Article

Leveraging multivariate information for community detection in functional brain networks

Check for updates
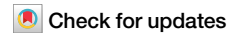
Maria Grazia Puxeddu [1,2] ✉, Maria Pope[3,4], Thomas F. Varley [1,5], Joshua Faskowitz[1] & Olaf Sporns[1,3,4]

Brain functioning relies on specialized systems whose integration enables cognition and behavior. Network science provides tools to model the brain as a set of interconnected brain regions wherein those segregated systems (modules) can be identified by optimizing the weights of pairwise connections within them. However, knowledge alone of these pairwise connections might not suffice: brain dynamics are also engendered by higher-order interactions that simultaneously involve multiple brain areas. Here, we propose a community detection algorithm that accounts for multivariate interactions and finds modules of brain regions whose activity is maximally redundant. We compared redundancy-dominated modules to those identified with conventional methods, uncovering a new organization of the transmodal cortex. Moreover, by identifying a spatial resolution where within-module redundancy and between-module synergy are maximally balanced, we captured a higher-order manifestation of the interplay between segregation and integration of information. Finally, we distinguish brain regions with high and low topological specialization based on their contribution to within- or between-module redundancy, and we observed how redundant modules reconfigure across the lifespan. Altogether, the results show a modular organization of the brain that accounts for higher-order interactions and pave the way for future investigations that might link it to cognition, behavior, or disease.

Emergent properties of complex systems arise from a balance between segregation and integration of the system's fundamental units[1]. A prominent example of such a system is the brain, which displays ongoing transitions between segregated and integrated activity[2–8]. On the one hand, functional specialization pushes the brain to segregate functionally related groups of neurons, neural populations, or brain areas. On the other hand, the integration of these systems promotes global communication required for coherent perception and behavior. Thus, identifying the principles able to model and recapitulate this interplay is a key goal of computational cognitive neuroscience.

The balance between local segregation and global integration can be viewed through the lens of network science. Under the network conceptualization, the brain can be modeled as an ensemble of distributed brain regions (nodes) linked by anatomical connections (anatomical networks) or dynamic interactions (functional networks)[9,10]. Segregation in functional brain networks has been studied primarily by observing how nodes are organized into modules[11,12]. Network modules, also referred to as

communities or clusters, are groups of nodes that form strong (or dense) connections to one another and weak (or sparse) connections to other nodes in the network. Strongly connected brain modules map out functional systems that are often invoked as building blocks of cognition and behavior[13]. Complementing brain modules are network hubs—nodes highly interconnected with the whole network—which enable information transmission between modules and thus functional integration[14,15].

The identification of modules, commonly called community detection, is usually conducted using the so-called functional connectivity matrix–the matrix formed by measuring covariance between pairs of neural elements. While valid, this approach is also limited in scope. By design, most community detection techniques search for groupings of elements according to pairwise similarities, without taking into account that more than two neural units can (and do) engage with each other, enabling higher-order interactions[16–18]. An increasing body of literature centers on these higher-order interactions as key features of complex systems, including the brain[19–23]. This surge in interest demands new tools able to reveal brain

[1]Department of Psychological and Brain Sciences, Indiana University, Bloomington, IN, USA. [2]Department of Computer, Control and Management Engineering "Antonio Ruberti", Sapienza University of Rome, Rome, Italy. [3]Program in Neuroscience, Indiana University, Bloomington, IN, USA. [4]Luddy School of Informatics, Indiana University, Bloomington, IN, USA. [5]Vermont Complex Systems Institute, University of Vermont, Burlington, VT, USA. ✉e-mail: mpuxeddu@iu.edu

organizational principles by leveraging this augmented space of possible interactions.

In the current paper, we tackle this issue by introducing a new framework to model and investigate how segregation emerges from higher-order interactions of brain regions. We build our work exploiting recent advances in information theory, which provides the mathematics to describe polyadic dependencies in multivariate systems[24]. Given a set of brain regions and their activity portrayed through time series, information theory can help us discern two types of interactions: redundant and synergistic[25]. While *redundant information* captures how much information is copied across the elements (i.e., brain regions), *synergistic information* is the information that is accessible only by considering the state of all the elements as a whole[26]. We postulate that the way the brain organizes itself into segregated subsystems can be captured by higher-order interactions within subsets of brain regions whose activity is maximally redundant, jointly sharing the same information. How information is integrated over the whole brain system can then be estimated by quantifying the level of synergy between elements of different subsets.

For this purpose, we developed a community detection algorithm that finds maximally redundant modules at multiple spatial scales. We assessed how they relate to canonical functional systems and to the modules identified with conventional methods that only consider pairwise interactions. We also quantified the extent to which the information carried by the subsets is integrated over the entire system through synergistic interactions. Moreover, after describing where these subsets are located in the cortex, we also defined a new local index of segregation and integration, based on subset interactions. Finally, we applied our algorithm to a human lifespan fMRI dataset to observe how this higher-order redundancy-dominated modular structure evolves from youth to senescence. Collectively, our work represents a methodological advancement towards a more comprehensive characterization of organizational principles in multivariate systems, and its application provides new insights into how the brain organizes itself into functionally segregated subsystems.

Results

Numerous studies investigated the modular structure of functional brain networks. One challenge is to move beyond the pairwise representation to reveal a more realistic organization of the brain's dynamics that includes multivariate interactions among multiple brain areas. Here, we address this issue by introducing an algorithm that groups together brain areas that share a significant amount of identical information in their neural activity, that is, we group together maximally redundant brain areas.

Throughout the analyses we used functional connectivity (FC) data derived from resting-state fMRI recordings of three different datasets: the Human Connectome Project (HCP)[27], the Microstructure-Informed Connectomics (MICA-MICs) data[28], and the Nathan Kline Institute (NKI) lifespan data[29]. In each dataset, the cerebral cortex was parcellated into 200 nodes[30] and covariance matrices encoding FC were derived from the BOLD time series. Because the data can be described as a multivariate Gaussian[31], we relied on Gaussian assumptions to estimate multivariate information theoretic measures from the covariance matrices. See further details on the datasets and signal retrieval methods on the "Methods" section.

Community detection via total correlation maximization

The most popular methods for community detection group together nodes based on the density and strength of their pairwise interactions. Among these, modularity maximization[32], implemented with the Louvain algorithm[33], is the most widely used in neuroscience applications. It is based on a quality function called 'modularity' (Q) that measures the goodness of a partition by counting how many connections fit within a given set of modules compared to chance, i.e., what we would observe in a null network model partitioned in the same way. The Q heuristic can be optimized outright to discover modular community structure, and the optimization can be tuned to find modules at different spatial scales, from coarser to finer (see also "Methods" "Multiscale modularity maximization"). While

modularity maximization has been proven useful in linking brain network topology to function and behavior, the communities that it finds do not account for higher-order interactions among subsets of nodes and therefore might miss important multivariate aspects of brain organization. Here, we overcame this issue by introducing a community detection algorithm that leverages higher-order interactions.

Information theory proposes the total correlation (TC) as a proxy for redundancy[1,34–36]. Given a multivariate system, TC is low when every element is independent and high when the joint state of the elements has low entropy, given the individual entropies, i.e., the system is dominated by redundant interactions. For a full mathematical description of TC see Eqs. (7) and (13) in "Methods". We hypothesized that modules are made of maximally redundant brain areas, that is, sets of nodes that share a large amount of information (see Fig. 1A for a schematic). Thus, we introduced the total correlation score (TC_score), a new quality function that, given a partition of the brain into subsets, estimates the redundancy of the modules compared to chance:

$$TC_{score} = \frac{1}{N} \sum_{m=1}^{M} (TC_m - E[TC_m]). \tag{1}$$

with N being the number of nodes, M the number of modules, TC_m the total correlation within module m, and $E[TC_m]$ the mean level of integration for randomly selected subsets of the same size of module m. Thus, TC_score is higher the more the TC within modules exceeds the average TC of equal size subsets. High TC_score values entail a good partition, whereas for particularly poor partitions, or very weakly integrated systems, this quality function can even be negative.

Analogously to Q, TC_score can be used to assess the goodness of a partition, or it can be maximized to infer an optimal modular structure. Here, we implemented an optimization algorithm based on simulated annealing, for which we report the pseudocode in Table 1. The first step of the algorithm consists of computing the TSE complexity curve[1]. This provides the maximum and mean levels of integration for any subset size for the network that we want to partition (all possible E[TC_m] in Eq. (1), and see Fig. 1B, C). Starting from a random (seed) partition then, we can compute the TC of the subsets identified by this partition and compute TC_score. At this point, the algorithm randomly switches node assignments (without changing the number of modules) and uses simulated annealing to search in the space of solutions for a partition that maximizes TC_score.

In Fig. 1D we report an example of what we obtained by running the algorithm on the HCP data. We projected on the cortex an initial random partition and the partition that optimizes TC_score. Analogously, we display in Fig. 1C the TC of the initial random modules and the TC of the optimized modules, which is much higher compared to the average TC of subsets of equivalent size. Finally, we ran the algorithm 100 times, varying the number of modules from 2 to 12. We computed the similarity between the 100 optimized partitions at each resolution in terms of variation of information (VI)[37]. Low VI values suggest that the algorithm delivered partitions that were highly similar and consistent across multiple attempts (Fig. 1E).

Relation between TC_{score} and Q

How do the two heuristics TC_score and Q, relate to each other? To answer this question, we applied multiresolution consensus clustering (MRCC)[38] to the HCP FC covariance matrix and we computed the TC_score on the resulting partitions. MRCC uses the Louvain algorithm to optimize Q at different spatial scales, thus providing partitions made of finer and coarser modules. By running MRCC, we obtained 990 partitions made of a number of modules within the range [2, 50] (see also Methods "Multiscale modularity maximization"). Each one of these partitions is associated with a Q-value that resulted from the optimization, which we compared to the TC_score values computed on the same partitions. TC_score and Q are significantly and strongly correlated ($r = 0.91$; $pval < 10$–15; Fig. 2A). However, if we divide the ensemble of data points into groups where partitions have the same number of modules, TC_score and Q result anti-correlated (Fig. 2B). This

means that TC_{score} and Q are linked by a Simpson's paradox: overall the two variables are positively correlated, but this correlation is reversed when the ensemble is divided into groups. Specifically, both TC_{score} and Q values increase for coarser partitions (i.e., partitions made of fewer modules), but if we consider partitions with equal levels of granularity the partition with higher TC_{score} is the one with lower Q and vice versa.

We explored what might cause the anticorrelation within partition groups. For each set of partitions with an equal number of modules we considered the two partitions with the highest and lowest TC_{score} (Fig. 2C) and computed the variance of the modules size and the symmetry of the modules between the hemispheres. The variance of modules size was computed as the standard deviation of the size of the modules (i.e., how many nodes belong to one module) so that higher values mean that the sizes of modules are heterogeneous within the partition. To assess the hemispheric symmetry of a partition, we calculated, for each module, the difference between the number of nodes in the left and right hemispheres, normalized by the total number of nodes in the module. The complement of the average of this measure across all modules serves as a proxy for hemispheric symmetry (HS, Eq. (2)). Higher HS values indicates greater symmetry in the partition.

$$HS = 1 - \frac{1}{M}\sum_{m=1}^{M}\frac{|N_{LH} - N_{RH}|}{N_{LH} + N_{RH}}. \qquad (2)$$

where N_{LH} and N_{RH} are the number of nodes in module m in the left and right hemispheres, respectively. The results, reported in Fig. 2D, E, show that

the partitions with the lowest TC_{score} (higher Q) consistently present modules with greater variance and a lower hemispheric symmetry, compared to the partitions with the highest TC_{score} (lowest Q).

Altogether, these findings suggest that our newly introduced TC_{score} fairly recapitulates Q in some aspects. However, TC_{score} rewards modular structures that are more symmetric between cortical hemispheres and composed of modules of similar size.

Lastly, although widely used, modularity maximization is not the only community detection algorithm. Thus, we carried out a supplementary analysis to compare TC_{score} to partitions obtained via a hierarchical version of the Stochastic Blockmodel (SBM)[39], which yields more general classes of partitions, including assortative, disassortative, and core-periphery (Fig. S2).

Relation with the canonical systems

In the previous section, we used TC_{score} as a quality function to assess the goodness of partitions found with modularity maximization to see how well our metric recapitulates conventionally detected modular structures. Here, we use our algorithm to maximize TC_{score} on FC networks. One question when inferring a new modular structure in the human brain is how it relates to resting state functional systems (RSFS). Throughout the neuroimaging literature, different labels have been applied to RSFC communities, nonetheless, the patterns of organization have been largely consistent[40–42]. However, all these studies have always overlooked higher-order interactions among brain regions, at least in terms of redundancy and synergy. In this section, we report how the modules derived from TCscore maximization,

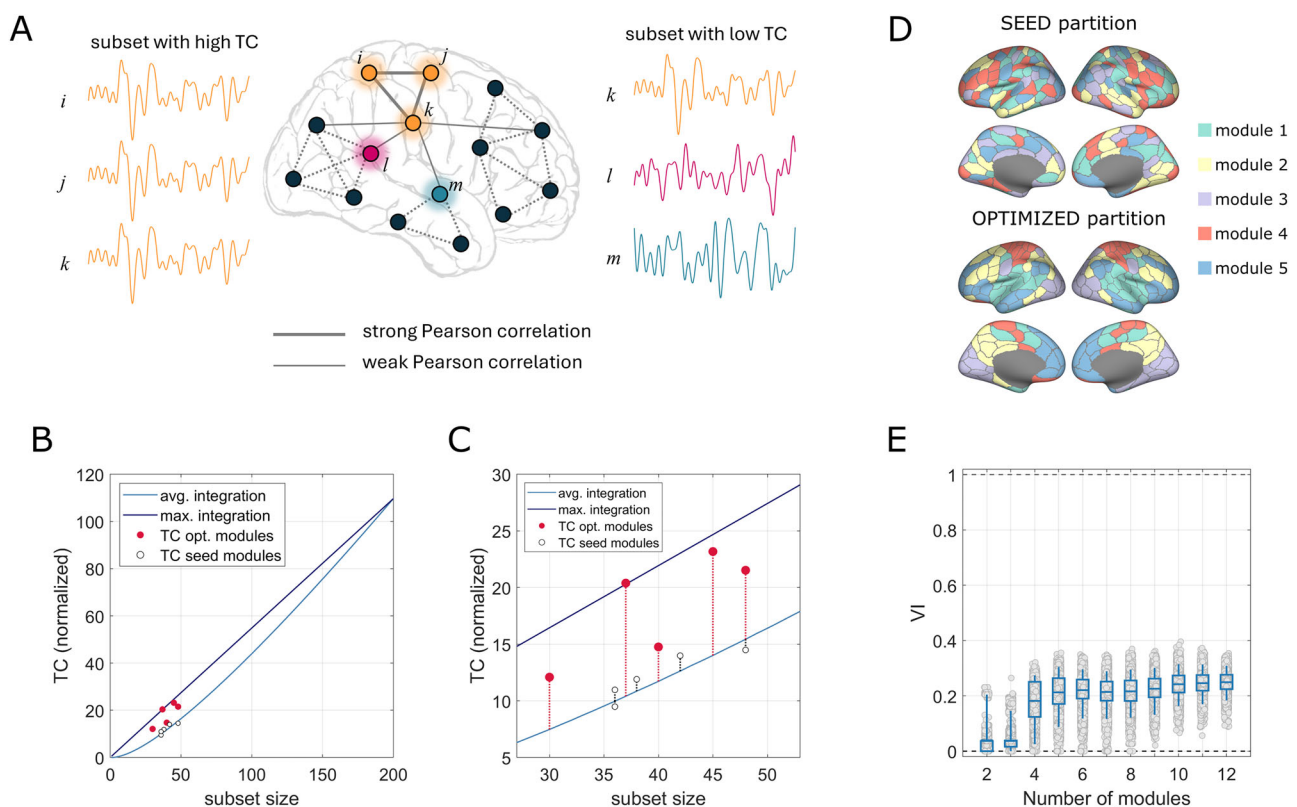**Fig. 1 | Schematic of the algorithm. A** Whereas with modularity maximization nodes are grouped together based on the strength of their pairwise connections (stronger connections fall within a module), with our algorithm we group together nodes if the TC of their time series is higher, on average, than the TC of an equally sized subsets of nodes. In the figure the yellow nodes constitute a module that satisfies both criteria for community detection. **B** TSE curves for the HCP covariance matrix. For any subset size, they provide an estimate of the average (blue) and maximum (purple) levels of integration of the nodes in the subsets in terms of total correlation. By partitioning the system into modules, we can compute the TC of the modules as the TC of the nodes constituting each module. Then, we can map them in

the same graph to see where these values are located with respect to the average level of integration for subsets of the same size. The white dots are the TC of the modules of a random partition, whereas the red dots are the TC of the modules of the optimized partition. **C** A zoomed in version of panel B, which clarifies that in the optimized partition the modules have a much higher level of integration. **D** An example of a random seed partition and the partition obtained after the optimization, projected on the cortex. **E** How similar the partitions obtained from different runs of the optimization are, measured at each spatial scale by using variation of information (VI) as the similarity metric.

Table 1 | Pseudocode of our algorithm for community detection via total correlation optimization

• compute TSE curve
• for $it = 1$ through $it = ITER$
• plant a random seed partition C_0
• compute TC_{score} on $C_0 \rightarrow score_0$
• for h=0 through h=H
• define temperature $T = T_0 \times (1 - h_{frac}/H)^h$
• switch module assignment for a random node $i \rightarrow C_i$
• compute new TC_{score} on $C_i \rightarrow score_i$
• if $score_i > score_0 \,\&\&\, \exp\left(\frac{score_i - score_0}{T}\right) > rand(0, 1)$
• update partition $\rightarrow C_0 = C_1$
• update $TC_{score} \rightarrow score_0 = score_i$

We set the parameters as follows: H = 100,000 (number of iterations in the annealing process); ITER = 100 (number of times we perform the optimization through annealing); h_{frac} = 10; T_0 = 1.
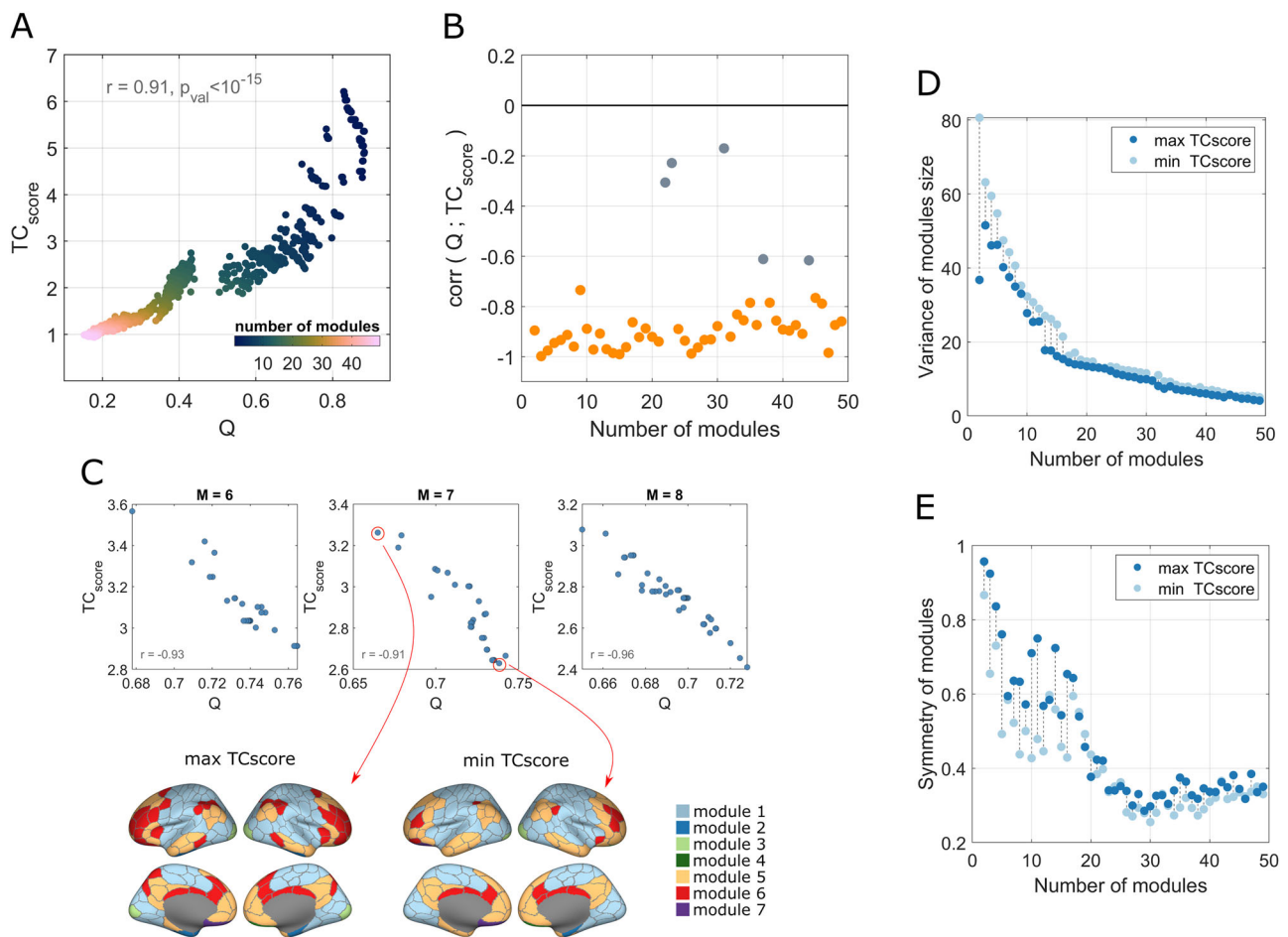

Fig. 2 | Relation between TC_{score} and Q. A Trend of TC_{score} relative to Q. Each point corresponds to a partition obtained optimizing Q on the HCP FC network, and on which we computed TC_{score} downstream. Colors code the spatial resolution of the partitions indicating the number of clusters. **B** For each group of partitions with an equal number of modules, we computed the correlation between TC_{score} and Q and reported here the correlation coefficients. Points colored in orange identify statistically significant correlations. **C, D**. Values of variance of the modules size and relate to previously reported RSFS, and specifically to well-established symmetry between hemispheres at each spatial resolution for the two partitions with highest and lowest TC_{score}. **E** Example of how the positive correlation observed in (**A**) is reversed if we consider groups of partitions with equal resolution (or equal number of modules). In this case, we report the reversed correlation for groups of partitions with 6, 7, and 8 modules. We also report a projection on the cortex surface of the modular structure of the 7-module partitions with the highest and lowest TC_{score}.

relate to previously reported RSFS, and specifically to well-established canonical systems[41].

We ran our algorithm initializing it with partition seeds characterized by a different number of communities (M = [2, 12]), so that also the number of communities of the optimized partitions varied across spatial scales. For each spatial resolution we ran the algorithm 100 times and, for each one of these iterations, we computed the similarity between the detected partition and the partition corresponding to the canonical systems. As an index of similarity, we used the Adjusted Mutual Information (AMI)[43], a measure that theoretically accounts for differences in the number of modules
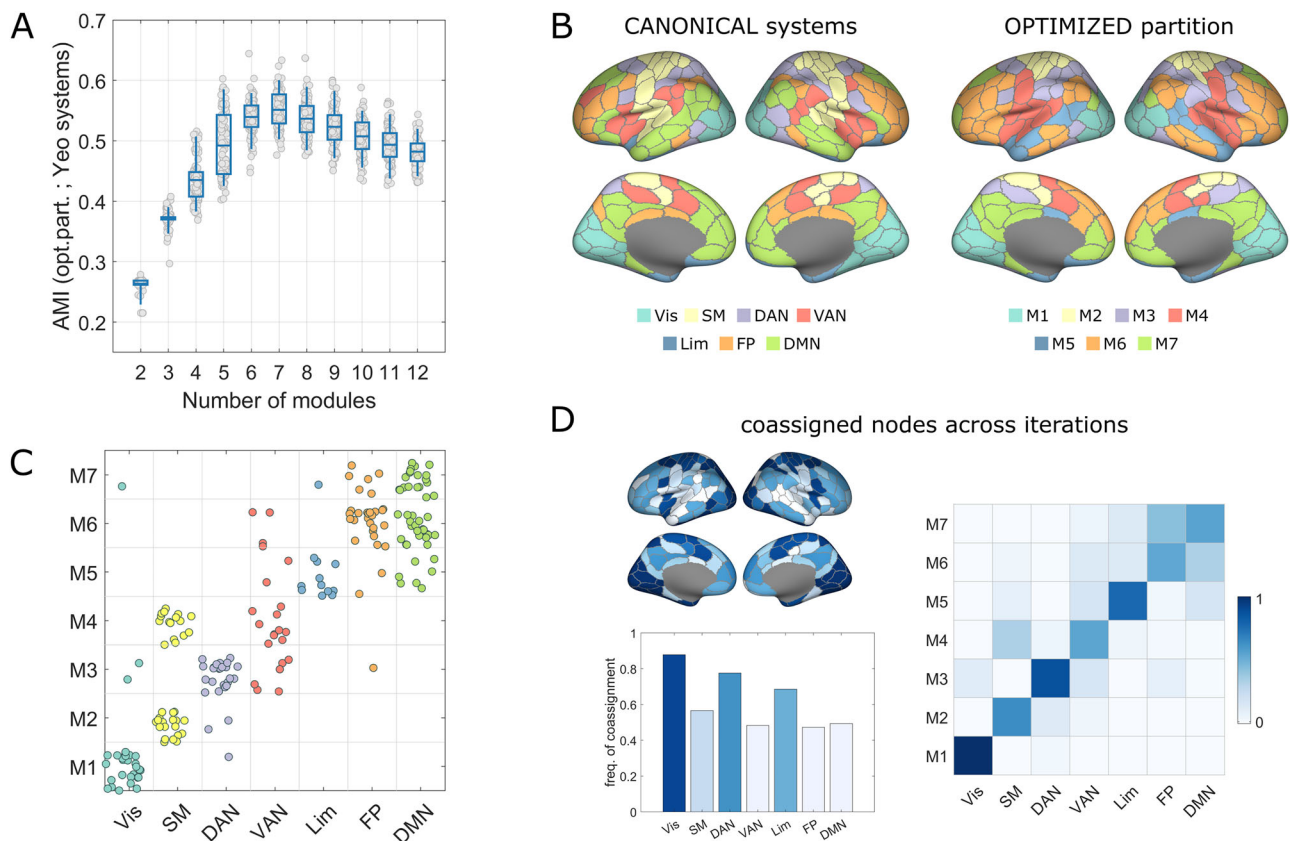
Fig. 3 | Relation with the canonical RSFS. A Similarity between the partitions inferred by maximizing TC_{score}, and the canonical systems in terms of Adjusted Mutual Information. The boxplots summarize the statistics of the 100 iterations of the algorithm at each spatial scale. **B** Projection on the cortex surface of the canonical systems (left) and the partition obtained optimizing TC_{score} (right). For the representation of the latter, we report the centroid partition (the partition most similar to all the others within the 100 iterations of the optimization algorithm) with 7 modules. **C** In this panel, each dot is a node and we show to which module/system it is assigned in the two partitions shown in (**B**). The colors follow the RSFS classification. **D** We further assessed the overlap across all 100 partitions obtained from the 100

optimization iterations. To do so, each partition was first aligned with the canonical systems by assigning numbers (1–7) to the systems and relabeling partitions to maximize node overlap between modules and their corresponding systems. We report on the cortex and through a bar plot the number of times nodes are coassigned between the canonical systems and their corresponding module across the 100 partitions. On the right, we present a visualization similar to (**C**), where we computed the overlap between each module and each canonical system, averaged across all iterations. Darker squares indicate a greater overlap between modules across iterations.

between two partitions. The AMI is bounded in the range [0, 1], with higher values indicating higer similarity (AMI = 1 means identical partitions). As reported in Fig. 3A, at each resolution the optimized partitions present some level of overlap with the canonical systems (the AMI is never 0), but at the same time, they are never identical (the AMI is never 1). As expected, the maximum similarity is obtained when M = 7, that is the number of the canonical systems.

Next, we investigated the relation of our partitions to canonical systems in more detail. We focused on TC-optimized partitions with 7 modules. A first visual comparison can be drawn from Fig. 3B, where we represented a projection on the cortex of the two partitions. We noticed that modules largely overlap, with some differences. We further mapped *how* they overlap by module/system in Fig. 3C, demonstrating that the greatest correspondence between module assignment is found for the visual system, the dorsal attention network (DAN), and the limbic system, which are well captured by modules 1, 3, and 5 of our optimized partition. The somatomotor system (SM) splits into modules 2 and 4, and part of it is co-assigned to the ventral attention system (VAN). Interestingly, recent studies highlighted high levels of co-activations between the two[44–46]. The default mode network (DMN) and frontoparietal networks (FP) instead are mixed into modules 6 and 7, suggesting that association areas are redistributed when considering higher-order interactions in the community detection process. These results are consistent among the 100 iterations of the algorithm (Fig. 3D), and are mostly replicated in the MICA dataset (Fig. S1, A–D).

Altogether these results suggest that partitions inferred by considering higher-order interactions via TC_{score} maximization partially preserve the organization recapitulated by the familiar canonical systems. However, they also suggest a different organization for regions comprising higher order association systems.

Redundancy dominated modules across spatial scales

In this section, we provide a more thorough description of the modular structure inferred by maximizing TC_{score}, analyzing how it balances segregation and integration of information at different spatial resolutions. For this purpose, we used the same set of optimized partitions of the previous analysis. This set comprises 100 instances for each spatial scale (identified by the number of clusters, M = [2, 12]), derived from the HCP FC covariance matrices.

First, we investigated how TC_{score} varies across spatial scales. It showed a decreasing trend with respect to M (Fig. 4A), with a peak in correspondence of partitions made of 3 modules. This means that, as the partitions become finer and modules smaller, modules are comprised of nodes carrying less redundant information. Thus, because the modules are less well-defined in terms of their redundant information content, we can say that the entire brain system becomes less segregated. The peak observed for three modules is a direct consequence of the TSE diagram in Fig. 1B where the maximum difference between average and maximum integration levels occurs for subsets of approximately 70 nodes—the expected size of three
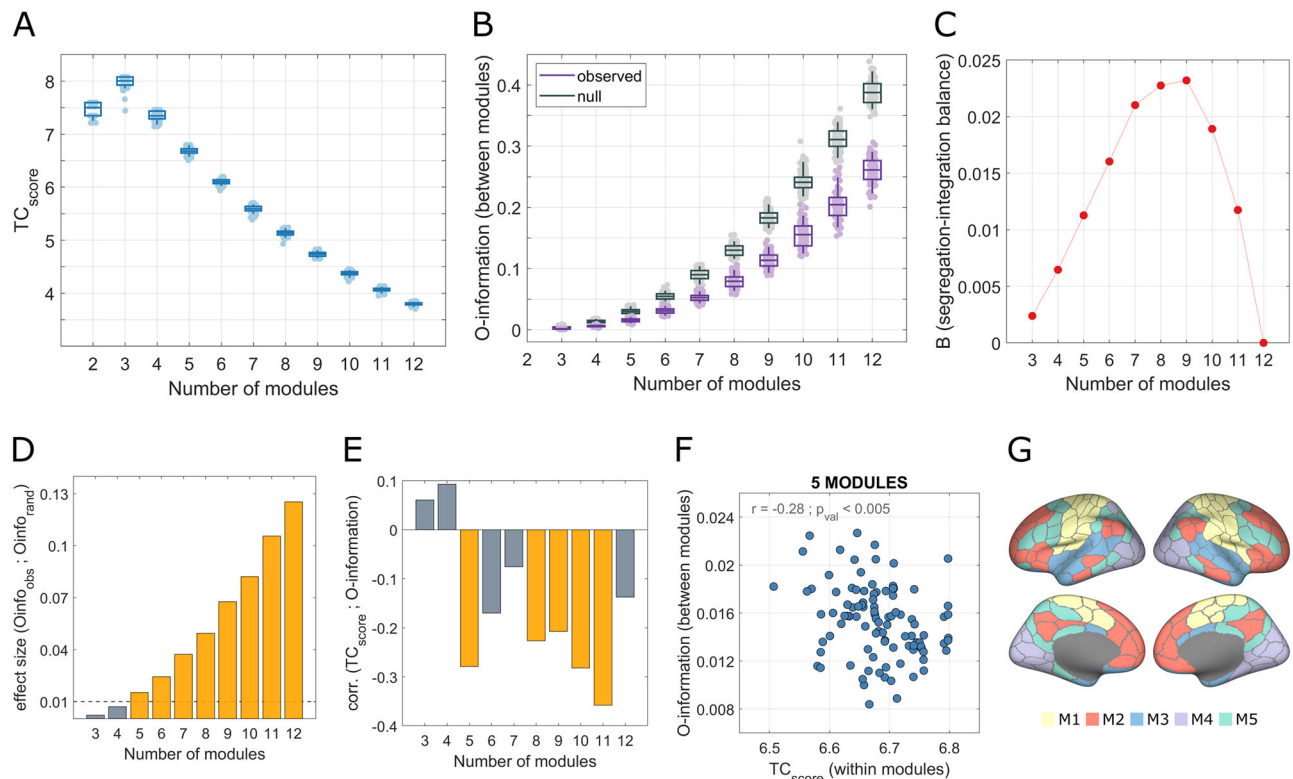
Fig. 4 | Redundancy and synergy across spatial resolutions. A Trend of TC_{score} versus the number of clusters. The statistics to display the boxplots were carried out considering 100 iterations of the optimization algorithm for each spatial scale. **B** Trend of O-information versus number of clusters in the optimized partitions (purple) and in randomized nodes (gray). **C** Trend of B, the index that recapitulates the balance between redundancy within modules and synergy between modules (segregation-integration balance in a higher order space), across spatial resolutions. **D** Median difference effect size of the observations from the two distributions of empirical and null model-derived O-information. **E** Correlation coefficients obtained by computing the Pearson correlation between TC_{score} and O-information among the 100 iterations at each spatial scale. Statistically significant correlations have been highlighted in yellow. **F** Example of such correlation for M = 5. **G** Projection on the cortex of the 5-module partition with highest TC_{score} and lowest O-information.

clusters in a 200-node network. We further analyzed the relationship between the covariance matrix, the leaf shape of the TSE curve, and TC_{score} in Fig. S3.

Then, we focused on how the information carried by single modules is integrated throughout the system at different spatial scales. We hypothesized that an informationally optimal community structure is one where modules carry distinct information that becomes accessible when observed together. This phenomenon is well captured by the notion of synergy, which we computationally quantify through the organizational information, or O-information[24]. In a multivariate system, the O-information represents the overall balance between redundancy and synergy. Synergy-dominated structures present negative values of O-information, with lower values indicating a higher degree of synergistic information (for details see "Methods" "Synergistic and redundant information in multivariate systems"). To have an estimate of how synergistic information manifests across modules, we randomly sampled nodes from different modules, for a total of 1000 times for each spatial scale, and computed the O-information on the obtained subsets. For instance, in partitions made of 7 modules, we sampled 7 nodes, each belonging to a different module, and computed the O-information in such subset. We carried out the same analysis on a null model, where we sampled the same number of nodes but at random, without considering the modular structure. Whereas TC_{score} is defined relative to a null model (the average level of integration in the TSE curve), the O-information lacks a similar comparison in its formulation, thus necessitating this extra null model comparison. We observed lower O-information in coarser partitions, however, in finer partitions, the O-information deviates more from the null model (Fig. 4B). We quantified this deviation by computing the effect size (ES) as the difference between the

means of the observed and the null O-information (Fig. 4D). This metric captures that in finer partitions the amount of synergistic information retrievable in a subset of nodes is more closely linked to the community structure - nodes from different modules actually carry distinct information - whereas at a coarser scale we might obtain lower O-information just because it is computed on a smaller number of elements.

Next, we investigated the relationship between TC_{score} and O-information for specific spatial scales. For any given resolution, we computed the correlation between the two variables across the 100 iterations (Fig. 4E, G). For most of the scales, TC_{score} and O-information was significantly anti-correlated: the partitions that have maximally redundant modules are also those where the synergy between modules is the greatest. This balance of redundancy within and synergy between modules can be view as an expression of the co-existence of segregation and integration in a higher-order space.

Once we observed the trends of redundancy and synergy across scales, we asked: Is there an optimal scale where redundancy within and synergy between modules are maximally balanced? To answer this question, we examined the interaction between TC_{score} and O-information. While these indices can be combined in multiple ways, here we defined the segregation-integration balance index (B) as follows:

$$B = TC_{score}[(1 - O \text{ } information)ES] \tag{3}$$

Ideally, we want the B to be as high as possible: when the three factors are high, we have maximum redundancy within and maximum synergy between modules. In other words, each module is made of nodes that share a large amount of information (high TC_{score}) and carry different information

to the network (high $(1 - \text{Oinformation}) \times ES$). Tracking this index along spatial resolutions (Fig. 4C) resulted in a bell shape with a peak in M = 9 and highest values in the range M = [7, 9]. This is a direct consequence of what observed in Fig. 4A, B. The redundancy within modules is higher in coarser partitions. The synergy between modules is also higher in coarser partitions, however, it deviates more from what we would observe in a null model (measured by ES) as the partitions become finer. Thus, it is intuitive that a combination of redundancy and synergy, that captures the higher-order balance between segregation and integration, can be observed at intermediate scales, when partitions are neither too coarse nor too fine.

Altogether, these results suggest that redundancy and synergy of brain activity within and between modules, and their interplay, vary across spatial resolution. An optimal balance has been found in partitions composed of 7 to 9 modules. All the findings are replicated in the MICA dataset (Fig. S1, E–H).

Local properties of the redundancy dominated modular structure
Having defined globally in which way the new modular structure balances redundant and synergistic information, we now want to characterize its local properties. Specifically, we want to investigate how much single brain regions participate in the redundant integration of the module to which they belong, relative to how much they participate in the redundant integration of the whole network. This concept is analogous to what is commonly known in network science as *participation coefficient*[47], but here we include the consideration of higher-order interactions. While computing nodal measures in bivariate networks is straightforward, it becomes more difficult when considering multivariate measures such as TC, which by definition, cannot be computed on single nodes.

To address this issue, we defined the Relative Integration Coefficient (RIC), which can be computed for each node i, and that is formulated as follows:

$$RIC_i = \frac{TC_M - TC_{M-i}}{TC_N - TC_{N-i}} \qquad (4)$$

Where TC_M is the TC computed on the module M to which node i belongs, TC_{M-i} is the TC of the same module when node i is removed, TC_N is the TC computed on the whole network, and TC_{N-i} is the TC computed on the whole network after removing node i. The idea is that, even if we cannot compute TC on single nodes, we can look at how much the TC increases (or decreases) when we remove node i from the modules and the network, in order to quantify how much it participates in the redundancy of its module compared to the redundancy of the whole system. High RIC values indicate that the node contributes to the redundancy of the module more than to the redundancy of the network (if we remove it, the TC of the module becomes low, hence the numerator is high). The role of these nodes is to share information predominantly with nodes of the same module, meaning they exhibit high topological specialization. This is particularly true when RIC > 1. Conversely, low RIC values indicate that a node either scarcely contributes to its module's redundancy or significantly to the network's overall redundancy (i.e., its removal increases the network's TC, lowering the denominator). These nodes share information across the entire network, indicating low topological specialization.

In Fig. 5A we displayed the RIC distribution computed on the centroid partition of the 100 instantiations with 7 modules. It shows a left and right tails of nodes with low and high topological specialization, respectively. In panels B–D, we mapped those nodes onto the cortex relative to their canonical systems. Nodes with high topological specialization are primarily found in the visual and somatomotor systems, i.e., in the sensory areas. Conversely, nodes with low topological specialization are mostly located in the VAN and DMN, i.e., in the transmodal and association cortex (although the DMN also includes a few nodes with high topological specialization). This distribution aligns with their role of integrating information from different areas to support cognitive functions. While most nodes in the limbic systems exhibit very low RIC values, this is mainly due to their

minimal contribution to redundancy, either in the module and in the whole network (see panel C), which is expected given the inherently noisy nature of the signals from that region of the brain. This local mapping is consistent across the 100 iterations of the algorithm and spatial scales (Fig. 5E, F). Moreover, a replication on the MICA dataset shows coherent results (Fig. S1, I–K).

Ultimately, with this local analysis, we could assign a role to single nodes, derived from their participation in higher-order interactions relative to the modular structure. We were able to identify brain areas more devoted to sharing information with similar nodes (i.e., nodes belonging to the same community) as well as brain areas whose primary purpose is integrating information across the whole network.

Redundancy dominated modular structure across the human lifespan
Lastly, we wanted to test the ability of TC-based community detection to capture variations in functional connectivity organization that could naturally arise when considering clinical cohorts, or large-scale multi-subject datasets. To this purpose, we applied our new methodology to a lifespan dataset, as the progressive modifications exhibited during development, adulthood, and senescence, are one of the most striking and well-documented examples of functional connectivity re-organization[48–50].

We leveraged the NKI dataset[29], comprised of 917 subjects with ages heterogeneously distributed between 6 to 85 years (Fig. S4, A). For each one of the subjects, we performed 100 iterations of the TC_{score} optimization at different spatial resolutions, obtaining partitions with a number of modules between 2 and 10. For consistency with previous analyses, we report the results for M = 7 (for each subject we considered the centroid partition after iterating the algorithm 100 times). However, results are consistent across spatial scale (Fig. S5) and specifically replicate for M=9, which corresponds to the resolution with optimal balance between segregation and integration (Fig. S6).

Here, we show how theoretical information measures computed on such partitions change with age. Note that, due to the shorter time series in the NKI dataset, the range of these measures differs from that observed in previous analyses on the HCP and MICA datasets. For a thorough examination of how information theory measures depends on the length of time series see Supplementary Material (Fig. S7), and the Technical Consideration section in the Discussion. Here, we invite the reader to focus on the trends over the lifespan rather than absolute values.

First, we inspected how redundancy within modules and synergy among them vary with age (Fig. 6A, B). We found that TC_{score} significantly decreases with age ($r_{(TC_{score}, age)} = -0.70$, $pval_{(TC_{score}, age)} < 10^{-15}$), whereas O-information between modules significantly increases ($r_{(Oinformation, age)} = 0.41$, $pval_{(Oinformation, age)} < 10^{-15}$). This result suggests that aging is associated with a decrease of segregation of the network, a phenomenon largely observed in previous studies[12]. Moreover, aging is associated with less stability of the algorithm (Fig. S8), which might indicate that it is more difficult to retrieve maximally redundant modules in older adults.

Next, we computed the RIC-values of each node, thus mapping age-related variations of the extent to which a node participates to the redundancy of its module. We found that the average RIC values decrease with age (Fig. 6C), meaning that later in life nodes participate more to the integration of the network, rather than individual modules. Unpacking the RIC's trend for each canonical system (Fig. 6D), we observed that nodes in the somatomotor and visual systems, previously classified as nodes with high topological specialization, have the steepest decline of RIC, which means that they drive the progressive integration of the network. On the other hand, nodes from the DMN and FP networks (low topological specialization) show a modest decline of RIC. Nodes from other canonical systems, not reported in the figure, presented an intermediate pattern.

A final question is how much and in what way the modular structure changes across the lifespan. To answer this question, we divided the dataset into 8 age groups, each spanning 10 years. Analogously to what has been
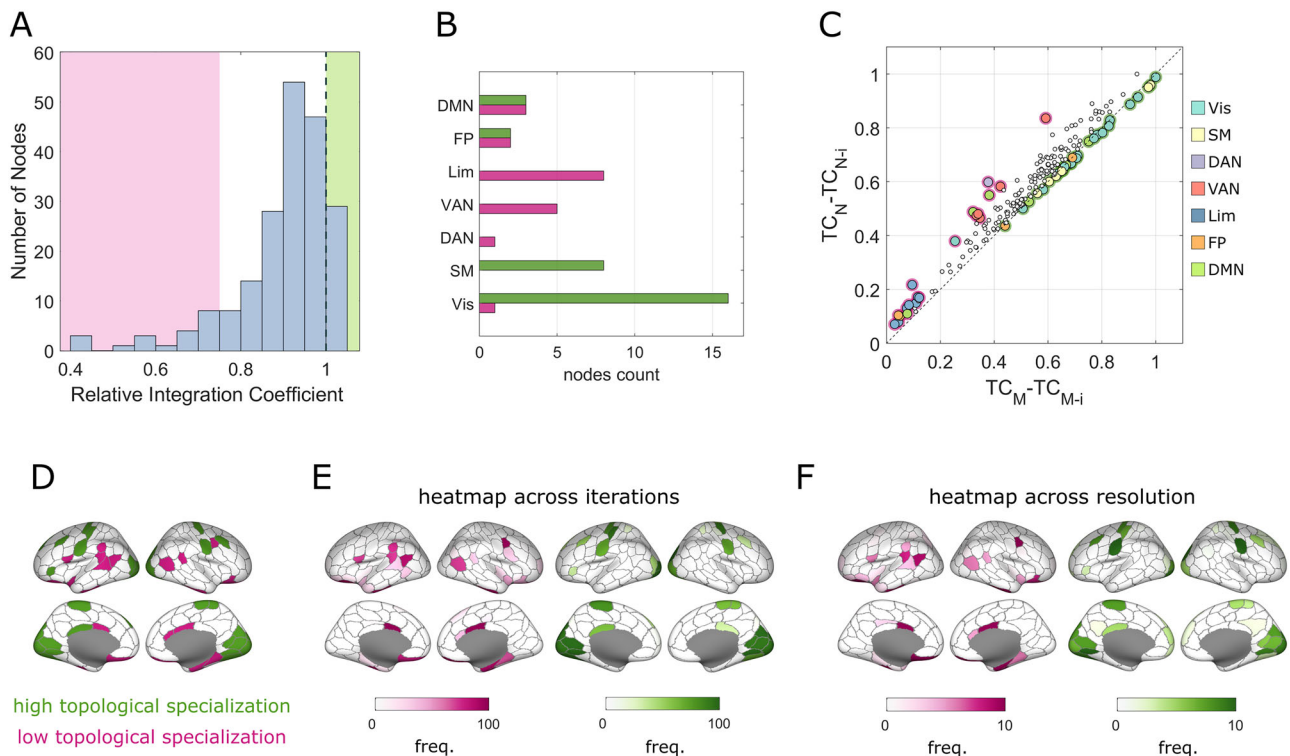
Fig. 5 | Relative integration coefficient. A Histogram of relative integration coefficient computed for each node. The area where RIC is greater than 1, denoting nodes with high topological specialization, has been highlighted in green, whereas the area containing 10% of the lowest RIC values, denoting nodes with low topological specialization, has been colored in pink. B Count of nodes with high and low topological specialization within the canonical systems. C Trend of $TC_N - TC_{N-i}$ versus $TC_M - TC_{M-i}$. These two measures are highly correlated, as a node contributing to its module's redundancy also inherently contributes to the network's

redundancy. Nodes with high and low topological specialization are color-coded according to their canonical system, with their contour indicating their functional role. D Projection on the cortex of the nodes with high (green) and low (pink) topological specialization in the centroid partition with 7 modules. E, F Projection on the cortex of the frequency with which nodes have resulted with high or low topological specialization among the 100 partitions with 7 modules (C) and the centroids partitions across spatial scales (D). The intensity of the color is proportional to that frequency.

done in ref. 51, we used a bootstrap strategy where for 500 iterations we sampled subjects in each group, to then identify which nodes change allegiance to modules across consecutive groups. This can be quantified via flexibility[52], a measure that captures how often a node changes module assignment across partitions. Here, we aligned the partition labels across age-groups, computed the nodal flexibility between consecutive groups, and averaged the results over the 500 iterations. We observed that the nodes belonging to the visual and somatomotor systems are the most stable across the lifespan (Fig. 6E, F). We also averaged this measure within the canonical systems (Fig. 6G, H), finding that in the late lifespan, the reconfiguration of the modular structure is greater than at the beginning (i.e., the entropy increases). Most of the systems reconfigure more with age, except for the DMN and the FP networks, which keep the most stable levels of nodal flexibility.

Overall, these results suggest that our newly introduced algorithm is able to track changes in functional organization. The sensory areas overlap with the most stable modules. However, the changes that they present are aimed at integrating the whole system. The DMN, FP and attention networks are those that reconfigure the most, but they keep stable levels of integration with the whole network from the beginning of the lifespan to the end.

Discussion

In this paper, we proposed an algorithm that uncovers modular structure in functional brain networks by accounting for higher-order interactions. Our approach leverages the mathematical framework of information theory to find highly redundant subsets of nodes. We provided a new view of brain organization and showed how it contrasts with canonical functional systems

and with the modular structure uncovered with traditional methods based on pairwise interactions. The discovery of a Simpson's paradox relating TC_{score} and modularity suggests that, when higher-order interactions are accounted for, the notion of 'structure' in complex systems becomes much richer than is typically observed in purely bivariate systems. This may be a fruitful new perspective, as the Simpson's paradox typically reflects nontrivial interactions between elements of a system, when viewed relative to certain bounds. We also presented two new indices: a global index that estimates how the redundant modules balance segregation and integration of information across the whole network, and a local index that characterizes node diversity in terms of contribution to network integration. Finally, we delineated the evolution of redundant modular structure across the human lifespan.

Redundancy dominated systems of the human brain

Decades of neuroimaging research have been devoted to defining functional systems of the human brain. Task-based experiments identified such systems as collections of co-activated brain areas[53,54]. More recently, the idea that brain function is intrinsic[55], constantly processing information for interpreting and responding to environmental demands, led to a description of functional systems of highly correlated spontaneous activity[40–42]. Proof that such systems are grounded in neurophysiology is widespread. Not only are they largely consistent across studies and methodologies, but some of them also match patterns of brain damage observed in clinical cohorts[56]. However, there exist co-fluctuating areas that have not been firmly aligned with independently validated brain systems[12]. Whether they correspond to undefined brain systems or are artifacts induced by network construction is yet to be determined. We argue that accounting for higher-order
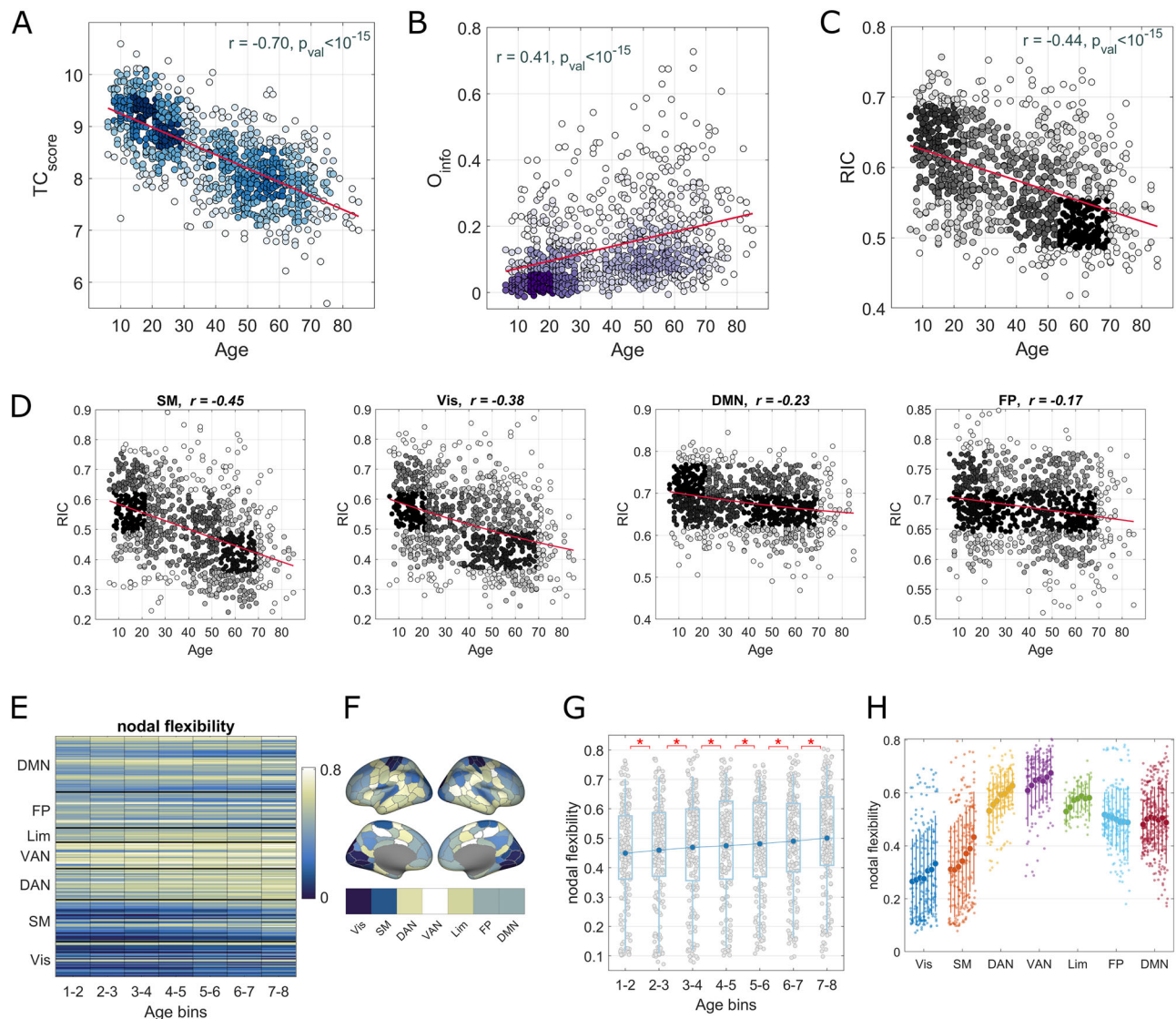
Fig. 6 | Application to the NKI dataset. A Values of the TC_{score} for the partitions inferred at the single-subject level, versus age. **B** Values of O-information between modules computed on single-subject partitions versus age. **C** RIC values averaged across nodes for each subject versus age. **D** RIC values averaged within the canonical systems, for each subject, versus age. Throughout panels A-D, because the population is not uniformly distributed across age, we represented values in age bins with higher population density with darker colors. **E** With the nodal flexibility we depict how much single nodes change module assignment across the lifespan. Yellow colors indicate high entropy, i.e., nodes that are more likely to change module, and vice versa colors toward blue indicate nodes whose module assignment is stable. **F** Projection on the cortex surface of the nodal flexibility, averaged across age-groups. **G** Distribution and average of nodal flexibility reported for each age group. There is a statistically significant difference between consecutive distribution; p_{values} of the six paired t-test: [0.0014, 0.0003, 0.0466, 0.0391, 0.0002, 0.0107]. **H** Values of nodal flexibility reported for each canonical systems and each age group. All systems show an increase of nodal flexibility (i.e., reconfiguration of module assignment) across the seven age groups, except for the Limbic system, the FP, and DMN.

interactions is an avenue to unravel more complex relationships among brain areas. This can lead to a different perspective on brain organization and elucidate the role of nodes that are more difficult to place in the classical framework.

We compared redundancy dominated modules to canonical resting-state systems. While some level of similarity was expected (canonical systems are dominated by redundant interactions[57]), we highlight how they differ. Our method led to a clear identification of the visual and somatomotor systems, and the dorsal attention network, which are also the systems more consistently detected across reports. However, the somatomotor system is split into two modules. This is consistent with a dorsal and a ventral representation of such system, reported in ref. 40. In our partition, some brain areas normally associated with the ventral attention system are co-assigned to nodes of the ventral part of the somatomotor area.

Interestingly, previous reports showed co-fluctuations among those regions[44–46]. This is an example of how beyond pairwise interactions offer a different angle on how co-activated brain regions could form coherent systems, still grounded in neurophysiology. As for the brain areas belonging to the frontoparietal system and DMN, these converge into two redundant modules where they are mixed. Future investigations will clarify the role of these new systems during tasks, cognitive functions, or disease progression.

A higher-order lens on segregation and integration of brain systems

The modular organization of the human brain supports segregation and integration of information among functional brain systems[3], enabling cognition and behavior. Recent studies[58] found proof that local within-systems connectivity is critical for motor execution, and between-

systems connectivity is crucial for cognitive tasks. Furthermore, brain organization benefits from the coexistence of segregation and integration. While lack of segregation hampers functional specialization and fails to protect the system from the spreading of disease[59], a complete disconnection between brain systems is associated with brain disorders (e.g., Alzheimer's[60] and Parkinson's[61] disease, or schizophrenia[62]) and leads to cognitive deficit[63].

In our work, we explored how the balance between segregation and integration is expressed when considering higher-order interactions. Notably, the TSE complexity[1], that inspired our TC_{score}, was designed to capture how brain dynamics optimizes segregation and integration of information. However, it did not account for any modular structure, which now we know is central to brain organization. Instead, we hypothesize that the dichotomy of segregation-integration is supported by a mesoscale organization made of subsystems where brain activity is redundant (copied over the system's elements) and whose interaction with each other leads to a new type of information, which is synergistic, that goes beyond functional and local specialization. Interestingly, we found these two measures to be anti-correlated in the HCP dataset. This suggests that the brain may be organized in a way where increased redundancy within subsystems corresponds to greater synergistic information between them. Spontaneously and even under a higher-order perspective, brain activity manifests in a manner where segregation and integration are maximized. We further characterized the balance between segregation and integration via an index that tracks how information encoded in brain activity is redundant within subsystems and synergistic between them. We explored how this index is conveyed across spatial resolutions, finding that a division of the neocortex into 9 subsystems optimizes the abovementioned balance, immediately followed by partitions into 8 and 7 subsystems.

Nodal diversity in redundancy-dominated modular structure

Given a division of the brain into subsystems, individual nodes can behave differently in how much they interact with nodes of the same system versus how much they interact with other nodes assigned to different systems. In this sense, we might say that they serve distinct functional roles, especially if they are highly interconnected hubs: they either foster local specialization (provincial hubs) or facilitate inter-module communication (connector hubs). The participation coefficient[47,64] was introduced to quantify the extent to which they cover one role or the other. Many reports located connector hubs in the association cortex[14,65,66] highlighting how they are involved in a wide repertoire of tasks and functions[67,68].

All these findings build on a dyadic view of the interactions among brain regions. Recent works[69,70] studied how being a hub affects computation capacity estimated in terms of multivariate information processing, and information theory has also been used to extend the formulation of participation coefficient to the case where multiple sets of partitions can describe the mesoscale organization of a complex system[71]. However, a measure able to capture the role of the nodes in the modular organization, which also accounts for higher-order interactions, is missing. Thus, we introduced the Relative Integration Coefficient. By shifting the concept of participation coefficient to a higher-order space, we identified nodes with high and low topological specialization based on the amount of information shared with co-assigned and non-coassigned nodes. Sensory areas mainly present high topological specialization, whereas regions from the DMN, ventral attention and frontoparietal networks have the lowest topological specialization. Notably, brain regions whose predominant role is to enhance modular redundancy were also found to be the most synergistic areas in the system, regardless of the modular structure[20,23]. A similar distinction between sensory and association areas has also been observed in ref. 72, where the study of co-fluctuations between sets of three and four regions saw the sensory areas as maximally coherent, as opposed to association areas. Future works will elucidate how these maximally synergistic, provincial, coherent areas engage during tasks or in different clinical cohorts.

Brain redundant subsystems have unique trajectories during the lifespan

The application of our algorithm to a lifespan dataset mainly served the purpose of validating its ability to detect changes in modular organization. Nonetheless, understanding how higher-order interactions change across the years, and with them the emerging modular structure, is of crucial importance and here we provided a first glance at it.

Alterations in the organization of brain connectivity have been observed, even in the absence of disease, in anatomical and functional networks, and in their relationship[48-51,73]. Thus, it is safe to assume that multivariate interactions among brain regions are also subject to change over the lifespan. By uncovering redundancy-dominated modules in subjects between 6 and 85 years old, we tested if, and to what extent, our algorithm is able to capture such changes. First, we observed a decrease within-module redundancy, and thus segregation of information, associated with aging. Evidence of this phenomenon is overwhelming in analyses of pairwise connectivity[74-78], but also in those accounting for higher-order interactions[79-81]. The fact that we could replicate this ensures that our algorithm is able to track changes in modular organization. We then extended our analysis by portraying how the heterogeneity of node processing with respect to the modular organization unravels over the lifespan. Nodes with high topological specialization become increasingly less specialized with age, whereas nodes with low topological specialization become only mildly more integrated in the system. This suggests that the decreased segregation is majorly driven by sensory areas that share more and more information with other systems. This might align with a reduction in the distinction between connector and non-connector nodes in less segregated systems[74]. We also described the reconfiguration of redundant brain systems, finding that regions in the dorsal attention and ventral attention networks are more prone to change module assignment. Taken together, these results indicate that sensory areas form brain systems that are the most stable throughout the lifespan, but at the same time, they drive the drop of segregation of information by sharing increasingly more information with regions outside of sensory cortex.

Technical considerations and future directions

A first technical consideration is that all the measures here reported are built upon the covariance matrices of couples of variables, so they come with an assumption of linearity of the data. Talking about redundant and synergistic interactions in this context might sound counterintuitive. However, prior works show that it is possible to search for higher-order interactions in a linear system, only capitalizing on its pairwise relationships. In[82] it has been shown that purely Gaussian systems can present higher-order synergies and that total correlation is tied to mutual information. Thus, polyadic dependencies can be observed even in linear systems. Even from[83] we know that the multivariate Gaussian is the maximum entropy distribution constrained by pairwise covariance. In other words, pairwise linear interactions set the structure upon which beyond-pairwise interactions coexist.

It is also important to note that the covariance matrices used in this paper are derived from BOLD time series of different lengths in the three datasets. In the HCP and MICA datasets we concatenated data from every subject at the node level, obtaining time series with a large number of samples; in the NKI we derived FC matrices at the subject level, from shorter time series. Information theory measures are sensitive to the length of the data on which they are computed, and from the current results, it is clear that they lie in different ranges when computed on the NKI single-subject data. Thus, we pursued a simulation study to test how having fewer data points skews such measures (Fig. S7). We found that TC is overestimated when computed on shorter time series, and this effect worsens the larger the subset size. However, this bias is consistent across iterations such that it does not invalidate our analyses. The effect shows constant magnitude for a given length of time series, which presumably is protracted across subjects in the same way, allowing between-subject comparisons.

Another source of ambiguity is the global signal regression (GSR) performed on the BOLD time series. Previous studies suggest that global

signal influences functional connectivity in a way that is spatially non-uniform, and there is a synergistic interaction between the two signals[84]. Moreover, GSR removes a large amount of redundancy present in the data[20]. Because we leveraged Gaussian assumptions to compute multivariate information theoretical measures on covariance matrices, possible confounds have to be taken into account. Nonetheless, the proposed framework is independent of the specific neural signal and could be replicated on datasets exploiting different neuroimaging modalities.

A possible limitation of our framework is that the number of communities provided as output is determined a priori by the number of communities that we impose in the seed partition. The algorithm, implemented as it is, cannot tune the spatial resolution. We tried to overcome this issue not only by examining redundant modular organization in a broad range of spatial scales, but also by identifying the most relevant scales, characterized by a balance between segregation and integration of information. Besides, an advantage of our framework is that it can be applied broadly to any set of multivariate time series. Its formulation is irrespective of the fact that our field of application is neuroscience, and it can be tested on other complex systems, such as biological[85], ecological[86], and epidemiological[87].

This paper opens the door to a multitude of future investigations. One question is how the balance between higher-order segregation and integration is manifested in clinical cohorts. Several brain diseases are known to lack this balance[88], and multivariate information theory has already been proven relevant in clinical and behavioral contexts[89–91]. Intuitively, given that neurotypical cognitive functioning requires the coordination of multiple brain regions[2,92], beyond pairwise frameworks are needed to discover biomarkers of neurodegeneration onsets. Another line of research revolves around the relationship between structural and functional connectivity. Ebbs and flows of modular organization[93] can be reanalyzed exploiting our measures, introducing sensitivity to higher-order interactions. Finally, one can pursue investigations where our framework is modified to accommodate, for instance, a different quality function that optimizes not only the redundancy within module, but also the synergy between modules. This would require preliminary steps where a graph similar to the TSE is built for the O-information, so that it can also be compared to a chance level. Further modifications can accommodate multilayer investigations. In the case of modularity maximization, having a multilayer formulation[94] has been proven convenient to track modular structure across time[52], subjects[95], or types of connectivity[96] because nodes labeling is consistent across layers. We do believe that a similar advantage is in fact achievable with our algorithm if we optimize individual layers planting the same seed. Future investigation will elucidate the feasibility of this direction.

Methods
Datasets and data processing
The analyses presented in this paper were conducted on three independent resting state fMRI datasets, derived from the Human Connectome Project (HCP)[27], the MICA57 (an open-source repository)[28], and the Nathan Kline Institute Rockland Sample (NKI)[29]. Several previous studies have used those datasets and described in detail their image acquisition and preprocessing pipelines (see for instance[97] for the HCP,[21] for MICA, and[98] for the NKI). In the following paragraphs, we recapitulate them.

HCP. The HCP data were derived from 100 unrelated subjects, who provided informed consent. The study protocols and procedures were approved by the Washington University Institutional Review Board. Resting-state fMRI (rs-fMRI) data were collected with a Siemens 3T Connectom Skyra (32-channel head coil) during four scans in two separate days. A gradient-echo echo-planar imaging (EPI) sequence (scan duration: 14:33 min; eyes open) was used with the following acquisition parameters: TR = 720 ms; TE = 33.1 ms; 52° flip angle; isotropic voxel resolution = 2 mm; multiband factor = 8.

Functional data were mapped into 200 regions covering the entire cerebral cortex, identified via the parcellation scheme proposed in ref. 30

and aligned to the Yeo canonical resting state networks[41]. A rigorous quality control of the images was performed and led to the exclusion of 5 of the 100 unrelated subjects. Exclusion criteria included mean and mean absolute deviation of the relative root mean square (RMS) motion across resting-state fMRI and diffusion MRI scans. Four subjects exceeding 1.5 times the interquartile range of the measure distributions were excluded. One subject was then excluded because of a software error during diffusion MRI preprocessing. Downstream quality control, the data included 95 subjects, 56% female, with a mean age of 29.29 ± 3.66.

Following a minimal preprocessing pipeline described in ref. 99, functional images were corrected for gradient distortion, susceptibility distortion, and motion, and then aligned to the T1-weighted data. The volume was corrected for intensity bias, normalized to a mean of 10,000, then projected onto the $32k\_fs\_LR$ mesh, and aligned to a common space with a multimodal surface registration91. Moreover, global signal regression (GSR), detrending and band pass filtering (0.008–0.08 HZ) were performed[100]. Finally, the first and last 50 frames of the time series were discarded, resulting in a scan length of 13.2 min and 1100 frames.

MICA. The MICA dataset was approved by the Ethics Committee of the Montreal Neurological Institute and Hospital, and includes 50 unrelated subjects who provided written informed consent. Resting-state fMRI data were collected using a 3T Siemens Magnetom Prisma-Fit (64-channel head coil). Data collections included a single scan session of 7 min during which participants were asked to look at a fixation cross. An EPI sequence was executed with the following acquisition parameters: TR = 600 ms; TE = 48 ms; 52° flip angle; isotropic voxel resolution = 3 mm; multiband factor = 6.

Data were mapped in 200 regions of the cerebral cortex following the same parcellation scheme used for the HCP data and described in ref. 30. Functional images were preprocessed as outlined in ref. 28. Briefly, data went through motion and distortion correction, as well as FSL's ICA FIX tool trained with an in-house classifier. Nodes were defined in the FreeSurfer surface and used to project each subject's time series. More details about preprocessing can be found in ref. 101, where the Micapipe93 processing pipeline is thoroughly described. Downstream, GSR was performed, as also done for the HCP data.

NKI. The NKI dataset consists of imaging data from a community sample of subjects encompassing a large portion of the lifespan. All participants gave written informed consent. The study received approval by the Institutional Review Boards of Nathan Kline Institute and Monclair State University. fMRI data were collected with a Siemens Magneton Trio (12-channel head coil), during a scan that lasted 9:40 seconds, where 971 participants were instructed to fixate a cross. Images were acquired using a gradient-echo planar imaging sequence with acquisition parameters set as follows: TR = 645 ms; TE = 30 ms; 60° flip angle; isotropic voxel resolution = 3 mm; multiband factor = 4.

Quality control of the data included excluding subjects if the scans exceeded 1.5 interquartile range in three or more of the following metrics: DVARS standard deviation, DVARS voxel-wise standard deviation, framewise displacement mean, AFNI's outlier ratio, and AFNI's quality index.

Also for this dataset, the images were mapped onto 200 cerebral regions using the parcellation scheme in ref. 30. The fMRI data were preprocessed using the fMRIPrep version 1.1.8[102], which comprises the briefly listed following steps, also described in its documentation. The workflow utilizes ANTs (2.1.0), FSL (5.0.9), AFNI (16.2.07), FreeSurfer (6.0.1), nipype[103], and nilearn[104]. FreeSurfer's cortical reconstruction workflow (version 6.0) was used to skull strip the T1w, which was then aligned to the MNI space. Functional data were slice time corrected and motion corrected, using AFNI 3dTshift and FSL mcflirt, respectively. "Fieldmap-less" distortion was performed by co-registering the functional image to the same-subject T1w with intensity inverted[105] constrained with an average fieldmap template[106], implemented with antsRegistration. Then, boundary-based registration[107] was performed using bbregister to co-register the corresponding T1w.

Motion correcting transformation, field distortion correcting warp, and BOLD-to-T1w transformation warp were concatenated and applied altogether using antsApplyTransforms with Lanczos interpolation. For each functional run, frame-wise displacement[108] was computed using Nipype. The first four frames of the BOLD data in the T1w space were discarded. After following the fMRIPrep pipeline, images were linearly detrended, band-pass filtered (0.008–0.08 Hz), confound regressed, and standardized. Furthermore, spike regressors for frames with motion greater than 0.5mm framewise displacement were applied. Finally, GSR was performed.

Covariance matrices definition

As in ref. 20, the HCP-based empirical analyses presented in this paper were conducted after combining all scans and subjects to obtain a single covariance (or FC) matrix representing a grand-average of the sample. For this purpose, BOLD time series were concatenated at the node level and Pearson correlation was computed for each node pair. The same grand-average FC matrix was computed with the MICA data. With the above-chosen parcellation of the cortex, which rendered 200 nodes representing different brain regions, the average FC matrices of the HCP and MICA datasets were highly correlated ($r = 0.851$, $pval < 10^{-15}$).

For the NKI-based analyses instead, we built FC matrices at the single subject level, as we wanted to be able to track individual differences linked to age. Thus, the Pearson correlation was computed for each pair of nodal time series, for each subject. After building the distribution of the mean FC values across subjects, we discarded those subjects whose mean FC exceeded the 95 percentile of the distribution. This left us with 917 covariance matrices relative to 917 subjects of age between 6 and 85 years old.

Redundant and synergistic information in multivariate systems

In this paper, we quantified higher-order interactions by means of information theory. Specifically, given a multivariate set of signals (e.g., BOLD time series), we used the mathematics provided by information theory to measure how much the information carried by the system is shared among the variable, i.e., redundant information, and how much it banks on each variable's contribution, i.e., synergistic information[25].

As with most measures introduced in information theory, redundancy and synergy capitalize on the basic notion of entropy[109], which quantifies the uncertainty associated with the state of a variable when only its distribution is known. Mathematically, if X is a discrete random variable, and $P(X = x)$ is the probability distribution of its states, the entropy is formulated as follows:

$$H_X = -\sum_{x \in X} P(X) \log_2 P(X) \tag{5}$$

In the case of a bivariate system, where we have access to two random variables X_1 and X_2, entropy is used to compute the Mutual Information (I)[109], which captures how much knowing the state of the variable X_1 reduces the uncertainty associated with the state of X_2. Mathematically, I is formulated as follows:

$$\begin{aligned} I(X_1; X_2) &= H(X_1) + H(X_2) - H(X_1, X_2) \\ &= H(X_1, X_2) - [H(X_1|X_2) + H(X_2|X_1)] \end{aligned} \tag{6}$$

Moving towards multivariate systems then, i.e., systems comprised by more than two variables, we can quantify higher-order dependencies by extending the MI formulation[110]. There exist multiple generalization of MI. Here, we focus on three of them that will be necessary to quantify redundant and synergistic information in neural data. These are: Total Correlation (TC)[1,34], Dual Total Correlation (DTC)[111], and Organizational Information (Ω)[24]. The TC is derived as in Eq. (7):

$$TC(\mathbf{X}) = \sum_{i=1}^{N} H(X_i) - H(\mathbf{X}) \tag{7}$$

where $\mathbf{X} = \{X_1, X_2, \ldots, X_N\}$ is a macro-variable comprising a set of N random variables and $H(\mathbf{X})$ is the joint entropy of $\mathbf{X}$. TC is low when every variable is independent and high when the joint-state of the whole system has low entropy. In other words, TC increases monotonically as the system $\mathbf{X}$ transitions from randomness to synchrony. TC can be used as a measure of redundancy: a multivariate system is dominated by redundant interactions when the variables share a large amount of information, hence, the state of a single variable considerably reduces the uncertainty associated with the state of every other variables, i.e., TC is high.

MI can be also generalized via the DTC as follows in Eq. (8):

$$DTC(\mathbf{X}) = H(\mathbf{X}) - \sum_{i=1}^{N} H(X_i|\mathbf{X}^{-i}) \tag{8}$$

where $H(X_i|\mathbf{X}^{-i})$ is the residual entropy, that is the uncertainty associated with the state of the variable X_i that is not disclosed by the state of any other variable, or subsets of variables, comprised in $\mathbf{X}$. With this difference, DTC captures all the entropy that is shared at least between two elements of $\mathbf{X}$. Contrarily to TC, DTC is high in systems where *some* information is shared, but is low in both cases where $\mathbf{X}$ comprises totally random or synchronized variables.

The organizational information is the difference between TC and DTC (Eq. (9)), so that when it is positive ($TC(\mathbf{X}) > DTC(\mathbf{X})$) redundant information among the variables is predominant, whereas when it is negative ($TC(\mathbf{X}) < DTC(\mathbf{X})$) the system is characterized by information that is both shared but not redundant.

$$\Omega(\mathbf{X}) = TC(\mathbf{X}) - DTC(\mathbf{X}) \tag{9}$$

Ω is negative in systems dominated by synergistic interactions and positive in systems dominated by redundant interactions, and we used it as a proxy for synergy in our empirical analyses. For a thorough discussion on how the O-information can be interpreted see also refs. 112,113.

Information theory in Gaussian systems

All the information theoretic measures reported above require the estimation of the entropy H(X), and specifically of P(x), which can be challenging on empirical continuous data, like fMRI BOLD signals. However, for normally distributed (Gaussian) multivariate systems, closed-form estimators exist[114] that make the computation of the joint entropy easier. Under the assumption that BOLD data follows a multivariate Gaussian distribution - an assumption supported by numerous studies[115–117] - we can exploit those close-form estimators and derive information theoretical measures directly from the covariance matrix.

For a single Gaussian random variable $X \sim N(\mu, \sigma)$, the entropy can be computed as:

$$H^N(X) = \frac{\ln(2\pi e \sigma^2)}{2} \tag{10}$$

For a multivariate Gaussian random variable $\mathbf{X} = \{X_1, X_2, \ldots, X_N\}$, the joint entropy can be computed as:

$$H^N(\mathbf{X}) = \frac{\ln\left[(2\pi e \sigma)^N |\Sigma|\right]}{2} \tag{11}$$

where $|\Sigma|$ denotes the determinant of the covariance matrix of $\mathbf{X}$.

Similarly to what has been shown in the previous section, we can generalize these formulations of entropy and derive the MI of a bivariate system comprising the variables X_1 and X_2. Then, given the Pearson correlation coefficient between X_1 and X_2, here referred to as ρ, the MI can be

written as:

$$I^N(X_1; X_2) = \frac{-\ln(1 - \rho^2)}{2} \tag{12}$$

Finally, the estimator for the total correlation of a multivariate Gaussian system has the following formulation:

$$TC^N(\mathbf{X}) = \frac{-\ln(|\Sigma|)}{2} \tag{13}$$

Then, from these equations, it is straightforward to rewrite in close-form both the dual total correlation and the organizational information. For a more thorough description of how information theoretical measures can be drawn in close-form see also refs. [20],[114].

Multiscale modularity maximization on brain networks

Modular structure was inferred on the HCP dataset via modularity maximization[32], which, given an adjacency matrix, returns partitions of the networks into modules that are maximally internally dense with respect to a chance level (or null model). Modularity is commonly indicated with Q and is formulated as follows:

$$Q(\gamma) = \sum_{ij}[W_{ij} - \gamma P_{ij}]\delta(\sigma_i, \sigma_j). \tag{14}$$

where W_{ij} and P_{ij} are the weights of the connections between node i and node j in the adjacency matrix W and in the null model P, γ is a resolution parameter, and $\delta(\sigma_i, \sigma_j)$ is a factor whose value is 1 when i and node j belong to the same community and 0 otherwise.

The choice of the null model P is non-trivial. The weight that we expect to find between two nodes strongly depends on the characteristics of the adjacency matrix W[118]. For covariance matrices, i.e., full weighted matrices, a reasonable choice is the Potts null model, where P_{ij} holds the same value for every pair of nodes and is modulated by γ. The choice of γ then defines the spatial resolution of the partition in terms of size and number of modules, so that low γ-values yield to a coarse modular structure, whilst high γ-values produce a finer modular structure.

In this paper, we applied a 2-step multiresolution approach, called multiresolution consensus clustering (MRCC)[38]. In the first step, we linearly sampled 1000 values of γ in the range [0, 1] and we ran modularity maximization with each one of these γ-values obtaining 1000 differently resolved partitions. At this point, two γ-values have been identified, γ_L and γ_H, which generated partitions with a number of modules between 2 and $N/4$ (with $N = 200$ being the number of nodes). Then, in the second step, we linearly sampled 1000 values of γ, but this time in the range [γ_L, γ_H], running modularity maximization with each of these values. Again, retaining partitions with at least 2 and maximum $N/4$ modules, we obtained 990 partitions (10 γ-values/partitions did not survive the limits imposed on clusters number) at different spatial resolutions that we used for our empirical analyses.

Multiscale modularity maximization was also applied to a modified adjacency matrix transformed based on Eq. (12), where the entries represent Mutual Information between pairs of time series instead of Pearson correlation. This complementary analysis was performed to assess the effect of this transformation on conventional community detection algorithms, and the results are reported in Fig. S9.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

All the data used in this study are available from the Human Connectome Project (http://www.humanconnectomeproject.org/), the Microstructure-Informed Connectomics Project (https://osf.io/j532r/) and the Nathan Kline Institute (fcon_1000.projects.nitrc.org/indi/enhanced/). Preprocessed network data used in the manuscript are available at https://github.com/mariagraziaP/comm_detection_via_redundancy_optimization/tree/main/src/mat. Numerical source data to reproduce the figures of the manuscript are available in the Supplementary Data.

Code availability

MATLAB code to reproduce the analysis is available at https://github.com/mariagraziaP/comm_detection_via_redundancy_optimization/tree/main.

References

1. Tononi, G., Sporns, O. & Edelman, G. M. A measure for brain complexity: relating functional segregation and integration in the nervous system. *Proc. Natl. Acad. Sci. USA* **91**, 5033–5037 (1994).
2. Shine, J. M. et al. The dynamics of functional brain networks: integrated network states during cognitive task performance. *Neuron* **92**, 544–554 (2016). Publisher: Elsevier.
3. Sporns, O. Network attributes for segregation and integration in the human brain. *Curr. Opin. Neurobiol.* **23**, 162–171 (2013).
4. Bullmore, E. & Sporns, O. The economy of brain network organization. *Nat. Rev. Neurosci.* **13**, 336–349 (2012).
5. Luppi, A. I. et al. Consciousness-specific dynamic interactions of brain integration and functional diversity. *Nat. Commun.* **10**, 4616 (2019).
6. Deco, G., Tononi, G., Boly, M. & Kringelbach, M. L. Rethinking segregation and integration: contributions of whole-brain modelling. *Nat. Rev. Neurosci.* **16**, 430–439 (2015).
7. Wang, R. et al. Segregation, integration, and balance of large-scale resting brain networks configure different cognitive abilities. *Proc. Natl. Acad. Sci. USA* **118**, e2022288118 (2021).
8. Fukushima, M. & Sporns, O. Structural determinants of dynamic fluctuations between segregation and integration on the human connectome. *Commun. Biol.* **3**, 1–11 (2020).
9. Bassett, D. S. & Sporns, O. Network neuroscience. *Nat. Neurosci.* **20**, 353–364 (2017).
10. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**, 186–198 (2009).
11. Sporns, O. & Betzel, R. F. Modular brain networks. *Annu. Rev. Psychol.* **67**, 613–640 (2016).
12. Wig, G. S. Segregated systems of human brain networks. *Trends Cogn. Sci.* **21**, 981–996 (2017).
13. Medaglia, J. D., Lynall, M. E. & Bassett, D. S. Cognitive network neuroscience. *J. Cogn. Neurosci.* **27**, 1471–1491 (2015).
14. van den Heuvel, M. P. & Sporns, O. Network hubs in the human brain. *Trends Cogn. Sci.* **17**, 683–696 (2013).
15. van den Heuvel, M. P. & Sporns, O. Rich-club organization of the human connectome. *J. Neurosci.* **31**, 15775–15786 (2011).
16. Battiston, F. et al. Networks beyond pairwise interactions: structure and dynamics. *Phys. Rep.* **874**, 1–92 (2020).
17. Rosas, F. E. et al. Disentangling high-order mechanisms and high-order behaviours in complex systems. *Nat. Phys.* **18**, 476–477 (2022).
18. Battiston, F. et al. The physics of higher-order interactions in complex systems. *Nat. Phys.* **17**, 1093–1098 (2021).
19. Varley, T. F., Pope, M., Puxeddu, MariaGrazia, Faskowitz, J. & Sporns, O. Partial entropy decomposition reveals higher-order information structures in human brain activity. *Proc. Natl. Acad. Sci. USA* **120**, e2300888120 (2023).
20. Varley, T. F., Pope, M., Faskowitz, J. & Sporns, O. Multivariate information theory uncovers synergistic subsystems of the human cerebral cortex. *Commun. Biol.* **6**, 1–12 (2023).

21. Pope, M., Varley, T. F., Puxeddu, M. G., Faskowitz, J. & Sporns, O. Time-varying synergy/redundancy dominance in the human cerebral cortex. *J. Phys. Complex.* **6**, 015015 (2025).

22. Luppi, A. I., Rosas, F. E., Mediano, PedroA. M., Menon, D. K. & Stamatakis, E. A. Information decomposition and the informational architecture of the brain. *Trends Cogn. Sci.* **28**, 352–368 (2024).

23. Luppi, A. I. et al. A synergistic core for human brain evolution and cognition. *Nat. Neurosci.* **25**, 771–782 (2022).

24. Rosas, F. E., Mediano, PedroA. M., Gastpar, M. & Jensen, H. J. Quantifying high-order interdependencies via multivariate extensions of the mutual information. *Phys. Rev. E* **100**, 032305 (2019).

25. Timme, N., Alford, W., Flecker, B. & Beggs, J. M. Synergy, redundancy, and multivariate information measures: an experimentalist's perspective. *J. Comput. Neurosci.* **36**, 119–140 (2014).

26. Clauw, K., Stramaglia, S. & Marinazzo, D. Higher-order mutual information reveals synergistic sub-networks for multi-neuron importance. Preprint at http://arxiv.org/abs/2211.00416 (2022).

27. Van Essen, D. C. et al. The WU-Minn human connectome project: an overview. *NeuroImage* **80**, 62–79 (2013).

28. Royer, J. et al. An open MRI dataset for multiscale neuroscience. *Sci. Data* **9**, 569 (2022).

29. Nooner, K. B. et al. The NKI-Rockland sample: a model for accelerating the pace of discovery science in psychiatry. *Front. Neurosci.* **6** (2012). ISSN 1662-453X. https://doi.org/10.3389/fnins.2012.00152. https://www.frontiersin.org/journals/neuroscience/articles/10.3389/fnins.2012.00152/full.

30. Schaefer, A. et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* **28**, 3095–3114 (2018).

31. Chen, C. C., Tyler, C. W. & Baseler, H. A. Statistical properties of BOLD magnetic resonance activity in the human brain. *NeuroImage* **20**, 1096–1109 (2003).

32. Newman, M. E. J. & Girvan, M. Finding and evaluating community structure in networks. *Phys. Rev. E* **69**, 026113 (2004).

33. Blondel, V. D., Guillaume, J. L., Lambiotte, R. & Lefebvre, E. Fast unfolding of communities in large networks. *J. Stat. Mech.: Theory Exp.* **2008**, P10008 (2008).

34. Watanabe, S. Information theoretical analysis of multivariate correlation. *IBM J. Res. Dev.* **4**, 66–82 (1960).

35. McGill, W. Multivariate information transmission. *Trans. IRE Professional Group Inf. Theory* **4**, 93–111 (1954).

36. Schneidman, E., Still, S., Berry, M. J. & Bialek, W. Network information and connected correlations. *Phys. Rev. Lett.* **91**, 238701 (2003).

37. Meilä, M. Comparing clusterings–an information based distance. *J. Multivar. Anal.* **98**, 873–895 (2007).

38. Jeub, LucasG. S., Sporns, O. & Fortunato, S. Multiresolution consensus clustering in networks. *Sci. Rep.* **8**, 3259 (2018).

39. Peixoto, T. P. Hierarchical block structures and high-resolution model selection in large networks. *Phys. Rev. X* **4**, 011047 (2014).

40. Power, J. D. et al. Functional network organization of the human brain. *Neuron* **72**, 665–678 (2011).

41. Thomas Yeo, B. T. et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).

42. Doucet, G. et al. Brain activity at rest: a multiscale hierarchical functional organization. *J. Neurophysiol.* **105**, 2753–2763 (2011).

43. Vinh, N. X., Epps, J. & Bailey, J. Information theoretic measures for clusterings comparison: Is a correction for chance necessary? In *Proceedings of the 26th Annual International Conference on Machine Learning*, ICML '09, pages 1073–1080, (2009). Association for Computing Machinery. ISBN 978-1-60558-516-1. https://doi.org/10.1145/1553374.1553511.

44. Faskowitz, J., Esfahlani, F. Z., Jo, Y., Sporns, O. & Betzel, R. F. Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture. *Nat. Neurosci.* **23**, 1644–1654 (2020).

45. Tanner, J. C. et al. Synchronous high-amplitude co-fluctuations of functional brain networks during movie-watching. *Imaging Neurosci.* **1**, 1–21 (2023).

46. Betzel, R. F., Chumin, E., Esfahlani, F. Z., Tanner, J., and Faskowitz, J. System-level high-amplitude co-fluctuations. Preprint at https://www.biorxiv.org/content/10.1101/2022.07.26.501262v1 (2022).

47. Guimer, R. & Nunes Amaral, L. A. Functional cartography of complex metabolic networks. *Nature* **433**, 895–900 (2005). Publisher: Nature Publishing Group.

48. Sun, L. et al. Human lifespan changes in the brain's functional connectome. *Nat. Neurosci.* **28**, 891–901 (2025).

49. Betzel, R. F. et al. Changes in structural and functional connectivity among resting-state networks across the human lifespan. *NeuroImage* **102**, 345–357 (2014).

50. Zuo, X. N. et al. Human connectomics across the life span. *Trends Cogn. Sci.* **21**, 32–45 (2017).

51. Puxeddu, M. G. et al. The modular organization of brain cortical connectivity across the human lifespan. *NeuroImage* **218**, 116974 (2020).

52. Bassett, D. S. et al. Dynamic reconfiguration of human brain networks during learning. *Proc. Natl. Acad. Sci. USA* **108**, 7641–7646 (2011).

53. Corbetta, M. & Shulman, G. L. Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* **3**, 201–215 (2002).

54. Dosenbach, NicoU. F. et al. A core system for the implementation of task sets. *Neuron* **50**, 799–812 (2006).

55. Raichle, M. E. The restless brain: how intrinsic activity organizes brain function. *Philos. Trans. R. Soc. B Biol. Sci.* **370**, 20140172 (2015).

56. Crossley, N. A. et al. The hubs of the human connectome are generally implicated in the anatomy of brain disorders. *Brain* **137**, 2382–2395 (2014).

57. Valenti, S. et al. Assessing High-Order Interdependencies Through Static O-Information Measures Computed on Resting State fMRI Intrinsic Component Networks. In Pier Luigi Mazzeo, Emanuele Frontoni, Stan Sclaroff, and Cosimo Distante, editors, *Image Analysis and Processing. ICIAP 2022 Workshops*, pages 386–397, (2022). Springer International Publishing. ISBN 978-3-031-13321-3. https://doi.org/10.1007/978-3-031-13321-3_34.

58. Cohen, J. R. & D'Esposito, M. The segregation and integration of distinct brain networks and their relationship to cognition. *J. Neurosci.* **36**, 12083–12094 (2016).

59. Salathé, M. & Jones, J. H. Dynamics and control of diseases in networks with community structure. *PLOS Comput. Biol.* **6**, e1000736 (2010).

60. Stam, C. J. et al. Graph theoretical analysis of magnetoencephalographic functional connectivity in Alzheimer's disease. *Brain* **132**, 213–224 (2009).

61. Bosboom, J. L. W., Stoffers, D., Wolters, E. C., Stam, C. J. & Berendse, H. W. MEG resting state functional connectivity in Parkinson's disease related dementia. *J. Neural Transm.* **116**, 193–202 (2009).

62. Lynall, M. E. et al. Functional connectivity and brain networks in schizophrenia. *J. Neurosci.* **30**, 9477–9487 (2010).

63. Warren, D. E. et al. Network measures predict neuropsychological outcome after brain injury. *Proc. Natl. Acad. Sci. USA* **111**, 14247–14252 (2014).

64. Guimerà, R. & Amaral, L. A. Nunes Cartography of complex networks: modules and universal roles. *J. Stat. Mech. Theory Exp.* **2005**, P02001 (2005).

65. Oldham, S. & Fornito, A. The development of brain network hubs. *Dev. Cogn. Neurosci.* **36**, 100607 (2019).

66. Power, J. D., Schlaggar, B. L., Lessov-Schlaggar, C. N. & Petersen, S. E. Evidence for hubs in human functional brain networks. *Neuron* **79**, 798–813 (2013).

67. Cole, M. W. et al. Multi-task connectivity reveals flexible hubs for adaptive task control. *Nat. Neurosci.* **16**, 1348–1355 (2013).

68. Bertolero, M. A., Yeo, B. T. Thomas & D'Esposito, M. The modular and integrative functional architecture of the human brain. *Proc. Natl. Acad. Sci. USA* **112**, E6798–E6807 (2015).

69. Timme, N. M. et al. High-degree neurons feed cortical computations. *PLOS Comput. Biol.* **12**, e1004858 (2016).

70. Faber, S. P., Timme, N. M., Beggs, J. M. & Newman, E. L. Computation is concentrated in rich clubs of local cortical networks. *Netw. Neurosci.* **3**, 384–404 (2019).

71. Cajic, P. et al. On the information-theoretic formulation of network participation. *J. Phys. Complex.* **5**, 015021 (2024).

72. Santoro, A., Battiston, F., Petri, G. & Amico, E. Higher-order organization of multivariate time series. *Nat. Phys.* **19**, 221–229 (2023).

73. Baum, G. L. et al. Development of structure-function coupling in human brain networks during youth. *Proc. Natl. Acad. Sci. USA* **117**, 771–778 (2020).

74. Chan, M. Y., Alhazmi, F. H., Park, D. C., Savalia, N. K. & Wig, G. S. Resting-state network topology differentiates task signals across the adult life span. *J. Neurosci.* **37**, 2734–2745 (2017).

75. Geerligs, L., Renken, R. J., Saliasi, E., Maurits, N. M. & Lorist, M. M. A brain-wide study of age-related changes in functional connectivity. *Cereb. Cortex* **25**, 1987–1999 (2015).

76. Song, J. et al. Age-related reorganizational changes in modularity and functional connectivity of human brain networks. *Brain Connect.* **4**, 662–676 (2014).

77. Spreng, R. N., Stevens, W. D., Viviano, J. D. & Schacter, D. L. Attenuated anticorrelation between the default and dorsal attention networks with aging: evidence from task and rest. *Neurobiol. Aging* **45**, 149–160 (2016).

78. Grady, C., Sarraf, S., Saverino, C. & Campbell, K. Age differences in the functional interactions among the default, frontoparietal control, and dorsal attention networks. *Neurobiol. Aging* **41**, 159–172 (2016).

79. Gatica, M. et al. High-order interdependencies in the aging brain. *Brain Connecti.* **11**, 734–744 (2021).

80. Gatica, M. et al. High-order functional redundancy in ageing explained via alterations in the connectome in a whole-brain model. *PLOS Comput. Biol.* **18**, e1010431 (2022).

81. Camino-Pontes, B. et al. Interaction information along lifespan of the resting brain dynamics reveals a major redundant role of the default mode network. *Entropy* **20**, 742 (2018).

82. Barrett, A. B. Exploration of synergistic and redundant information sharing in static and dynamical Gaussian systems. *Phys. Rev. E* **91**, 052802 (2015).

83. Cover, T. M. & Thomas, J. A. *Elements of Information Theory, 2nd Edition*. ISBN 978-0-471-24195-9. https://www.wiley.com/en-it/Elements+of+Information+Theory%2C+2nd+Edition-p-9780471241959 (2006).

84. Colenbier, N. et al. Disambiguating the role of blood flow and global signal with partial information decomposition. *NeuroImage* **213**, 116699 (2020).

85. Sanchez-Gorostiaga, A., Bajić, D., Osborne, M. L., Poyatos, J. F. & Sanchez, A. High-order interactions distort the functional landscape of microbial consortia. *PLoS Biol.* **17**, e3000550 (2019).

86. Grilli, J., Barabás, G., Michalska-Smith, M. J. & Allesina, S. Higher-order interactions stabilize dynamics in competitive network models. *Nature* **548**, 210–213 (2017).

87. Scarpino, S. V. & Petri, G. On the predictability of infectious disease outbreaks. *Nat. Commun.* **10**, 898 (2019).

88. Chang, Z., Wang, X., Wu, Y., Lin, P. & Wang, R. Segregation, integration and balance in resting-state brain functional networks associated with bipolar disorder symptoms. *Hum. Brain Mapp.* **44**, 599–611 (2023).

89. Herzog, R. et al. Genuine high-order interactions in brain networks and neurodegeneration. *Neurobiol. Dis.* **175**, 105918 (2022).

90. Koçillari, L. et al. Behavioural relevance of redundant and synergistic stimulus information between functionally connected neurons in mouse auditory cortex. *Brain Inform.* **10**, 34 (2023).

91. Varley, T. F., Sporns, O., Schaffelhofer, S., Scherberger, H. & Dann, B. Information-processing dynamics in neural networks of macaque cerebral cortex reflect cognitive state and behavior. *Proc. Natl. Acad. Sci. USA* **120**, e2207677120 (2023).

92. Demertzi, A. et al. Human consciousness is supported by dynamic complex patterns of brain signal coordination. *Sci. Adv.* **5**, eaat7603 (2019).

93. Fukushima, M. et al. Structure-function relationships during segregated and integrated network states of human brain functional connectivity. *Brain Struct. Funct.* **223**, 1091–1106 (2018).

94. Mucha, P. J., Richardson, T., Macon, K., Porter, M. A. & Onnela, Jukka-Pekka Community structure in time-dependent, multiscale, and multiplex networks. *Science* **328**, 876–878 (2010).

95. Betzel, R. F. et al. The community structure of functional brain networks exhibits scale-specific patterns of inter- and intra-subject variability. *NeuroImage* **202**, 115990 (2019).

96. Puxeddu, M. G., Faskowitz, J., Sporns, O., Astolfi, L. & Betzel, R. F. Multi-modal and multi-subject modular organization of human brain networks. *NeuroImage* **264**, 119673 (2022).

97. Zamani Esfahlani, F., Faskowitz, J., Slack, J., Mišić, B. & Betzel, R. F. Local structure-function relationships in human brain networks across the lifespan. *Nat. Commun.* **13**, 2053 (2022).

98. Jo, Y., Tanner, J., Seguin, C., Faskowitz, J. & Betzel, R. F. Variation in high-amplitude events across the human lifespan. Preprint at https://www.biorxiv.org/content/10.1101/2024.05.15.594378v1 (2024).

99. Glasser, M. F. et al. The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80**, 105–124 (2013).

100. Parkes, L., Fulcher, B., Yücel, M. & Fornito, A. An evaluation of the efficacy, reliability, and sensitivity of motion correction strategies for resting-state functional MRI. *NeuroImage* **171**, 415–436 (2018).

101. Cruces, R. R. et al. Micapipe: a pipeline for multimodal neuroimaging and connectome analysis. *NeuroImage* **263**, 119612 (2022).

102. Esteban, O. et al. fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat. Methods* **16**, 111–116 (2019).

103. Gorgolewski, K. et al. Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Front. Neuroinformatics* (Frontiers, 2011). ISSN 1662-5196. https://doi.org/10.3389/fninf.2011.00013. https://www.frontiersin.org/articles/10.3389/fninf.2011.00013.

104. Abraham, A. et al. Machine learning for neuroimaging with scikit-learn. *Front. Neuroinformatics* (Frontiers, 2014). ISSN 1662-5196. https://doi.org/10.3389/fninf.2014.00014. https://www.frontiersin.org/articles/10.3389/fninf.2014.00014.

105. Wang, S. et al. Evaluation of Field Map and Nonlinear Registration Methods for Correction of Susceptibility Artifacts in Diffusion MRI. *Front. Neuroinformatics* (Frontiers, 2017). ISSN 1662-5196. https://doi.org/10.3389/fninf.2017.00017. https://www.frontiersin.org/articles/10.3389/fninf.2017.00017.

106. Treiber, J. M. et al. Characterization and correction of geometric distortions in 814 diffusion weighted images. *PLoS ONE* **11**, e0152472 (2016).

107. Greve, D. N. & Fischl, B. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage* **48**, 63–72 (2009).

108. Power, J. D. et al. Methods to detect, characterize, and remove motion artifact in resting state fMRI. *NeuroImage* **84**, 320–341 (2014).

109. Shannon, C. E. A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 379–423 (1948).

110. Finn, C. & Lizier, J. T. Generalised measures of multivariate information content. *Entropy* **22**, 216 (2020).

111. Abdallah, S. A. & Plumbley, M. D. A measure of statistical complexity based on predictive information with application to finite spin systems. *Phys. Lett. A* **376**, 275–281 (2012).

112. Varley, T. F. Generalized decomposition of multivariate information. *PLOS ONE* **19**, e0297128 (2024).

113. Williams, P. L. & Beer, R. D. Generalized measures of information transfer. Preprint at http://arxiv.org/abs/1102.1507 (2011).

114. Lizier, J. T. JIDT: An Information-Theoretic Toolkit for Studying the Dynamics of Complex Systems. *Front. Robot. AI* (Frontiers, 2014). ISSN 2296-9144. https://doi.org/10.3389/frobt.2014.00011. https://www.frontiersin.org/articles/10.3389/frobt.2014.00011.

115. Schulz, M. A. et al. Different scaling of linear models and deep learning in UKBiobank brain images versus machine-learning datasets. *Nat. Commun.* **11**, 4238 (2020).

116. Hlinka, J., Paluš, M., Vejmelka, M., Mantini, D. & Corbetta, M. Functional connectivity in resting-state fMRI: Is linear correlation sufficient? *NeuroImage* **54**, 2218–2225 (2011).

117. Liégeois, R., Yeo, B. T. Thomas & Van De Ville, D. Interpreting null models of resting-state functional MRI dynamics: not throwing the model out with the hypothesis. *NeuroImage* **243**, 118518 (2021).

118. Zamani Esfahlani, F. et al. Modularity maximization as a flexible and generic framework for brain network exploratory analysis. *NeuroImage* **244**, 118607 (2021).

Author contributions

M.G.P. and O.S. conceived and designed the study. M.G.P. performed the data analysis. M.G.P., M.P., T.F.V., and O.S. contributed to the development of the methodology. J.F. preprocessed and curated the fMRI data. M.G.P. wrote the first draft of the manuscript. All authors contributed to manuscript editing and revision.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-025-08198-2.

Correspondence and requests for materials should be addressed to Maria Grazia Puxeddu.

Peer review information *Communications Biology* thanks Michael Reimann, Andrea Luppi and the other anonymous reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Benjamin Bessieres. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.