate of the FID matrix (Fig.S4, left). The surrogates were obtained by randomly shuffling the FID matrix across rows and columns, while 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.S4, right), 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. S. Bassett, *et al.*, Dynamic reconfiguration of human brain networks during learning. *Proc. Natl. Acad. Sci.* **108**, 7641–7646 (2011).
5. 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).
6. 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).
7. 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).
8. 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.
9. G. Hahn, A. F. Bujan, Y. Frégnac, A. Aertsen, A. Kumar, Communication through Resonance in Spiking Neuronal Networks. *PLoS Comput. Biol.* **10**, 1003811 (2014).
10. P. Fries, Rhythms For Cognition: Communication Through Coherence HHS Public Access. *Neuron* **88**, 220–235 (2015).
11. G. Buzsaki, Neuronal Oscillations in Cortical Networks. *Science* **304**, 1926–1929 (2004).
12. 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).
13. 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).
14. A. S. Pillai, V. K. Jirsa, Symmetry Breaking in Space-Time Hierarchies Shapes Brain Dynamics and Behavior. *Neuron* **94**, 1010–1026 (2017).
15. J. D. Murray, *et al.*, A hierarchy of intrinsic timescales across primate cortex. *Nat. Neurosci.* **17**, 1661–1663 (2014).
16. R. Gao, R. L. van den Brink, T. Pfeffer, B. Voytek, Neuronal timescales are functionally dynamic and shaped by cortical microarchitecture. *eLife* **9**, e61277 (2020).
17. J. D. Murray, *et al.*, Stable population coding for working memory coexists with

heterogeneous neural dynamics in prefrontal cortex. *Proc. Natl. Acad. Sci.* **114**, 394–399 (2017).

18. G. Shafiei, *et al.*, Topographic gradients of intrinsic dynamics across neocortex. *eLife* **9**, e62116 (2020).
19. E. J. Müller, *et al.*, Core and matrix thalamic sub-populations relate to spatio-temporal cortical connectivity gradients. *NeuroImage* **222**, 117224 (2020).
20. F. Z. Esfahlani, *et al.*, High-amplitude cofluctuations in cortical activity drive functional connectivity. *Proc. Natl. Acad. Sci.* **117**, 28393–28401 (2020).
21. S. Baillet, Magnetoencephalography for brain electrophysiology and imaging. *Nat. Neurosci.* **20**, 327–339 (2017).
22. F. Lopes da Silva, EEG and MEG: Relevance to neuroscience. *Neuron* **80**, 1112–1128 (2013).
23. F. Battiston, V. Nicosia, V. Latora, Structural measures for multiplex networks. *Phys. Rev. E - Stat. Nonlinear Soft Matter Phys.* **89**, 032804 (2014).
24. K. Mandke, *et al.*, Comparing multilayer brain networks between groups: Introducing graph metrics and recommendations. *NeuroImage* **166**, 371–384 (2018).
25. M. Vaiana, S. F. Muldoon, Multilayer Brain Networks. *J. Nonlinear Sci.* **30**, 2147–2169 (2020).
26. D. J. C. Mackay, “Information Theory, Inference, and Learning Algorithms” (1995) (July 7, 2020).
27. 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).
28. A. K. Engel, P. Fries, W. Singer, Dynamic predictions: Oscillations and synchrony in top–down processing. *Nat. Rev. Neurosci.* **2**, 704–716 (2001).
29. 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).
30. K. Friston, The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* **11**, 127–138 (2010).
31. A. M. Bastos, *et al.*, Visual Areas Exert Feedforward and Feedback Influences through Distinct Frequency Channels. *Neuron* **85**, 390–401 (2015).
32. 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.
33. P. Tewarie, *et al.*, Tracking dynamic brain networks using high temporal resolution MEG measures of functional connectivity. *NeuroImage* **200**, 38–50 (2019).
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. R. Leech, D. J. Sharp, The role of the posterior cingulate cortex in cognition and disease. *Brain* **137**, 12–32 (2014).
49. X. Teng, X. Tian, D. Poeppel, Testing multi-scale processing in the auditory system. *Sci. Rep.* **6**, 1–13 (2016).
50. 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).
51. J. Gross, *et al.*, Good practice for conducting and reporting MEG research. *NeuroImage* **65**, 349–63 (2013).
52. 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.
53. A. de Cheveigné, J. Z. Simon, Denoising based on time-shift PCA. *J. Neurosci. Methods* **165**, 297–305 (2007).
54. Sadasivan P.K., Narayana Dutt D., SVD based technique for noise reduction in electroencephalographic signals. *Signal Process.* **55**, 179–189 (1996).
55. 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).
56. 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).
57. 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).
58. 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).
59. 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).
60. M. Paluš, Nonlinearity in normal human EEG: cycles, temporal asymmetry, nonstationarity and randomness, not chaos. *Biol. Cybern.* **75**, 389–396 (1996).
61. C. J. Stam, Nonlinear dynamical analysis of EEG and MEG: review of an emerging field.

Clin. Neurophysiol. Off. J. Int. Fed. Clin. Neurophysiol. **116**, 2266–2301 (2005).
62. 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).
63. 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).
64. M. Wibral, J. Lizier, S. Vögler, V. Priesemann, R. Galuske, Local active information storage as a tool to understand distributed neural information processing. *Front. Neuroinformatics* **0** (2014).

Figures and figure captions

AAL Atlas Human

IDs	ROIs	MPC	IDs	ROIs	MPC
1	Left Rectus	0.33	40	Right Rectus	0.35
2	Left Olfactory	0.27	41	Right Olfactory	0.3
3	Left Frontal Sup Orb	0.22	42	Right Frontal Sup Orb	0.29
4	Left Frontal Med Orb	0.23	43	Right Frontal Med Orb	0.28
5	Left Frontal Mid Orb	0.15	44	Right Frontal Mid Orb	0.17
6	Left Frontal Inf Orb	0.13	45	Right Frontal Inf Orb	0.12
7	Left Frontal Sup	-0.01	46	Right Frontal Sup	0.02
8	Left Frontal Mid	-0.01	47	Right Frontal Mid	0.02
9	Left Frontal Inf Oper	0.01	48	Right Frontal Inf Oper	0.03
10	Left Frontal Inf Tri	0.04	49	Right Frontal Inf Tri	0.03
11	Left Frontal Sup Medial	0.12	50	Right Frontal Sup Medial	0.13
12	Left Supplementary Motor Area	-0.03	51	Right Supplementary Motor Area	-0.07
13	Left Paracentral Lobule	-0.11	52	Right Paracentral Lobule	-0.1
14	Left Precentral	-0.1	53	Right Precentral	-0.08
15	Left Rolandic Operculum	0.05	54	Right Rolandic Operculum	0.01
16	Left Postcentral	-0.14	55	Right Postcentral	-0.09
17	Left Parietal Sup	-0.12	56	Right Parietal Sup	-0.1
18	Left Parietal Inf	-0.13	57	Right Parietal Inf	-0.09
19	Left SupraMarginal	-0.06	58	Right SupraMarginal	-0.1
20	Left Angular	-0.11	59	Right Angular	-0.12
21	Left Precuneus	-0.07	60	Right Precuneus	-0.04
22	Left Heschl	0.07	61	Right Heschl	0.06
23	Left Temporal Sup	0.08	62	Right Temporal Sup	0.04
24	Left Temporal Mid	0.06	63	Right Temporal Mid	0.03
25	Left Temporal Inf	0.11	64	Right Temporal Inf	0.02
26	Left Temporal Pole Sup	0.16	65	Right Temporal Pole Sup	0.07
27	Left Temporal Pole Mid	0.22	66	Right Temporal Pole Mid	0.19
28	Left ParaHippocampal	0.18	67	Right ParaHippocampal	0.14
29	Left Fusiform	0.08	68	Right Fusiform	0.04
30	Left Cingulum Ant	0.21	69	Right Cingulum Ant	0.21
31	Left Cingulum Mid	0.002	70	Right Cingulum Mid	-0.02
32	Left Cingulum Post	0.04	71	Right Cingulum Post	0.04
33	Left Insula	0.06	72	Right Insula	0.06
34	Left Occipital Sup	-0.16	73	Right Occipital Sup	-0.17
35	Left Occipital Mid	-0.09	74	Right Occipital Mid	-0.17
36	Left Occipital Inf	-0.04	75	Right Occipital Inf	-0.08
37	Left Calcarine	-0.1	76	Right Calcarine	-0.11
38	Left Cuneus	-0.14	77	Right Cuneus	-0.14
39	Left Lingual	-0.01	78	Right Lingual	-0.06

Table 1. List of regions of AAL Atlas. Each region number is colored according to the lobe. Color codes are reported in Fig.2.a. The MPC values of the regions are reported.

Supplementary Material

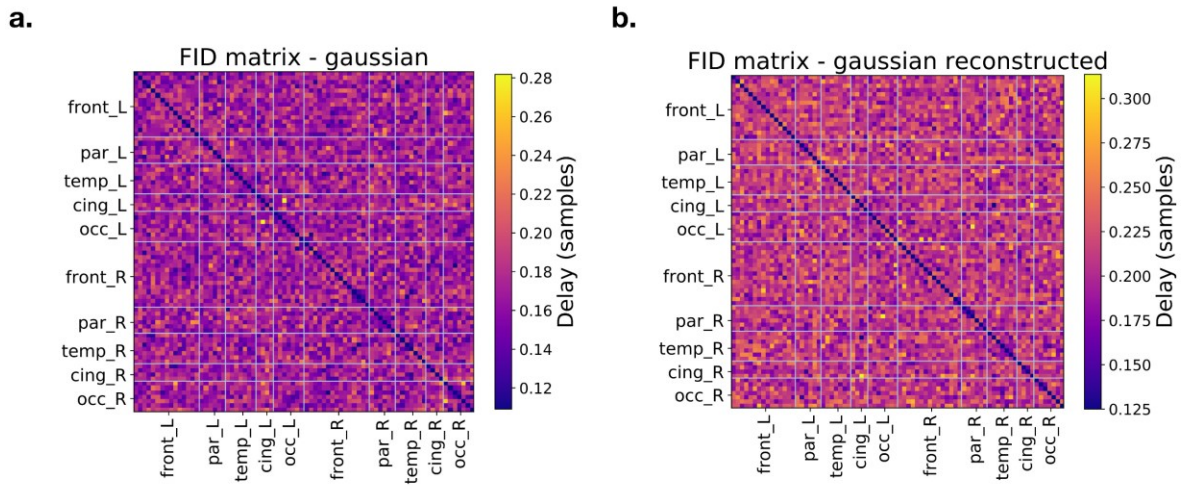

Fig. S1 Leakage analysis pipeline. **a.** FID matrix computed from Gaussian surrogates. The decay times were estimated from random surrogates analogously to the procedure used with the original data. Hence, rows and columns represent brain regions, while matrix entries are the corresponding decay minima. **b.** Starting from the Gaussian surrogates to the left, the surrogates were outprojected, for each subject, using the corresponding mixing matrix, as to obtain new sensor-level surrogates. White noise, correlated as 1/*distance between sensors*, was added to the sensor-level surrogate with a SNR =12. The sensor-level surrogates were then inverted and new source-level data were generated. The minima of the decay of the AMI were the recomputed on the newly obtained source-level surrogates, to check that the spatial mixing of otherwise independent processes did not introduce the spatial patterns observed in the empirical data.

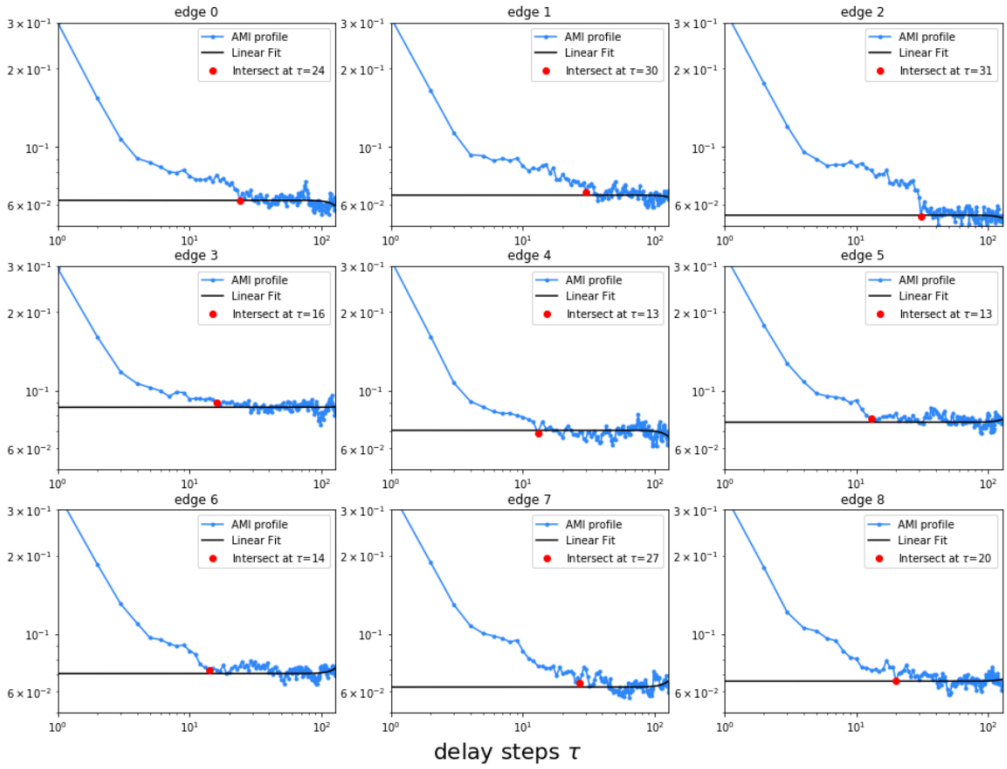

Fig. S2: Examples of the AMI decays of 9 randomly selected edges in a single trial (i.e., form the same 10 seconds-long epoch from a randomly selected subject) (log-log scale). In order to estimate the time after which the co-activation time series and its delayed version become maximally independent, 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.

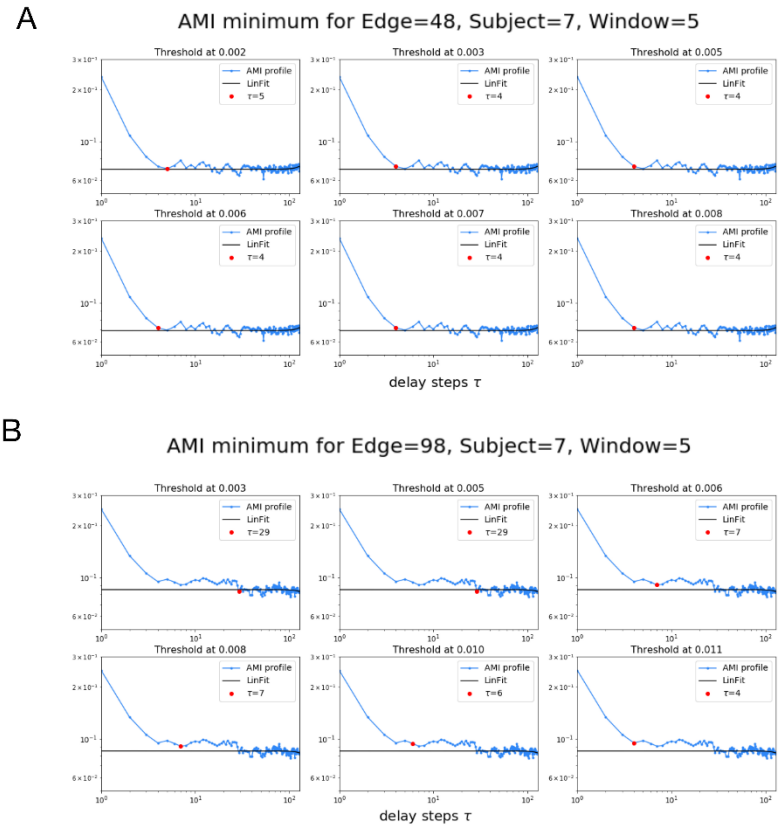

Fig. S3: The blue lines show the AMI decay profile for a single epoch. The red dots represent the estimate of the decay-times. The decay-time is measured as the first time that the AMI profile approaches the null-line (fitted to the tail of the AMI distribution, in black). The distance from the null-line, 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.

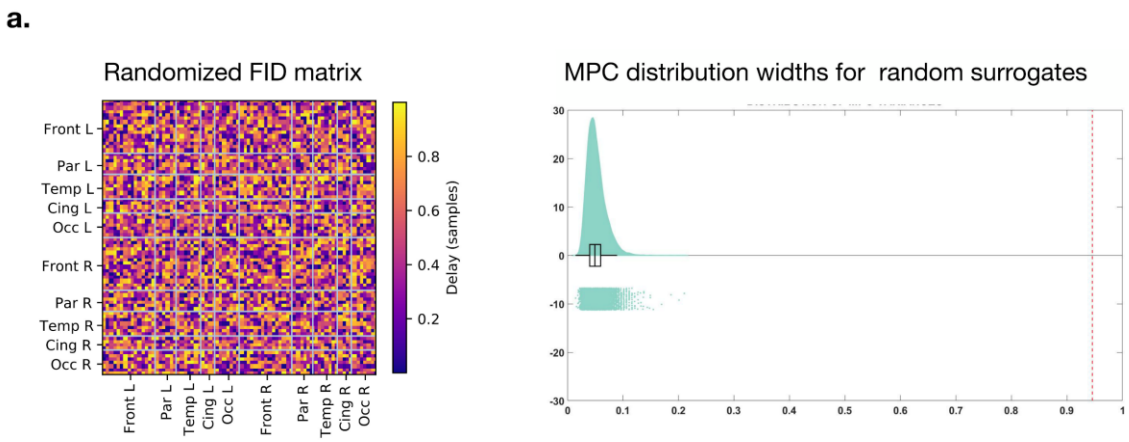

Fig. S4. Deviation from a random null model. To check the statistical significance of the results, the MPC was calculated on 10000 randomizations of the FID matrix (e.g., left). In the null models, the distribution width across regions of the MPC ($\in [0,1]$) is very low, meaning that, at random, all the regions equally belong to both layers. That is, they have the same probability of sharing an edge with either the SSN and the LSN. 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.