

Assessing imbalance in signed brain networks

Marzio Di Vece^{1,2,*} Emanuele Agrimi^{2,3,4} Samuele Tatullo² Tommaso Gili²
 Miguel Ibáñez-Berganza^{2,5} and Tiziano Squartini^{1,2,5}

¹*Scuola Normale Superiore, P.zza dei Cavalieri 7, 56126 Pisa, Italy*

²*IMT School for Advanced Studies, P.zza San Francesco 19, 55100 Lucca, Italy*

³*Institute for Advanced Study, University of Amsterdam, Oude Turfmarkt 145–147, 1012 GC Amsterdam, The Netherlands*

⁴*Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Science Park 105–107, 1098 XG Amsterdam, The Netherlands*

⁵*Istituto Nazionale di Alta Matematica ‘Francesco Severi’ (INdAM-GNAMP), P.le Aldo Moro 5, 00185 Rome, Italy*



(Received 1 October 2025; accepted 15 June 2026; published 11 September 2026)

Many complex systems—be they financial, natural, or social—are composed of units—such as stocks, neurons, or agents—whose joint activity can be represented as a multivariate time series. An issue of both practical and theoretical importance concerns the possibility of inferring the presence of a *static* relationship between any two units solely from their *dynamic* behavior. The present contribution aims at tackling such an issue within the framework of traditional hypothesis testing: Briefly speaking, our suggestion is that of linking any two units if behaving in a sufficiently similar way. To achieve such a goal, we project a multivariate time series onto a signed graph by (1) comparing the empirical properties of the former with those expected under a suitable benchmark and (2) linking any two units with a positive (negative) edge in case the corresponding series shares a significantly large number of concordant (discordant) values. To define our benchmarks, we adopt an information-theoretic approach that is rooted into the constrained maximization of Shannon entropy, a procedure inducing an ensemble of multivariate time series that preserves some of the empirical properties on average, while randomizing everything else. We showcase the possible applications of our method by addressing one of the most timely issues in the domain of neurosciences, i.e., that of determining whether brain networks are frustrated or not, and, if so, to what extent. As our results suggest, this is indeed the case, with the major contribution to the underlying negative subgraph coming from the subcortical regions (and, to a lesser extent, from the limbic ones). At the mesoscopic level, the minimization of the Bayesian information criterion, instantiated with the signed stochastic block model, reveals that brain regions gather into modules aligning with the statistical variant of the relaxed balance theory.

DOI: [10.1103/p6qp-2s9f](https://doi.org/10.1103/p6qp-2s9f)

I. INTRODUCTION

Time series are ubiquitous and versatile representations of dynamical complex systems, be they financial, natural, or social. A major challenge is that of inferring the underlying dependencies between units from the (often nonlinear) dynamics characterizing them. Being capable of transforming a time series into a symbolic sequence, while preserving essential information about the “original” generative process, is of practical relevance, as the *dynamical* properties of the former can be revealed by focusing on the *structural* properties of the latter [1–3]. To this aim, sequential data are often represented as traditional graphs, with *nodes* replacing individual units (such as stocks, neurons, or agents) and *edges* reflecting their activity-driven interactions.

This choice plays a major role in the neuroscientific domain, where network-based approaches have revolutionized

our understanding of brain functioning, especially through the study of functional connectivity (FC) networks derived from blood oxygen level-dependent (BOLD) functional Magnetic Resonance Imaging (fMRI) signals. Rooted in graph theory and statistical physics, this perspective, often referred to as *network neuroscience*, aims at revealing the organizational principles of cognition and behavior [4] by modeling the brain as an integrated system of interacting regions [5].

In the following subsection, a short overview of the approaches that have been proposed so far to extract relationships between nodes directly from the time series of the original units is provided: In Sec. IA, methods defining an intermediate proximity matrix are discussed; in Secs. IB and IC, methods that map univariate and multivariate time series onto graphs without a proper validation procedure are discussed; and in Sec. ID, the need to approach such an issue within the framework of traditional hypothesis testing is motivated.

A. From time series to proximity and from proximity to graphs

Given a set of time series, the most common approach is that of extracting a *proximity matrix* that is subsequently used for defining a graph. Although several definitions of proximity

*Contact author: marzio.divece@sns.it

have been proposed [6], the commonest choice is identifying it with the notion of *pairwise correlation*, as returned by the calculation of the Pearson correlation coefficient

$$r_{ij} = \frac{\sum_{t=1}^T (X_{it} - \underline{X}_i) \times (X_{jt} - \underline{X}_j)}{\sqrt{\sum_{t=1}^T (X_{it} - \underline{X}_i)^2} \times \sqrt{\sum_{t=1}^T (X_{jt} - \underline{X}_j)^2}}, \quad (1)$$

where

$$\underline{X}_i = \frac{\sum_{t=1}^T X_{it}}{T} \quad \text{and} \quad \underline{X}_j = \frac{\sum_{t=1}^T X_{jt}}{T} \quad (2)$$

represent the (sample) average of series $X_i = \{X_{it}\}_{t=1, \dots, T}$ and $X_j = \{X_{jt}\}_{t=1, \dots, T}$.

A less popular, but still worth mentioning, definition of proximity is based on the notion of *partial correlation* [7,8]. If we keep indicating the matrix of pairwise correlations with $\mathbf{R}$, the generic entry of the matrix of partial correlations, then, reads $\hat{J}_{ik} = J_{ik} / \sqrt{J_{ii} \times J_{kk}}$, with $\mathbf{J} = \mathbf{R}^{-1}$ being the so-called *precision matrix*. Estimations of the FC based on partial correlations are affected by a smaller intersubject variability [9,10], as they have been shown to resemble more closely the so-called structural connectivity patterns [11–14].

Fixed thresholding algorithms. The simplest approach of this kind is based on interpreting a correlation matrix (be it partial or not) as a fully connected, weighted matrix. Non-trivial structures can be, thus, obtained by implementing sparsification procedures, the simplest of which is thresholding correlations [15]. This procedure ultimately boils down to remove the weaker connections—usually the most affected by experimental noise [16] while retaining the stronger ones; still, as reported by the authors of Ref. [17], thresholding may prevent the detection of a modular organization, as the correlations within modules may be cut out even if significantly stronger than those between modules.

In the financial literature, the network resulting from the application of a threshold is also known as the *asset graph* [1,2,18]: In Ref. [1], the authors derived an asset graph from the correlation matrix of a set of stocks by filling an empty graph, edge by edge, starting from the one corresponding to the largest correlation and stopping at the N th largest correlation (with N being the number of nodes); upon doing so, the considered asset graph is comparable in size with the most popular *asset tree*, although being less constrained by requests such as the one of forbidding cycles.

Since letting the network density drive the thresholding procedure appears as arbitrary as setting a threshold directly on correlations, the proposal to determine the resulting network by requiring it to remain connected has been advanced: Such a data-driven procedure removes the weaker correlations until the value at which the resulting network disconnects is reached [19,20]. Although this kind of analysis has been popularized with the name of *percolation analysis*, the aforementioned value is, technically speaking, the *connectivity threshold*.

If the previous approach prescribes to stop pruning correlations once the connectivity threshold is reached, a second approach focuses on the mesoscopic organization of the network and prescribes to consider a correlation-induced distance, draw the corresponding dendrogram [21] and cut it in correspondence with a properly defined threshold [22,23]:

A possible criterion to individuate it is that of maximizing the modular character of the resulting network structure, analyzed upon running one of the many algorithms that have been proposed to this aim (Louvain, Infomap, Surprise, etc.) [24–26].

Filtering algorithms. A different class of algorithms attempts to filter a correlation matrix by letting its values “dress” peculiar graph topologies, supposedly isolating the most relevant correlations—a threshold is, thus, still individuated, although only indirectly.

The most popular filtering approach returns the so-called *maximum spanning tree* (MST), a technique retaining the $N - 1$ largest correlations while ensuring that (1) loops are discarded, (2) each node is reachable from any other node via a path that is the widest between its end points, and (3) the total weight is greater than, or equal to, the total weight of any other spanning tree. This method can be applied upon running the Kruskal’s algorithm and is often employed to distinguish the connector regions (i.e., the most central ones) from the provincial regions (i.e., the most peripheral ones) [22,27–30].

The Kruskal’s algorithm can be slightly modified by adding the extra condition that no edge between any two nodes that are already connected is allowed. Upon doing so, a *maximum spanning forest* (MSF) is obtained, i.e., a set of disconnected trees usually interpreted as the backbone of a set of communities. Both the MST and the MSF are rather different from their random counterparts [31], their structural organization providing a topological signature of the emergence of diseases such as schizophrenia [30].

An alternative approach, discarding less information, returns the so-called planar maximally filtered graph (PMFG) [32,33]. The criterion informing such a method prescribes to retain the correlations ensuring that the resulting network is a planar graph, i.e., a graph that can be drawn with no links crossing each other—more precisely, one in which edges may intersect only at common end points (nodes); since the PMFG always contains the MST, the former complements the latter by providing additional (and not just different) information. The PMFG has also been described as the simplest instance of a more general procedure embedding high-dimensional data into lower-dimensional manifolds with a controllable number of “holes” called “genus,” the PMFG corresponding to the case in which the genus is zero [33]. Although versatile, it suffers from some degree of arbitrariness too, affecting the properties of the postulated, approximating structure: The value of the genus must, in fact, be fixed *a priori*; besides, there is no obvious reason why time series should find a natural embedding into a bidimensional plane.

Testing correlations. A more principled approach to extract information from a correlation matrix is that of comparing (some of) its properties with the ones expected under a properly defined benchmark, the aim of this class of techniques being that of identifying nonrandom collective information about the set of time series induced by $\mathbf{C}$. Interestingly, these techniques do not discard negative correlations, which are known to be a consequence of spatial and temporal heterogeneities induced by nontrivial, neural-mediated hemodynamic mechanisms affecting the synchronization of neural activity, thus contributing to the stability of resting-state brain networks [34]. Although isolated attempts exist—such as the one proposed in Ref. [12], whose authors test the null

hypothesis that the average value of the distribution of partial correlations, for each pair of regions and across subjects, is zero—methods of this kind can be split into those based on *random matrix theory* [35–38] and those based on *entropy maximization* [39,40].

B. From univariate time series to graphs

Different (classes of) approaches seek to extract a network from one or more time series while ignoring (the intermediate step represented by) the proximity matrix [41].

Visibility-induced mapping. According to such a mapping, the time steps of a univariate time series are interpreted as nodes to be connected if “visible” to each other. Different types of visibility algorithms exist, designed to obtain undirected/directed, binary/weighted graphs.

The first algorithm of this kind was proposed in Ref. [42] and is based on the idea of representing each observation as a vertical bar whose height equals the numerical value of the observation itself: If the tops of any two bars can be *directly* connected via a “visibility line,” the corresponding nodes are, then, connected in the resulting network.

Several variants of the so-called *visibility graph* (VG) have been proposed: In Ref. [43], a *horizontal visibility graph* (HVG) is constructed by considering (only) horizontal visibility lines—as a consequence, the HVG is always a subgraph of the VG, hence containing less information, although its construction is faster; in Ref. [44], a *directed horizontal visibility graph* (DHVG) is constructed by letting any pair of time steps $t_i < t_j$ induce an edge $i \rightarrow j$; in Ref. [45], a *limited penetrable visibility graph* (LPVG) is constructed by connecting any two nodes if the visibility line joining the tops of the two associated bars crosses a number of “obstacles” that is smaller than, or equal to, a fixed value L —naturally, the LPVG reduces to the VG when $L = 0$, i.e., when there are no “obstacles”; in Ref. [46], a *weighted visibility graph* is constructed by enriching each edge of a DHVG with a weight matching the “view angle” of the observing node, its sign with respect to the horizontal axis indicating whether the values of the time series increase (if positive) or decrease (if negative); and in Ref. [47], a *differential visibility graph* is constructed by removing the edge set of the HVG from the edge set of the VG.

Transition-induced mapping. According to such a mapping, connections are established between (symbols assigned to) different portions of univariate time series, intended to represent states between which transitions occur: Methods of this kind, in fact, represent a time series as a directed, weighted graph.

The mapping introduced in Ref. [48] is based on a partition of the distribution of the values assumed by a time series in quantiles, to each of which a symbol is assigned [49]; the mapping introduced in Ref. [50] is named *coarse-grained phase space graphs* and requires the definition of a time window ω to determine the number of values of the series to be represented as an ω -dimensional vector, whose components are numerically determined by adopting a rule similar to the quantile-based one; a third variant of the same kind of mapping is the one introduced in Ref. [51] and named *visibility graphlet networks*: The time steps from t to $t + \omega$ are, then, mapped to a DHVG.

Proximity-induced mapping. According to such a mapping, connections are established between (groups of) time steps of a univariate time series, according to measures of distance/similarity between them.

The first method of this kind was introduced in Ref. [52] to represent pseudoperiodic time series. First, it segments a time series into cycles (e.g., dictated by a periodic behavior) without requiring them to have the same length; then, correlations between pairs of cycles are computed—by shifting the shortest one, if needed—and the largest correlation value is compared against a threshold: If found to be larger, an edge between the two cycles, represented as nodes in the resulting network, is established.

An alternative class of methods is the one gathering the so-called *recurrence networks* (RNs), which is based on the idea that the points of a time series are recurrent if their distance, in a properly defined embedding space, is “short enough”: First, each time step is mapped to a point of an embedding space and, then, distances are computed by using one of the many norms that can be defined. The way the embedding space is analyzed distinguishes RNs in three subcategories, i.e., *k-nearest-neighbor networks* [53], *adaptive nearest-neighbor networks* [54], and ϵ -*recurrence networks* [55].

C. From multivariate time series to graphs

Methods of this kind aim at extracting a network from a multivariate time series. While some of them output single-layer networks, each time series being represented as a node, others output multilayer networks, each time series being mapped to a different layer by employing any of the methods reviewed in the previous section.

Single-layer mapping. This first kind of mapping collects generalizations of the aforementioned RNs, hereby named *cross recurrence networks* (CRNs), *intersystem recurrence networks* (ISRNs), and *joint recurrence networks* (JRNs) [56–58]. CRNs are defined for pairs of time series that do not need to be of the same length [59,60]; such a recipe can be further extended to deal with more than two time series: In this case, we have the so-called ISRNs, inducing adjacency matrices constituted by (matrices representing the) RNs along the diagonal and by (matrices representing the) CRNs off the diagonal [56]. A third alternative is represented by the JRNs [57,58], inducing adjacency matrices that, similarly to what has been observed for the ISRNs, are blockwise, each diagonal block referring to a single time series and the off-diagonal blocks referring to pairs of time series.

Multiple-layer mapping. This second kind of mapping, on the other hand, associates a multivariate time series with a multilayer network. A first example of such a mapping is provided by Ref. [61]: Here, a multiplex visibility graph (MVG) is obtained by mapping each univariate time series into a graph by employing a VG; afterward, the MVG is aggregated into a single-layer weighted graph, each node representing a layer-specific VG and each weight proxying the similarity between the corresponding layers. The authors of Ref. [62] follow the same pipeline but replace VGs with RNs, thus employing *multiplex recurrence networks* to study paleoclimatic time series.

D. From multivariate time series to validated graphs

The most recent approaches emphasize the importance of probabilistic (null, as well as generative) models to benchmark network features and disentangle fundamental, structural properties from statistical artifacts [63,64]. These models play a crucial role in understanding whether empirical features (e.g., a modular structure, a small-world architecture, and a rich-club organization) are indicators of a genuine degree of self-organization or simple by-products of lower-order constraints [65].

In the present contribution, we focus on the issue of representing sequential data as relational data, approaching the following question: How can a set of time series be represented as a traditional graph? from the perspective of traditional hypothesis testing. After comparing the statistical properties of the former against those expected under a suitable benchmark, our suggestion is to project the set of time series onto a signed graph by linking any two series with a positive (negative) edge in case they share a statistically significant number of concordant (discordant) values. More formally, we adopt an information-theoretic approach that is rooted in the constrained maximization of Shannon entropy, a procedure inducing an ensemble of multivariate time series that preserves some of their empirical properties on average while randomizing everything else [66–68].

As our method deals with signed data [69–71], one may wonder how relevant the information provided by signs is within the neuroscientific framework. Several studies have recently investigated the role played by the negative correlations characterizing resting-state fMRI data: The authors of Ref. [34] show that negative correlations are essential for the stability of resting-state brain networks against perturbations or changes; the authors of Ref. [72] notice that the interplay between positive and negative correlations induces frustrated patterns preventing the brain from reaching a state of minimum energy and enhancing its adaptability to dynamical changes; finally, the authors of Ref. [73] show that the default mode network (DMN) works “in opposition” with the various task-based networks, a supposedly crucial feature to maintain cognitive functions and consciousness.

II. SETTING UP THE FORMALISM

Let us start by setting up the formalism. Since a generic time series can be represented as $X_i = \{X_{it}\}_{t=1, \dots, T}$, a set of N time series whose duration is T will be represented as an $N \times T$ rectangular matrix $\mathbf{X} = \{X_{it}\}_{i=1, \dots, N, t=1, \dots, T}$, whose generic entry X_{it} represents the value of the i th time series at time t . Since time series are typically preprocessed, let us explicitly distinguish the raw values above from the transformed ones (e.g., after standardization [74]) and indicate the (value of the) i th random variable at time t with the symbol w_{it} . Since w_{it} can be either *positive* or *negative*, to ease mathematical manipulations, let us employ Iverson’s brackets—a notation ensuring all quantities of interest are non-negative [69] and define

$$w_{it}^- = |w_{it}| \times [w_{it} < 0], \quad \forall i, t, \quad (3)$$

$$w_{it}^+ = |w_{it}| \times [w_{it} > 0], \quad \forall i, t, \quad (4)$$

i.e., mutually exclusive variables that induce the non-negative matrices $\mathbf{W}^+$ and $\mathbf{W}^-$ obeying

$$w_{it} = w_{it}^+ - w_{it}^-, \quad \forall i, t, \quad (5)$$

$$|w_{it}| = w_{it}^+ + w_{it}^-, \quad \forall i, t, \quad (6)$$

in an entrywise fashion—or, more compactly, $\mathbf{W} = \mathbf{W}^+ - \mathbf{W}^-$ and $|\mathbf{W}| = \mathbf{W}^+ + \mathbf{W}^-$. Their purely binary counterparts read

$$b_{it}^- = [w_{it} < 0], \quad \forall i, t, \quad (7)$$

$$b_{it}^+ = [w_{it} > 0], \quad \forall i, t, \quad (8)$$

in turn, inducing the non-negative matrices $\mathbf{B}^+$ and $\mathbf{B}^-$ that obey the analogous relationships

$$b_{it} = b_{it}^+ - b_{it}^-, \quad \forall i, t, \quad (9)$$

$$|b_{it}| = b_{it}^+ + b_{it}^-, \quad \forall i, t, \quad (10)$$

in an entrywise fashion—or, more compactly, $\mathbf{B} = \mathbf{B}^+ - \mathbf{B}^-$ and $|\mathbf{B}| = \mathbf{B}^+ + \mathbf{B}^-$. Adopting the terminology introduced for bipartite networks, we will also refer to the rectangular matrix $\mathbf{B}$ as the *biadjacency matrix* associated with the considered set of time series.

The present analysis solely focuses on binarized time series: Besides easing the mathematical tractability of the latter, this simplified representation has been shown to retain the salient features of resting-state brain dynamics—for instance, the *(a)synchrony* of time series—while being robust against intersubject variability [75,76].

A. Local properties and dependencies of time series

Let us start by defining a number of quantities of interest for our study. In analogy with the properties traditionally considered to analyze bipartite signed networks, the *total number of positive values* reads

$$B^+ = \sum_{i=1}^N \sum_{t=1}^T b_{it}^+ = \sum_{i=1}^N k_i^+ = \sum_{t=1}^T \kappa_t^+, \quad (11)$$

an expression further suggesting the definition of two (sets of) local positive quantities, i.e.,

$$k_i^+ = \sum_{t=1}^T b_{it}^+, \quad (12)$$

or *positive degree of series i* , capturing the total number of positive entries characterizing the i th series, and

$$\kappa_t^+ = \sum_{i=1}^N b_{it}^+, \quad (13)$$

or *positive degree at time t* , capturing the total number of positive entries characterizing the t th time step and providing information about the correlation between the members of the whole set of time series. Analogously, the *total number of negative values* reads

$$B^- = \sum_{i=1}^N \sum_{t=1}^T b_{it}^- = \sum_{i=1}^N k_i^- = \sum_{t=1}^T \kappa_t^-, \quad (14)$$

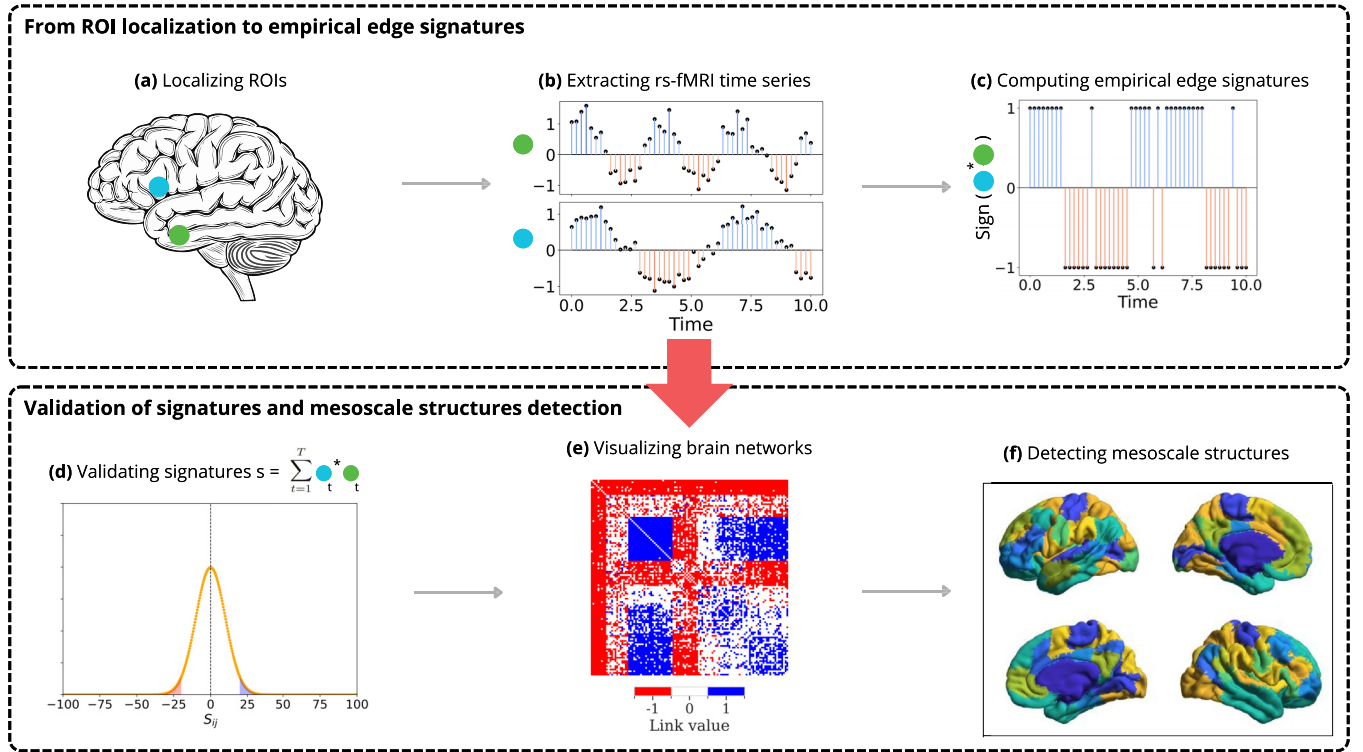


FIG. 1. Infographics illustrating the pipeline of our analysis. Pictorial representation of the pipeline we follow in the present contribution, from the registration of brain activity to the identification of communities in statistically validated signed projections: (a) rs-fMRI signals are recorded; (b) pairwise interactions between brain regions are considered; (c) the scalar product of each pair of (standardized and binarized) time series is calculated to quantify their *signature*; (d) its empirical value is validated against a null model to remove statistical noise; (e) a significantly large positive (negative) signature induces a positive (negative) link between the two involved brain regions; and (f) a BIC-based community detection reveals the modules partitioning our statistically validated signed projections. The first brain image was adapted from Canva (additional elements were added by the authors), while the last one was obtained through the Enigma Toolbox [77].

with obvious meaning of the symbols. The density of positive and negative values, reading $c^+ = B^+/(N \times T)$ and $c^- = B^-/(N \times T)$ and referred to as *positive* and *negative bipartite connectance*, remains naturally defined.

B. Concordant and discordant temporal motifs

Let us, now, define the second-order properties known as *dyadic motifs*. In the binary case,

$$B_{ij}^{++} = \sum_{t=1}^T b_{it}^+ b_{jt}^+ \quad (15)$$

counts the number of time steps in correspondence with which both series i and j are characterized by positive values and

$$B_{ij}^{--} = \sum_{t=1}^T b_{it}^- b_{jt}^- \quad (16)$$

counts the number of time steps in correspondence with which both series i and j are characterized by negative values. It is quite intuitive to ascribe them to the class of *concordant motifs*, i.e., patterns capturing the “agreement” between any two series. Analogously, it is quite intuitive to ascribe

$$B_{ij}^{+-} = \sum_{t=1}^T b_{it}^+ b_{jt}^- \quad (17)$$

and

$$B_{ij}^{-+} = \sum_{t=1}^T b_{it}^- b_{jt}^+, \quad (18)$$

counting the number of time steps in correspondence with which series i is characterized by a positive value and series j is characterized by a negative value or vice versa, to the class of *discordant motifs*, i.e., patterns capturing the “disagreement” between any two series.

Any two nodes whose associated time series exhibit a significantly large number of concordant motifs will be connected by a +1, while any two nodes whose associated time series exhibit a significantly large number of discordant motifs will be connected by a −1. We would also like to stress that our projection scheme is applicable only in the absence of null values—a condition that is naturally satisfied by BOLD time series, representing the focus of this study.

C. A scheme for the validated projection of multivariate time series

As depicted in Fig. 1, our validation algorithm works as follows:

(1) focus on a specific pair of series, say i and j , and measure their (binary) similarity (see Sec. III);

(2) quantify the statistical significance of the measured similarity, with respect to a properly defined benchmark, by computing the corresponding p -value, say p_{ij} (see Sec. IV);

(3) repeat the step above for each pair of series;

(4) apply a multiple hypothesis testing procedure and insert a (binary) link between the nodes i and j if and only if p_{ij} is found to be less than, or equal to, the threshold value p_{th} (see Sec. V).

Let us now describe each step in detail.

III. COMPUTING THE SIMILARITY OF ANY TWO SERIES

The first step of our method prescribes measuring the degree of similarity of series i and j . To this aim, let us consider the quantity named *signature* [69], defined as the scalar product of the two series: Upon defining $\mathbf{b}_i$ as the i th row of the matrix $\mathbf{B}$, the signature reads

$$S_{ij} = \mathbf{b}_i \times \mathbf{b}_j = \sum_{t=1}^T b_{it} b_{jt} = \sum_{t=1}^T (b_{it}^+ - b_{it}^-)(b_{jt}^+ - b_{jt}^-) \\ = \sum_{t=1}^T [(b_{it}^+ b_{jt}^+ + b_{it}^- b_{jt}^-) - (b_{it}^+ b_{jt}^- + b_{it}^- b_{jt}^+)]; \quad (19)$$

in words, the signature is the difference between two quantities, i.e., the *concordance* of series i and j , reading

$$C_{ij} = \sum_{t=1}^T C_{ijt} = \sum_{t=1}^T (b_{it}^+ b_{jt}^+ + b_{it}^- b_{jt}^-) = B_{ij}^{++} + B_{ij}^{--}, \quad (20)$$

and counting the number of concordant motifs, and the *discordance* of nodes i and j , reading

$$D_{ij} = \sum_{t=1}^T D_{ijt} = \sum_{t=1}^T (b_{it}^+ b_{jt}^- + b_{it}^- b_{jt}^+) = B_{ij}^{+-} + B_{ij}^{-+}, \quad (21)$$

and counting the number of discordant motifs: In symbols, $S_{ij} = C_{ij} - D_{ij}$, $\forall i < j$. A naive way of projecting a multivariate binary time series would prescribe to stop here and pose

$$a_{ij}^{\text{naive}} = \text{sgn}[S_{ij}], \quad (22)$$

i.e., apply the sign function to the signature, hence connecting series i and j with a positive link if $S_{ij} > 0$, i.e., $C_{ij} > D_{ij}$, and with a negative link if $S_{ij} < 0$, i.e., $C_{ij} < D_{ij}$.

IV. TESTING THE SIMILARITY OF ANY TWO SERIES

The second step of our method prescribes evaluating the statistical significance of the similarity of any two series. To this aim, let us first notice that

$$-T \leq S_{ij} \leq T, \quad (23)$$

where $S_{ij} = -T$ if $C_{ij} = 0$ (i.e., if each dyadic motif is composed by a -1 and a $+1$) and $S_{ij} = T$ if $D_{ij} = 0$ (i.e., if each dyadic motif is composed by either two -1 s or two $+1$ s) and, then, provide the derivation of two kinds of benchmarks, i.e., a homogeneous and a heterogeneous one—In what follows, we will employ *exponential random graphs* (ERGs) [66,78,79] defined by either global or local

linear constraints, hence treating links as independent random variables.

A. Homogeneous benchmark

Let us start by considering the homogeneous benchmark defined by the finite scheme [80]

$$b_{it} \sim \begin{pmatrix} -1 & +1 \\ p^- & p^+ \end{pmatrix}, \quad \forall i, t, \quad (24)$$

which, in turn, induces the finite scheme

$$b_{it} b_{jt} \sim \begin{pmatrix} -1 & +1 \\ 2p^+ p^- & (p^+)^2 + (p^-)^2 \end{pmatrix} \\ = \begin{pmatrix} -1 & +1 \\ q^- & q^+ \end{pmatrix}, \quad \forall i, j, t, \quad (25)$$

satisfying the relationships

$$1 = q^+ + q^-, \quad (26)$$

$$\langle b_{it} b_{jt} \rangle = q^+ - q^-, \quad (27)$$

$$\text{Var}[b_{it} b_{jt}] = (q^+ + q^-) - (q^+ - q^-)^2 \\ = 1 - (q^+ - q^-)^2 = 4q^+ q^-. \quad (28)$$

Let us, now, notice that S_{ij} is a sum of independent and identically distributed (i.i.d.) Bernoulli random variables [81], the outcomes of each elementary event being, in fact, -1 and $+1$ instead of 0 and $+1$. Such a consideration allows us to find the probability distribution obeyed by S_{ij} quite straightforwardly, by starting from the one describing C_{ij} alone, i.e.,

$$P(C_{ij} = k) = \binom{T}{k} (q^+)^k (q^-)^{T-k}. \quad (29)$$

Let us, in fact, consider the change of variable $k = (T + s)/2$, leading to the binomial

$$P(S_{ij} = s) = \binom{T}{\frac{T+s}{2}} (q^+)^{\frac{T+s}{2}} (q^-)^{\frac{T-s}{2}} \quad (30)$$

that ranges from $-T$ to T . Let us, now, convince ourselves that this transformation is indeed the correct one, by providing some explicit examples: If nodes i and j establish $k = 0$ concordant motifs, they establish T discordant motifs and the signature reads $s = -T$; if nodes i and j establish $k = T/2$ concordant motifs, they establish the same number of discordant motifs and the signature reads $s = 0$; and if nodes i and j establish $k = T$ concordant motifs, they establish zero discordant motifs and the signature reads $s = T$. In other words, the aforementioned change of variable transforms a discrete distribution whose support extends from 0 to T into a discrete distribution whose support extends from $-T$ to T : More intuitively, we have moved from considering sequences of 0 s and $+1$ s to considering sequences of -1 s and $+1$ s, where each $+1$ can be identified with a concordant motif and each -1 can be identified with a discordant motif.

A comment about the parity of T and s is needed: Let us notice that, in case T is odd, s itself must be odd, assuming the values $s = -T, -T + 2, -T + 4, \dots, T$ (corresponding to $k = 0, 1, 2, \dots, T$); in case T is even, s itself must be even,

assuming the same values as above in correspondence with the same values of k .

B. Heterogeneous benchmark

Let us, now, consider the heterogeneous benchmark

$$b_{it} \sim \begin{pmatrix} -1 & +1 \\ p_{it}^- & p_{it}^+ \end{pmatrix}, \quad \forall i, t, \quad (31)$$

which, in turn, induces the finite scheme

$$\begin{aligned} b_{it} b_{jt} &\sim \begin{pmatrix} -1 & +1 \\ p_{it}^+ p_{jt}^- + p_{it}^- p_{jt}^+ & p_{it}^+ p_{jt}^+ + p_{it}^- p_{jt}^- \end{pmatrix} \\ &= \begin{pmatrix} -1 & +1 \\ q_{ijt}^- & q_{ijt}^+ \end{pmatrix}, \quad \forall i, j, t, \end{aligned} \quad (32)$$

satisfying the relationships

$$1 = q_{ijt}^+ + q_{ijt}^-, \quad (33)$$

$$\langle b_{it} b_{jt} \rangle = q_{ijt}^+ - q_{ijt}^-, \quad (34)$$

$$\begin{aligned} \text{Var}[b_{it} b_{jt}] &= (q_{ijt}^+ + q_{ijt}^-) - (q_{ijt}^+ - q_{ijt}^-)^2 \\ &= 1 - (q_{ijt}^+ - q_{ijt}^-)^2 = 4q_{ijt}^+ q_{ijt}^-. \end{aligned} \quad (35)$$

Since S_{ij} is, now, a sum of independent and not identically distributed (i.n.i.d.) Bernoulli random variables, the probability distribution obeyed by it is the Poisson binomial

$$P(S_{ij} = s) = \sum_{C_k} \left[\prod_{\mu \in C_k} q_{ij\mu}^+ \prod_{\nu \notin C_k} q_{ij\nu}^- \right], \quad (36)$$

where C_k is the set of k -tuples of which μ and ν are instances. Let us provide some explicit examples:

$$P(S_{ij} = -T) = \prod_{\alpha=1}^T q_{ij\alpha}^-, \quad (37)$$

$$P(S_{ij} = 2 - T) = \sum_{\beta=1}^T \left[q_{ij\beta}^+ \prod_{\substack{\alpha=1 \\ \alpha \neq \beta}}^T q_{ij\alpha}^- \right], \quad (38)$$

$$P(S_{ij} = 4 - T) = \sum_{\beta=1}^T \sum_{\substack{\gamma=1 \\ \gamma > \beta}}^T \left[q_{ij\beta}^+ q_{ij\gamma}^+ \prod_{\substack{\alpha=1 \\ \alpha \neq \beta, \gamma}}^T q_{ij\alpha}^- \right], \quad (39)$$

and so on. In words, if nodes i and j establish $k = 0$ concordant motifs, they establish T discordant motifs and the signature reads $s = -T$; if nodes i and j establish $k = 1$ concordant motif, they establish $T - 1$ discordant motifs and the signature reads $s = 2 - T$; and if nodes i and j establish $k = 2$ concordant motif, they establish $T - 2$ discordant motifs and the signature reads $s = 4 - T$. The same line of reasoning can be repeated for all the admissible values of k .

C. The statistical significance of similarity: a two-sided test of hypothesis

Once we have calculated the distribution of the similarity for each pair of series, we must evaluate the statistical significance of the empirical signature: Since we have a signed

quantity, we need both tails of such a distribution to carry out a two-sided test of hypothesis. In fact, we need to answer the two related questions: (1) Is the empirical value of the signature significantly different from the one expected under the chosen benchmark? and (2) if so, is the deviation negative (hence, the signature is significantly smaller than expected) or positive (hence, the signature is significantly larger than expected)?

The first question can be answered upon calculating the two-sided p -value, reading

$$p_{ij} = 2 \times \min\{F(S_{ij}^*), 1 - F(S_{ij}^*)\}, \quad (40)$$

with F being the cumulative distribution function, defined as $F(S_{ij}^*) = \sum_{x \leq S_{ij}^*} P(S_{ij} = x)$: Such a number evaluates the probability of observing a deviation from the expected value in either direction.

The second question can be answered upon calculating the sign of such a deviation, by determining whether either

$$F(S_{ij}^*) < 1 - F(S_{ij}^*) \quad (41)$$

or

$$F(S_{ij}^*) > 1 - F(S_{ij}^*) \quad (42)$$

holds true. In the first case, $F(S_{ij}^*) < 1/2$, the empirical value of the signature is smaller than the median of the distribution and the deviation is negative; in the second case, $F(S_{ij}^*) > 1/2$, the empirical value of the signature is larger than the median of the distribution and the deviation is positive. Intuitively [82],

(1) the two conditions $F(S_{ij}^*) < 1/2$ and $p_{ij} \leq p_{th}$ indicate that nodes i and j have established a number of discordant motifs that is so large to induce a significantly negative signature—and, potentially, a negative link in the projection ($a_{ij} = -1$). In a sense, the two series are “in counterphase” most of the time, an expression that, in a more neuroscientific jargon, is intended to mean “found to be *asynchronously* active a significantly large amount of the time”;

(2) the two conditions $F(S_{ij}^*) > 1/2$ and $p_{ij} \leq p_{th}$ indicate that nodes i and j have established a number of concordant motifs that is so large to induce a significantly positive signature—and, potentially, a positive link in the projection ($a_{ij} = +1$). In a sense, the two series are “in phase” most of the time, an expression that, in a more neuroscientific jargon, is intended to mean “found to be *synchronously* active a significantly large amount of the time”;

(3) the condition $p_{ij} > p_{th}$ individuates a value of the signature (i.e., of concordant/discordant motifs) that is *compatible* with the one predicted by the chosen benchmark: Stated otherwise, the empirical number of motifs could have been observed in configurations generated by the benchmark itself—hence, inducing a null link in the projection ($a_{ij} = 0$).

Let us stress that the signs populating a projection carry a relative meaning, i.e., positive and negative signatures are so with respect to a benchmark: More explicitly, two series are revealed “in counterphase” if “not enough in phase” with respect to the benchmark. For a graphical representation of our validation procedure, see Fig. 2.

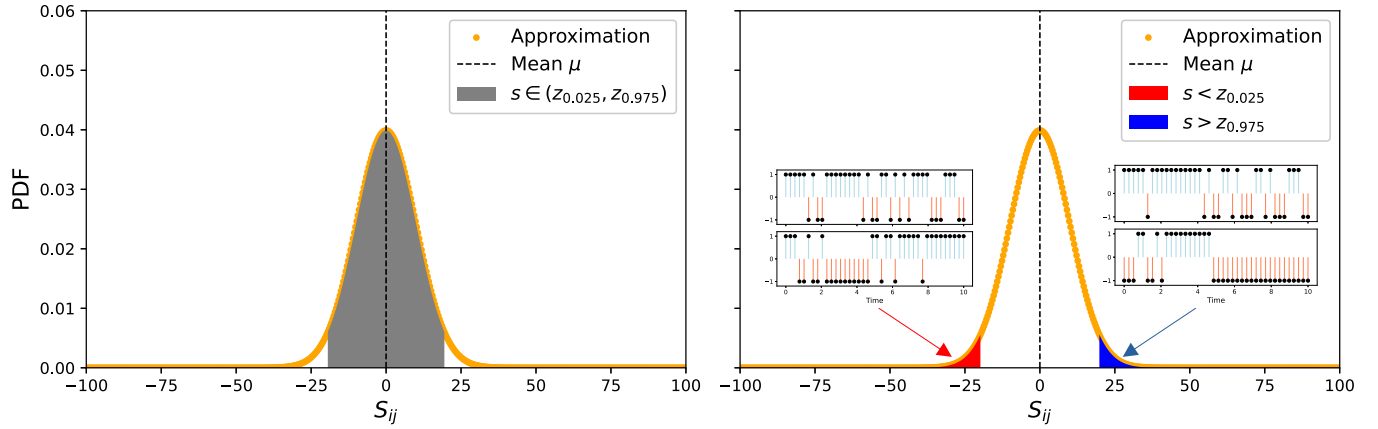


FIG. 2. Infographics illustrating our validation procedure. Probability distribution of the signature and its Gaussian approximation: While the left panel provides a graphical answer to the question “Is the empirical value of the signature significantly different from the one expected under the chosen benchmark?,” the right panel provides a graphical answer to the question “Is the deviation negative (hence, the signature is significantly smaller than expected) or positive (hence, the signature is significantly larger than expected)?” The red area corresponds to the region of validation of the negative links, induced by pairs of series “in counterphase” (i.e., *asynchronously* active) most of the time; the blue area corresponds to the region of validation of the positive links, induced by pairs of series “in phase” (i.e., *synchronously* active) most of the time.

V. VALIDATING THE PROJECTION OF A MULTIVARIATE TIME SERIES

The second step of our method returns a symmetric matrix of p -values. Individuating the ones associated with the hypotheses to be actually rejected requires a procedure to deal with the comparison of multiple hypotheses at the same time. In very general terms, a threshold must be set: If the specific p -value is smaller than it, the associated event is interpreted as statistically significant and the corresponding nodes are connected in the projection.

One approach is that of adopting the “standard” threshold $\alpha = 0.05$ and carrying out each test independently from the others but such a procedure leads to increase [83] the number of incorrectly rejected null hypotheses (i.e., validated links). An alternative approach is that of adopting the Bonferroni correction [84–86] but such a procedure is known to severely increase [87] the number of incorrectly retained null hypotheses (i.e., discarded links).

In the present paper, we apply the so-called false discovery rate (FDR) procedure [88]. FDR allows one to control for the expected number of “false discoveries” (i.e., incorrectly rejected null hypotheses or incorrectly validated links), regardless of the independence of the hypotheses tested. Whenever $|H|$ different hypotheses $H_1, H_2, \dots$, characterized by $|H|$ different p -values, must be tested at the same time, FDR prescribes to first sort the $|H|$ p -values in increasing order

$$p\text{-value}_1 \leq \dots \leq p\text{-value}_{|H|} \quad (43)$$

and, then, identify the largest integer $\hat{i}$ satisfying the condition

$$p\text{-value}_{\hat{i}} \leq \frac{\hat{i}\alpha}{|H|}, \quad (44)$$

with α representing the usual single-test significance level, hereby set to the value $\alpha = 0.05$: Since H_i is the hypothesis that the distribution of dyadic motifs established by the i th pair of nodes follows a binomial/Poisson binomial, rejecting

it amounts to connecting the corresponding nodes in the projection. Notice that $|H| = N(N-1)/2$.

Naturally, deciding which test is the best-suited one for the problem at hand depends on the importance assigned to false positives and false negatives: As a rule of thumb, the Bonferroni correction can be deemed as appropriate when few tests are expected to be significant (i.e., when even a single false positive would be problematic); when, on the contrary, many tests are expected to be significant, using the Bonferroni correction may, in turn, produce a too large number of false negatives.

VI. BENCHMARKS FOR THE ANALYSIS OF MULTIVARIATE TIME SERIES

To evaluate the statistical significance of the similarity between any two series i and j , a benchmark is required. As anticipated, a natural choice is that of adopting the class of null models known as ERGs. Within such a framework, the constrained maximization of Shannon entropy

$$S = - \sum_{\mathbf{B} \in \mathbb{B}} P(\mathbf{B}) \ln P(\mathbf{B}) \quad (45)$$

leads to assign the generic biadjacency matrix $\mathbf{B}$ the probability

$$P(\mathbf{B}) = \frac{e^{-H(\theta, \mathbf{C}(\mathbf{B}))}}{Z(\theta)}, \quad (46)$$

whose value is determined by the vector $\mathbf{C}(\mathbf{B})$ of topological constraints via the Hamiltonian $H(\theta, \mathbf{C}(\mathbf{B})) = \theta \times \mathbf{C}(\mathbf{B}) = \sum_i \theta_i C_i(\mathbf{B})$. The use of linear constraints allows $P(\mathbf{B})$ to be written in a factorized form, i.e., as the product of pair-specific probability coefficients [71].

In order to determine the unknown parameters θ , the likelihood maximization recipe can be adopted: Given an observed biadjacency matrix $\mathbf{B}^*$, it translates into solving the system of

equations

$$\langle C_i \rangle(\theta) = \sum_{\mathbf{B} \in \mathbb{B}} P(\mathbf{B}) C_i(\mathbf{B}) = C_i(\mathbf{B}^*), \quad \forall i, \quad (47)$$

which prescribes to equate each ensemble average to its observed counterpart.

Let us now generalize such a framework to accommodate models for studying binary multivariate time series. To this aim, we will follow Refs. [69,89], which, in turn, builds upon the analytical approach introduced in Ref. [66] and further developed in Ref. [68].

Let us, now, make a general observation: As each of the N time series is defined by T time steps, in corresponding with which either the value -1 or the value $+1$ can be assumed, the ensemble is constituted by all possible $|\mathbb{B}| = 2^{N \times T}$ binary, signed biadjacency matrices.

A. Signed random graph model

The first null model we consider is induced by the Hamiltonian

$$H(\mathbf{B}) = \alpha B^+(\mathbf{B}) + \gamma B^-(\mathbf{B}), \quad (48)$$

i.e., by the two global constraints $B^+(\mathbf{B})$ and $B^-(\mathbf{B})$. We will refer to it as Bipartite Signed Random Graph Model (bSRGM) [69].

According to the bSRGM, each element of a multivariate time series is a random variable whose behavior is described by the coefficients

$$P(b_{it} = +1) = \frac{e^{-\alpha}}{e^{-\alpha} + e^{-\gamma}} = p^+, \quad (49)$$

$$P(b_{it} = -1) = \frac{e^{-\gamma}}{e^{-\alpha} + e^{-\gamma}} = p^-. \quad (50)$$

In other words, b_{it} obeys a generalized Bernoulli distribution whose probability coefficients are determined by the (Lagrange multipliers of the) imposed constraints: Each positive link appears with probability p^+ and each negative link appears with probability p^- .

In order to employ the bSRGM for studying real-world multivariate time series, the parameters that define it need to be properly tuned: More specifically, one needs to ensure

$$\langle B^+ \rangle_{\text{bSRGM}} = B^+(\mathbf{B}^*), \quad (51)$$

$$\langle B^- \rangle_{\text{bSRGM}} = B^-(\mathbf{B}^*), \quad (52)$$

with the symbol $\mathbf{B}^*$ indicating the empirical multivariate time series under analysis. To this aim, one can maximize the likelihood function $\mathcal{L}_{\text{bSRGM}}(\alpha, \gamma) = \ln P_{\text{bSRGM}}(\mathbf{B}^* | \alpha, \gamma)$, where $P_{\text{bSRGM}}(\mathbf{B}^* | \alpha, \gamma)$ is defined as in Eq. (A6), with respect to the unknown parameters that define it [67]. Such a recipe leads us to find

$$p^+ = B^+(\mathbf{B}^*) / (N \times T), \quad (53)$$

$$p^- = B^-(\mathbf{B}^*) / (N \times T), \quad (54)$$

with obvious meaning of the symbols (see also Appendix A).

B. Signed configuration model

The second null model we consider is induced by the Hamiltonian

$$H(\mathbf{B}) = \sum_{i=1}^N [\alpha_i k_i^+(\mathbf{B}) + \gamma_i k_i^-(\mathbf{B})] + \sum_{t=1}^T [\delta_t \kappa_t^+(\mathbf{B}) + \eta_t \kappa_t^-(\mathbf{B})], \quad (55)$$

i.e., by the $2(N + T)$ local constraints $\{k_i^+(\mathbf{B})\}_{i=1}^N, \{k_i^-(\mathbf{B})\}_{i=1}^N, \{\kappa_t^+(\mathbf{B})\}_{t=1}^T$, and $\{\kappa_t^-(\mathbf{B})\}_{t=1}^T$. We will refer to it as Bipartite Signed Configuration Model (bSCM) [69].

According to the bSCM, each element of a multivariate time series is a random variable whose behavior is described by the coefficients

$$P(b_{it} = +1) = \frac{e^{-(\alpha_i + \delta_t)}}{e^{-(\alpha_i + \delta_t)} + e^{-(\gamma_i + \eta_t)}} = p_{it}^+, \quad (56)$$

$$P(b_{it} = -1) = \frac{e^{-(\gamma_i + \eta_t)}}{e^{-(\alpha_i + \delta_t)} + e^{-(\gamma_i + \eta_t)}} = p_{it}^-. \quad (57)$$

In other words, b_{it} obeys a generalized Bernoulli distribution whose probability coefficients are determined by the (Lagrange multipliers of the) imposed constraints: The series i , at time t , assumes the value $+1$ with probability p_{it}^+ and the value -1 with probability p_{it}^- .

In order to tune the parameters defining the bSCM to ensure

$$\langle k_i^+ \rangle_{\text{bSCM}} = k_i^+(\mathbf{B}^*), \quad \forall i, \quad (58)$$

$$\langle k_i^- \rangle_{\text{bSCM}} = k_i^-(\mathbf{B}^*), \quad \forall i, \quad (59)$$

$$\langle \kappa_t^+ \rangle_{\text{bSCM}} = \kappa_t^+(\mathbf{B}^*), \quad \forall t, \quad (60)$$

$$\langle \kappa_t^- \rangle_{\text{bSCM}} = \kappa_t^-(\mathbf{B}^*), \quad \forall t, \quad (61)$$

let us maximize the likelihood function $\mathcal{L}_{\text{bSCM}}(\alpha, \gamma, \delta, \eta) = \ln P_{\text{bSCM}}(\mathbf{B}^* | \alpha, \gamma, \delta, \eta)$, where $P_{\text{bSCM}}(\mathbf{B}^* | \alpha, \gamma, \delta, \eta)$ is defined as in Eq. (A14), with respect to the unknown parameters that define it [67] (see also Appendix A). Such a recipe leads us to find

$$k_i^+(\mathbf{B}^*) = \sum_{t=1}^T \frac{e^{-(\alpha_i + \delta_t)}}{e^{-(\alpha_i + \delta_t)} + e^{-(\gamma_i + \eta_t)}} = \langle k_i^+ \rangle, \quad \forall i, \quad (62)$$

$$k_i^-(\mathbf{B}^*) = \sum_{t=1}^T \frac{e^{-(\gamma_i + \eta_t)}}{e^{-(\alpha_i + \delta_t)} + e^{-(\gamma_i + \eta_t)}} = \langle k_i^- \rangle, \quad \forall i, \quad (63)$$

$$\kappa_t^+(\mathbf{B}^*) = \sum_{i=1}^N \frac{e^{-(\alpha_i + \delta_t)}}{e^{-(\alpha_i + \delta_t)} + e^{-(\gamma_i + \eta_t)}} = \langle \kappa_t^+ \rangle, \quad \forall t, \quad (64)$$

$$\kappa_t^-(\mathbf{B}^*) = \sum_{i=1}^N \frac{e^{-(\gamma_i + \eta_t)}}{e^{-(\alpha_i + \delta_t)} + e^{-(\gamma_i + \eta_t)}} = \langle \kappa_t^- \rangle, \quad \forall t, \quad (65)$$

a system that can be solved only numerically, along the guidelines provided in Ref. [90] (see also Appendix B).

VII. DATASET DESCRIPTION

We have applied the methodology above to 3T resting-state fMRI data (rs-fMRI) data from 100 unrelated subjects of the Young Adult Human Connectome Project (HCP) [91].

Overall, this dataset includes 1200 participants aged 22–35 of whom 1087 underwent at least one rs-fMRI scan. Spontaneous, slowly fluctuating brain activity was measured during fMRI, using a multiband accelerated echo-planar imaging two-dimensional BOLD sequence with whole-brain coverage, while subjects maintained fixation on a central cross and were instructed to lie still and rest quietly—TR = 720 ms, TE = 33.1 ms, flip angle = 52°, FOV = 208 mm × 180 mm (RO × PE), matrix = 104 × 90 (RO × PE), slice thickness = 2.0 mm, number of slices = 72, resolution = 2.0 mm isotropic voxels, multiband factor = 8, echo spacing = 0.58 ms, BW = 2290 Hz/px, frames per run = 1200, and run duration (min:s) = 14:33. Within each session, oblique axial acquisitions alternated between phase encoding in a right-to-left direction in one run and phase encoding in a left-to-right direction in the other run.

A. Preprocessing of time series

fMRI recordings underwent a two-step preprocessing procedure. First, the data were motion corrected, intensity adjusted, and normalized [92]. Afterward, structured noise was corrected using the FIX denoising pipeline [93]. This cleanup included 24 confound time series derived from the motion estimation (the 6 rigid-body parameter time series, their backward-looking temporal derivatives, and all 12 resulting regressors squared [94]). On the motion parameters, the temporal high-pass filtering was previously applied. They are, then, regressed out of the data aggressively, as they are not expected to contain variance of interest. To improve the length of BOLD time series, the right-to-left and left-to-right runs were concatenated after the mean across time removal. Finally, each brain was parcellated into 116 regions of interest (ROIs) over which the signal was averaged. More specifically, the cortical regions were parcellated following the Schaefer 100 atlas [95] and the subcortical (SUBC) ones following the Tian 16 atlas [96]. As a final result, we obtained a collection of 100 sets of time series, each one having dimensions $N \times T = 116 \times 2400$.

Let us, finally, notice that preprocessing procedures such as the global signal regression (GSR) have been criticized since biasing data to balance the number of positive and negative signs [97]: For such a reason, we have opted for not adopting it here.

B. Prewhitening of time series

A documented property of rs-fMRI signals concerns the presence of a non-negligible amount of temporal autocorrelations: In other words, the sluggish nature of the hemodynamic response, which extends over timescales of $\simeq 20$ s, lets the BOLD response at time t to depend on its response at previous time steps [98–100]; since this may bias the number of co-occurrences, here we address such an issue by implementing a prewhitening procedure aimed at removing it.

More quantitatively, we model the time series X_{it} as

$$X_{it} = \sum_{k=1}^{p_i} \phi_k^{(i)} X_{it-k} + \varepsilon_{it}, \quad (66)$$

i.e., as an autoregressive process of order p_i , where $\phi_1^{(i)} \dots \phi_{p_i}^{(i)}$ are the coefficients and ε_{it} is the residual, i.e., the component of the i th signal at time t that cannot be (linearly) predicted by employing the p_i previous values: The autoregressive coefficients have been estimated via ordinary least squares separately for each of the $N = 116$ ROIs [101]; orders have been evaluated separately as well, the optimum $\hat{p}_i$ having been selected via Bartlett's white noise test [102–104] (see also Appendix C). Once the coefficients and the order have been estimated, the prewhitened residuals have been computed as

$$\varepsilon_{it} = X_{it} - \sum_{k=1}^{\hat{p}_i} \hat{\phi}_k^{(i)} X_{it-k}, \quad t = \hat{p}_i + 1, \dots, T. \quad (67)$$

Since different ROIs have different orders, the series of residuals have different lengths—for instance, series i has a length of $T - \hat{p}_i$: As a consequence, the series of residuals must be “aligned.” To this aim, we have adopted a *trimming strategy*, implementing the so-called *trailing trim*, which prescribes to retain the last $T_{\text{eff}} = T - \max_i \hat{p}_i$ points of each series.

Overall, thus, we are left with an $N \times T_{\text{eff}}$ rectangular matrix of residuals: The prewhitened binarized series are, then, obtained by applying Iverson's brackets to each entry of such a standardized matrix, i.e.,

$$w_{it} = \frac{\varepsilon_{it} - \underline{\varepsilon}_i}{\sigma_i[\varepsilon_i]}, \quad \forall i, t, \quad (68)$$

where $\underline{\varepsilon}_i$ and $\sigma_i[\varepsilon_i]$ represent the (sample) average and standard deviation of series $\varepsilon_i = \{\varepsilon_{it}\}_{t=1, \dots, T}$; in formulas,

$$b_{it}^- = [w_{it} < 0], \quad \forall i, t, \quad (69)$$

$$b_{it}^+ = [w_{it} > 0], \quad \forall i, t. \quad (70)$$

VIII. RESULTS

Let us, now, employ our validation procedure to carry out a “binary-to-binary” analysis, i.e., obtain statistically validated binary projections of binary multivariate time series, defined by the “rule” that any two nodes sharing a significantly large number of concordant (discordant) motifs are connected by a positive (negative) edge.

A. Consistency checks on the signature

Before moving to analyze our data, let us clarify the nature and role of S_{ij} , verifying the extent to which S_{ij} retains the sign of the Pearson correlation coefficient [see Eq. (1)] between the time series i and j , across all subjects. To this aim, we have proceeded in two ways: the first one, by identifying i and j with $w_i = \{w_{it}\}_{t=1, \dots, T}$ and $w_j = \{w_{jt}\}_{t=1, \dots, T}$; and the second one, by identifying i and j with $b_i = \{b_{it}\}_{t=1, \dots, T}$ and $b_j = \{b_{jt}\}_{t=1, \dots, T}$. In the first case, the percentage of concordant, positive signs amounts to $\simeq 97\%$, while the percentage of concordant, negative signs amounts to $\simeq 78\%$; and in the second case, the percentages above rise to $\simeq 99\%$.

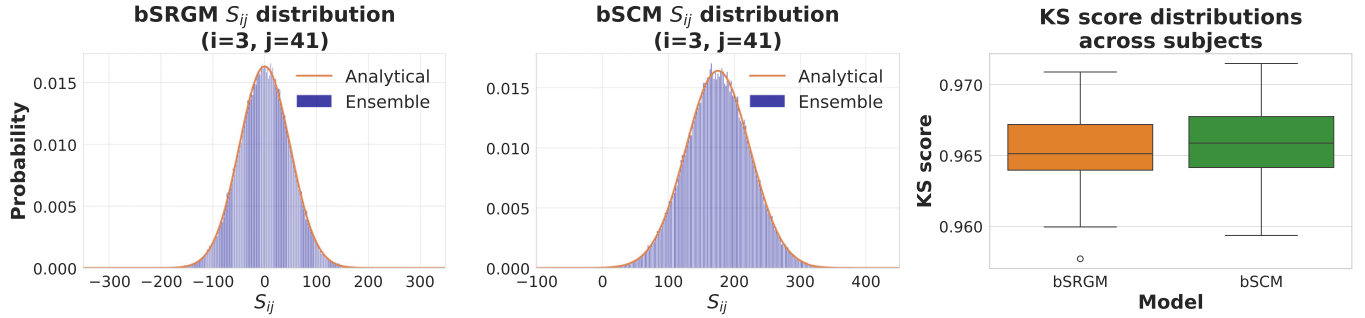


FIG. 3. Consistency checks regarding the distribution of the signature for subject No. 100307. Comparison between the distribution of the dyadic signature induced by the bSRGM/bSCM (left panel/middle panel) and their numerical counterparts, obtained by explicitly sampling 10^5 realizations from the corresponding ensemble: In both cases, the KS test rejects the hypothesis that the two distributions coincide. The box plots summing up the distributions of the KS scores for the bSRGM and the bSCM, obtained by explicitly sampling 10^3 realizations from the corresponding ensemble (right panel), show that, under both models, the median KS score amounts to $\text{KS}_{5\%} \simeq 0.965$, the 95% of the values ranging between the 2.5th and 97.5th percentiles reading $[q_{2.5}, q_{97.5}] = [0.961, 0.970]$. In words, our sampling procedure can be considered satisfactory enough to reproduce the analytical distributions defined by Eqs. (30) and (36), even in case of statistical disagreement.

Let us, now, characterize our projections in a more quantitative way: To this aim, we will indicate the adjacency matrix of a generic projection with $\mathbf{A}$ and refer to the naive ones [see Eq. (22)], those validated via the bSRGM, and those validated via the bSCM with $\mathbf{A}_{\text{naive}}$, $\mathbf{A}_{\text{bSRGM}}$, and $\mathbf{A}_{\text{bSCM}}$, respectively.

Let us, now, verify that the ensemble distribution of each pair-specific signature is consistent with its analytical counterpart, for all benchmarks considered here. Specifically, let us focus on each subject and test (1) whether the distribution defined by Eq. (30) is compatible with the ensemble distribution of S_{ij} , $\forall i < j$ induced by the bSRGM; and (2) whether the distribution defined by Eq. (36) is compatible with the ensemble distribution of S_{ij} , $\forall i < j$ induced by the bSCM. Each subject can be associated with a Kolmogorov-Smirnov (KS) score, i.e., the percentage of times the KS test (at the level of 5%) does not reject the hypothesis of compatibility: While a value of 1 indicates perfect compatibility, a value of 0 indicates perfect incompatibility. The results are summed up by the box plots depicted in Fig. 3: The median KS score is $\text{KS}_{5\%} \simeq 0.965$ under both the bSRGM and the bSCM, the 95% of these distributions ranging between the 2.5th and the 97.5th percentiles reading $[q_{2.5}, q_{97.5}] = [0.961, 0.970]$; in words, our sampling procedure satisfactorily reproduces the analytical distributions defined by Eqs. (30) and (36) even when the KS test rejects the hypothesis of compatibility (e.g., for subject No. 100307).

B. Patterns of brain (im)balance at the microscale

1. Connectance

When considering the unsigned case, the connectance of the projection returned by the naive approach is typically large [105]: This holds true here as well, for all the considered subjects (see Fig. 4). It is worth noticing that many validated links are positive. This leads us to conclude that the time series associated with the related pairs of regions are concordant for the vast majority of the snapshots constituting the considered time window.

The positive connectance, defined as $\rho^+ = 2L^+/N(N-1)$ with $L^+ = \sum_{i=1}^N \sum_{j(>i)} a_{ij}^+$, decreases as we employ

stricter benchmarks: On average, in fact, $\rho_{\text{naive}}^+ = 0.93$ while $\rho_{\text{bSRGM}}^+ = 0.75$ and $\rho_{\text{bSCM}}^+ = 0.22$. Such a result can be explained upon considering that concordant motifs represent the majority of motifs: Filtering to a larger extent, thus, leads to a projection characterized by a smaller amount of positive links—precisely the ones induced by the aforementioned patterns. On the other hand, the negative connectance, defined as $\rho^- = 2L^-/N(N-1)$ with $L^- = \sum_{i=1}^N \sum_{j(>i)} a_{ij}^-$, shows a (much) less trivial behavior as, on average, $\rho_{\text{naive}}^- = 0.07$ while $\rho_{\text{bSRGM}}^- = 0.01$ and $\rho_{\text{bSCM}}^- = 0.32$.

Such a result can be explained upon considering that the global, bSRGM-based filter is defined by a binomial distribution centered on not-that-far right values of the signature: in fact, $c^- = B^-/(N \times T) \gtrsim c^+ = B^+/(N \times T)$, as Fig. 4 confirms. Since the bipartite density of negative links is only slightly larger than the bipartite density of positive links, $\langle D_{ij} \rangle \gtrsim \langle C_{ij} \rangle$, i.e., $\langle S_{ij} \rangle \lesssim 0$. As a consequence, several “naively negative” signatures are no longer deemed as significant and the resulting negative links are cut; positive links, instead, are overrepresented as the “validation threshold” becomes smaller: The global filter is, thus, stricter with negative than with positive links.

The local, bSCM-based filter, instead, accounts for nodes heterogeneity: Whenever the positive degree of a node is (much) larger than its negative degree, the distribution of its similarity with any other node moves to the (far) right, hence validating negative links to a much larger extent; the local filter is, thus, stricter with positive than with negative links, somehow rebalancing their number. In other words, when two regions are each strongly connected to most of the rest of the brain, yet share only a slight tendency toward “in-phase” dynamics with one another, we assign a negative connection between them—loosely speaking, the bSCM work as an “enhancer” of negative signs, increasing the likelihood of observing asynchronies between pairs of regions.

2. Tendency to make hubs

A related index is the one named *tendency to make hubs*, introduced in Ref. [34] and quantifying the tendency of a given degree distribution—when dealing with signed graphs,

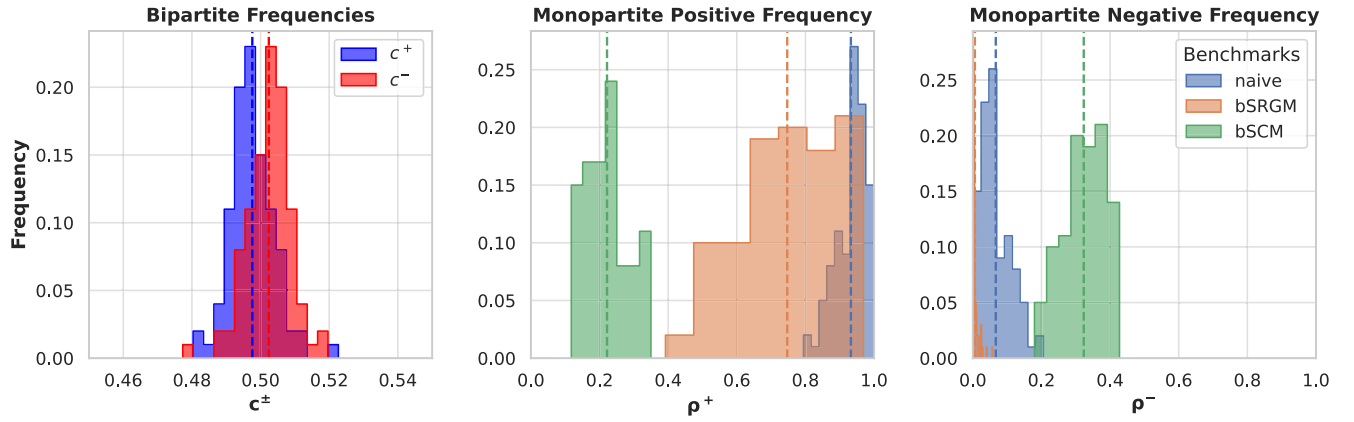


FIG. 4. Distributions of the signed, bipartite, and monopartite connectance. Left panel: Distributions of the positive and negative bipartite connectance, respectively, defined as $c^+ = B^+/(N \times T)$ and $c^- = B^-/(N \times T)$, across our 100 subjects. Middle and right panels: Distributions of the positive and negative monopartite connectance, respectively, defined as $\rho^+ = 2L^+/N(N-1)$ and $\rho^- = 2L^-/N(N-1)$, across the three different projections of our 100 subjects. Bin widths are computed according to the Freedman-Diaconis rule. Averages are plotted as vertical, dashed lines. How can a large number of bipartite, negative links coexist with the large number of positive links characterizing the naive projection? Bipartite, negative links are evidently arranged into concordant motifs that, in turn, let the naive and bSRGM-induced projections be populated by a majority of positive links. Employing the bSCM mitigates this situation, as the number of positive and negative links is, now, rebalanced.

both the positive and the negative degree distributions need to be considered—to host hubs. Hereby, we adopt a slightly different index, rooted in statistical theory and reducing the functional form of a given degree distribution to just one number: named *coefficient of variation* (CV), it is defined as the ratio between the standard deviation and the expected value of the reference distribution, i.e.,

$$CV^+ = \frac{\text{std}[k^+]}{k^+} = \frac{\text{std}[k^+]}{2L^+/N} \quad (71)$$

and

$$CV^- = \frac{\text{std}[k^-]}{k^-} = \frac{\text{std}[k^-]}{2L^-/N}, \quad (72)$$

while a CV smaller than 1 indicates that the expected value is representative of the entire distribution (as its dispersion around the average is smaller than it), observing a CV larger than 1 indicates that the expected value is not representative of the entire distribution (as its dispersion around the average is larger than it)—from this perspective, a fat-tailed distribution is expected to obey $CV > 1$.

As evident from Fig. 5, none of our projections seems to be characterized by the presence of positive hubs ($CV_{\text{naive}}^+ = 0.09$, $CV_{\text{bSRGM}}^+ = 0.25$, $CV_{\text{bSCM}}^+ = 0.75$); when considering the negative degrees, instead, naive and bSRGM-induced distributions display a right skewness that suggests the presence of nodes whose number of connections is (much) larger than that of the others ($CV_{\text{naive}}^- = 1.64$, $CV_{\text{bSRGM}}^- = 3.16$): In a sense, thus, the subgraph induced by the positive links tends to be more homogeneous than the subgraph induced by the negative ones [106]. The bSCM, however, levels out such differences ($CV_{\text{bSCM}}^- = 0.66$).

We also explicitly show the distributions of the positive and negative degrees pooled across our subjects. As KS tests (at the level of 5%) reveal, practically all pairwise comparisons point out our three (naive, bSRGM-induced, and

bSCM-induced) positive degree distributions as significantly different; moreover, the appearance of positive hubs is enhanced when the global filter is used.

A related result concerns the degree distributions of cortical and subcortical regions: As evident from Fig. 18, subcortical regions tend to have a larger negative degree than cortical regions.

3. Spatial coherence of signed connections

From what we have learned so far, the positive subgraphs are (overall) structurally homogeneous, while the negative ones are (overall) structurally heterogeneous. Let us, now, ask ourselves how similar positive and negative subgraphs are across subjects.

In order to answer this question, we calculated the Jaccard similarity between pairs of (adjacency matrices representing) individual projections, per benchmark. As Fig. 6 shows, the positive variant of such an index (i.e., the percentage of positive links located in the same position) ranges between 0.4 and 1 while its negative variant (i.e., the percentage of negative links located in the same position) ranges between 0 and 0.4, on both the naive and bSRGM-induced projections: In words, the spatial coherence of positive links is larger than the spatial coherence of negative links, which appear as those (better) characterizing single subjects. This picture changes when bSCM-induced projections are considered: In such a case, in fact, the spatial coherence of the two kinds of links appears as very similar.

4. Assessing the brain (im)balance at the microscale

Let us now inspect the arrangement of negative links within projections. To this aim, two choices are possible: (1) the first one amounts at considering the index named *strong degree of balance* [69] and counting the triads characterized by an even number of negative links, i.e., either zero or two—in

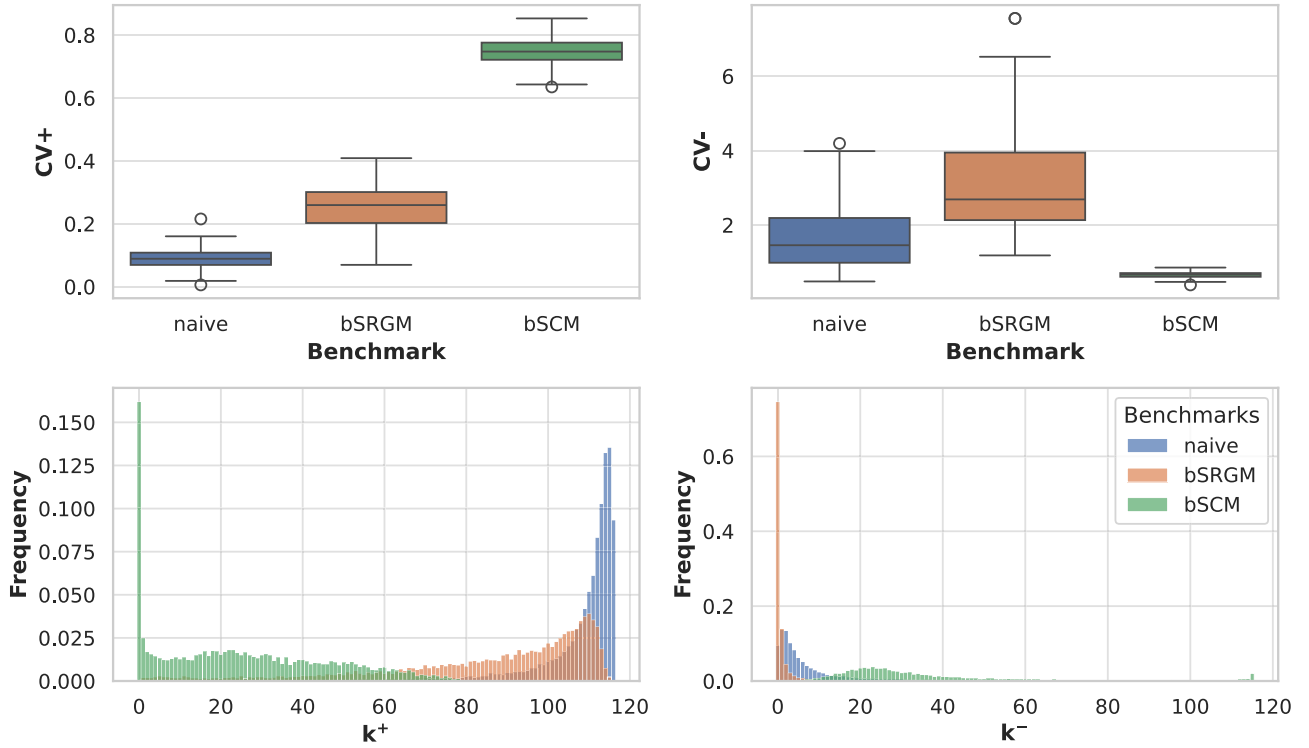


FIG. 5. Distributions of CV^+ , CV^- , and signed degrees. Top panels: Box plots summing up the distributions of the coefficient of variation for positive degrees (CV^+ , left) and negative degrees (CV^- , right) across the three different projections of our 100 subjects. Each box plot illustrates the distribution of the corresponding metric, indicating its central tendency and dispersion for each model: The positive variant of the CV suggests that the corresponding subgraphs are characterized by a quite homogeneous structure, while the negative one suggests that the opposite holds true; yet, these differences are leveled out when considering bSCM-induced projections. Bottom panels: Distributions of the positive (left) and negative (right) degrees populating our projections, pooled across our subjects. While naive and bSRGM-induced positive distributions appear as left skewed, their negative counterparts appear as right skewed; bSCM-induced distributions, instead, are much flatter in both cases, suggesting a larger heterogeneity of the degrees of this kind. The leftmost peak on k^+ and the rightmost peak on k^- are due to subcortical regions.

other words, the ones that are balanced according to the strong balance theory; and (2) the second one amounts at considering the index named *weak degree of balance* [69] and counting the

triads characterized by zero, two, or three negative links—in other words, the ones that are balanced according to the weak balance theory.

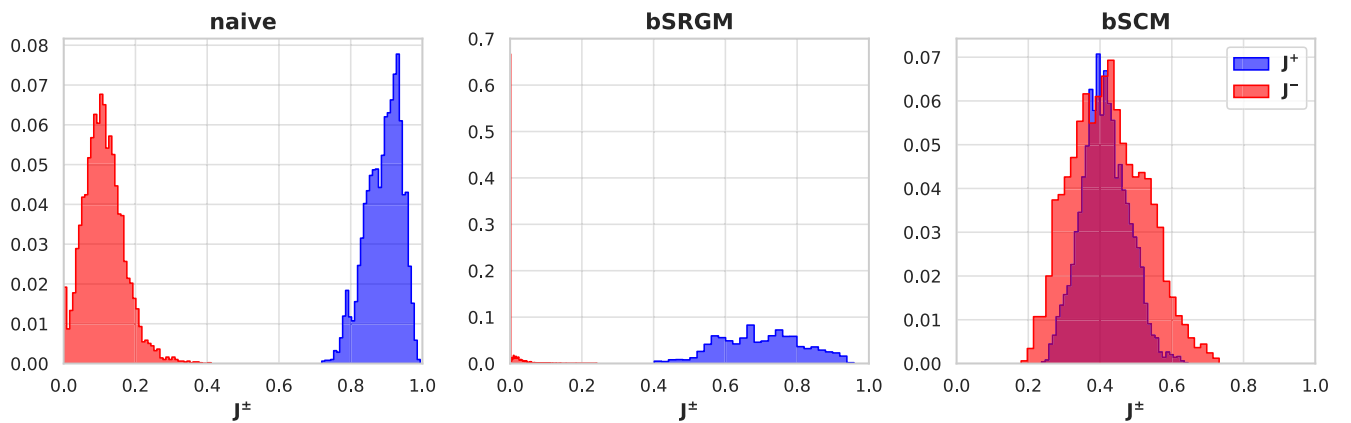


FIG. 6. Distribution of the Jaccard similarity between pairs of projections. The numerator of such an index represents the number of positive (negative) links occupying the same position in the adjacency matrix of the two subjects to compare and its denominator represents the cardinality of the union of the two sets of positive (negative) links. Bin widths are computed according to the Freedman-Diaconis rule. Since a value of 1 indicates perfect overlap while a value of 0 indicates no overlap, positive links yield a larger Jaccard similarity than negative links on both the naive and bSRGM-induced projections, a result indicating the former ones as more spatially coherent than the latter ones. When considering bSCM-induced projections, instead, the spatial coherence of the two kinds of links appears very similar.

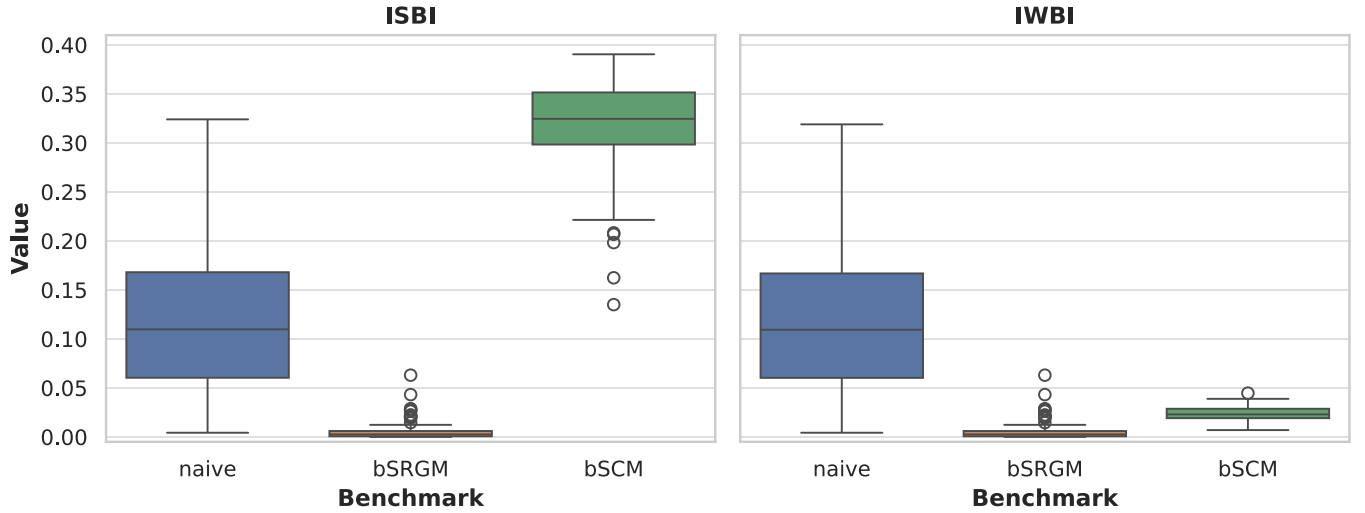


FIG. 7. Distributions of the IBI values across subjects. Box plots summing up the distributions of the values of the strong (left) and weak (right) variants of the IBI across the three different projections of our 100 subjects. Each box plot proxies the distribution of the corresponding metric, indicating its central tendency and dispersion: In both cases, the bSRGM-induced projections are characterized by the smallest median, while the naive projections are peaked in correspondence with larger values, although showing more variability. The difference between the ISBI and the IWBI values characterizing the bSCM-induced projections is due to the sensitivity of this benchmark to the abundance of the triad (---), whose presence is enhanced by the predominantly negative relationships observed between subcortical regions as well as between limbic regions (see also Fig. 8). Overall, however, each median IBI value is positive, a result confirming the non-null level of frustration characterizing humans' brain networks.

The indices above induce two variants of the index of brain imbalance (IBI—practically, their complementary to 1), i.e., the *index of strong brain imbalance*

$$\text{ISBI} = \frac{T^{(+++)} + T^{(---)}}{T^{(+++)} + T^{(---)} + T^{(+-+)} + T^{(-+-)}} \quad (73)$$

and the *index of weak brain imbalance*

$$\text{IWBI} = \frac{T^{(++)}}{T^{(+++)} + T^{(---)} + T^{(+-+)} + T^{(-+-)}}, \quad (74)$$

with

$$T^{(+++)} = \frac{\text{Tr}[(\mathbf{A}^+)^3]}{6}, \quad (75)$$

$$T^{(---)} = \frac{\text{Tr}[\mathbf{A}^- \mathbf{A}^+ \mathbf{A}^-]}{2}, \quad (76)$$

$$T^{(+-+)} = \frac{\text{Tr}[\mathbf{A}^+ \mathbf{A}^- \mathbf{A}^+]}{2}, \quad (77)$$

$$T^{(-+-)} = \frac{\text{Tr}[(\mathbf{A}^-)^3]}{6}. \quad (78)$$

Naturally, the closer their value to 1, the larger the amount of unbalanced triads.

As Fig. 7 shows, the bSRGM-induced projections are the ones characterized by the largest amount of balanced triads—a result that is not surprising once considering the small amount of negative links populating them; the naive projections, instead, are characterized by a number of balanced triads that amounts to 5–6 times the number of unbalanced triads. Finally, and most interestingly, the behavior of the bSCM-induced projections is sensitive to that of the triad (---), which can be quite abundant: Such a behavior characterizes the subcortical and the limbic regions, partly reconciling our findings with those in Ref. [107,108]. Overall,

however, each median IBI value is positive [109], a result confirming the non-null level of frustration characterizing humans' brain networks: If, according to Ref. [34], the value $\text{IBI} = 0$ is interpreted as the one characterizing the energy ground state, the aforementioned results imply that humans' brain networks are found in some sort of excited state.

Let us, now, consider the brain parcellation carried out in Ref. [110], where the authors propose a functional partition of brain regions by distinguishing communities of regions (hereby, *Yeo's networks* or *communities*) involved in sensory tasks (e.g., the visual and the somatomotor networks) from those involved in higher-order, cognitive tasks [e.g., the *default mode network* (DMN)]: They are named *control* (CONT), *dorsal attention* (DA), *limbic* (LIM), *somatomotor* (SM), *saliency/ventral attention* (SVA), *visual* (VIS) networks, and DMN. While the analysis carried out in Ref. [73] highlights the presence of negative correlations between the DMN and other networks, the one carried out in Ref. [72] shows that, in healthy individuals, unbalanced triads predominantly occur between different Yeo's networks.

In order to verify the first of the aforementioned findings, we have constructed two 8×8 matrices: Each entry of the first (second) one encodes the number of positive (negative) links between the corresponding Yeo's networks—to which the *subcortical regions* have been added—aggregated across individuals. As Fig. 8 shows, (1) the largest number of positive links is found within the DMN; and (2) a large number of negative links is found between the DMN and every other network, as well as between the SUBC and every other network. While the first result confirms the empirical “in-phase” dynamics of the ROIs constituting the default mode network, the second one seems to (at least) partly reflect a filter-induced sign reversal; indeed, the subcortical regions are characterized

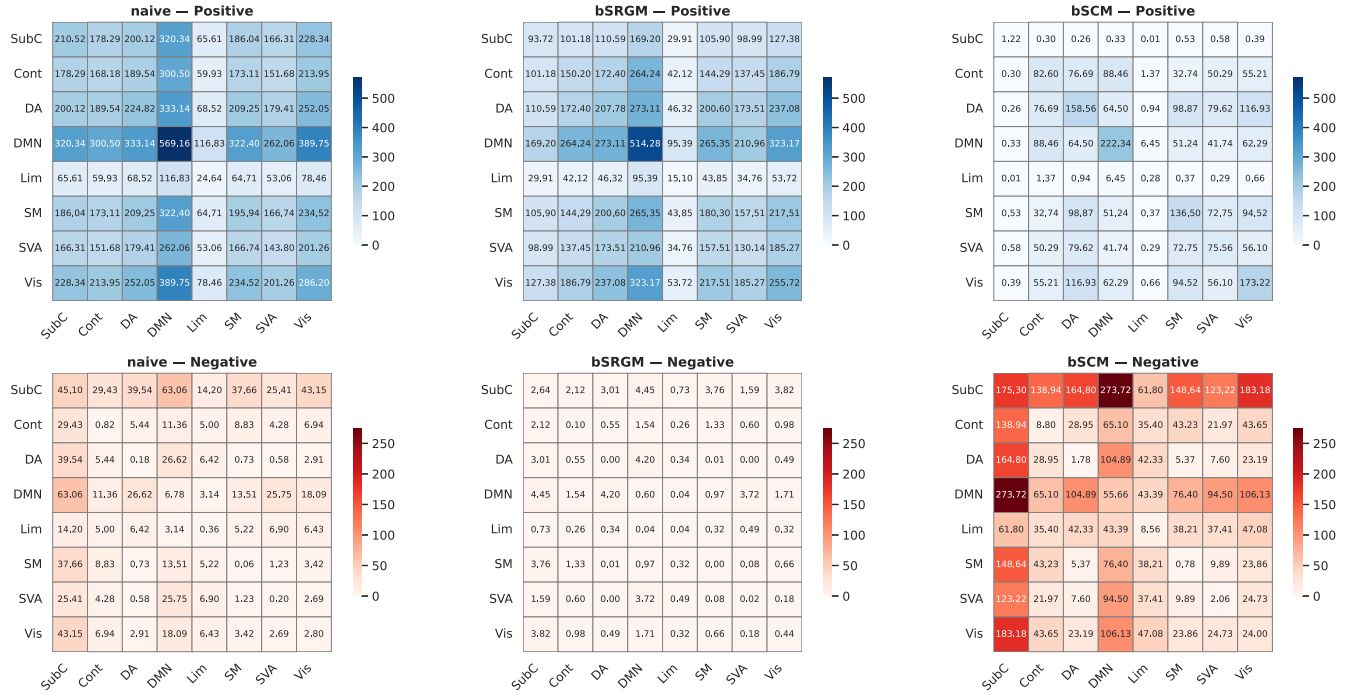


FIG. 8. Heatmaps of the signed connections within and between Yeo's brain networks. Values are averaged across subjects for each of our benchmarks. The intensity of each color is proportional to the magnitude of the corresponding entry. Network labels stand for subcortical (SUBC), control (CONT), dorsal attention (DA), limbic (LIM), somatomotor (SM), salience/ventral attention (SVA), visual (VIS) networks, and default mode network (DMN). The neatest signal is observed for the bSCM-induced projections: While the largest number of positive links is found within the DMN, a large number of negative links is found between the DMN and every other network, as well as between the SUBC and every other network—the largest number of negative connections being located precisely between the DMN and the SUBC. Notice that, concerning the blocks lying on the main diagonal, $L^- > L^+$ (only) for the SUBC and LIM networks (see also Fig. 7).

by one of the highest degrees of independence from the remaining ROIs, comparable only to that of the limbic regions: As a consequence, the corresponding value of the signature falls below the validation threshold imposed by the local filter.

In order to verify the second finding above, we have inspected the distributions of the ISBI values associated with the admissible triplets of regions: As evident from Fig. 9, (1) the SUBC triplets are those exhibiting the highest level of imbalance [111]; (2) the cortical triplets belonging to LIM exhibit the highest level of imbalance, followed by those where the DMN is present once or, at most, twice; and (3) the cortical triplets belonging to the same Yeo's network are those exhibiting the lowest level of imbalance (the smallest values characterizing the triplets involving nodes from the DA, SM, and SVA networks)—with the exception of the regions belonging to LIM and, to a lesser extent, those belonging to DMN and CONT.

C. Patterns of brain (im)balance at the mesoscale

Following Ref. [70], one can adopt an “agnostic” attitude and explore the mesoscale organization of a projection without aligning with any specific conceptual framework. A principled approach to achieve such a goal is that of minimizing the quantity $\text{BIC} = \kappa \ln V - 2\mathcal{L}$ where, now, $V = N(N - 1)/2$. Since we aim at describing a projection mesoscale organization, a natural choice is that of instantiating the Bayesian information criterion (BIC) with the signed stochastic block

model (SSBM), defined by the likelihood function

$$\begin{aligned} \mathcal{L}_{\text{SSBM}} = & \prod_{r=1}^k (p_{rr}^+)^{L_{rr}^+} (p_{rr}^-)^{L_{rr}^-} (1 - p_{rr}^+ - p_{rr}^-)^{\binom{N_r}{2} - L_{rr}^-} \\ & \times \prod_{r=1}^k \prod_{s=1, s>r}^k (p_{rs}^+)^{L_{rs}^+} (p_{rs}^-)^{L_{rs}^-} (1 - p_{rs}^+ - p_{rs}^-)^{N_r N_s - L_{rs}^-}, \end{aligned} \quad (79)$$

and a number of parameters $\kappa_{\text{SSBM}} = k(k + 1)$, with k indicating the number of modules into which the projection is partitioned. Naturally, N_r is the number of nodes constituting block r , p_{rr}^+ is induced by the empirical number of positive links within block r , i.e.,

$$p_{rr}^+ = \frac{2L_{rr}^+}{N_r(N_r - 1)}, \quad (80)$$

p_{rr}^- is induced by the empirical number of negative links within block r , i.e.,

$$p_{rr}^- = \frac{2L_{rr}^-}{N_r(N_r - 1)}, \quad (81)$$

p_{rs}^+ is induced by the empirical number of positive links between blocks r and s , i.e.,

$$p_{rs}^+ = \frac{L_{rs}^+}{N_r N_s}, \quad (82)$$

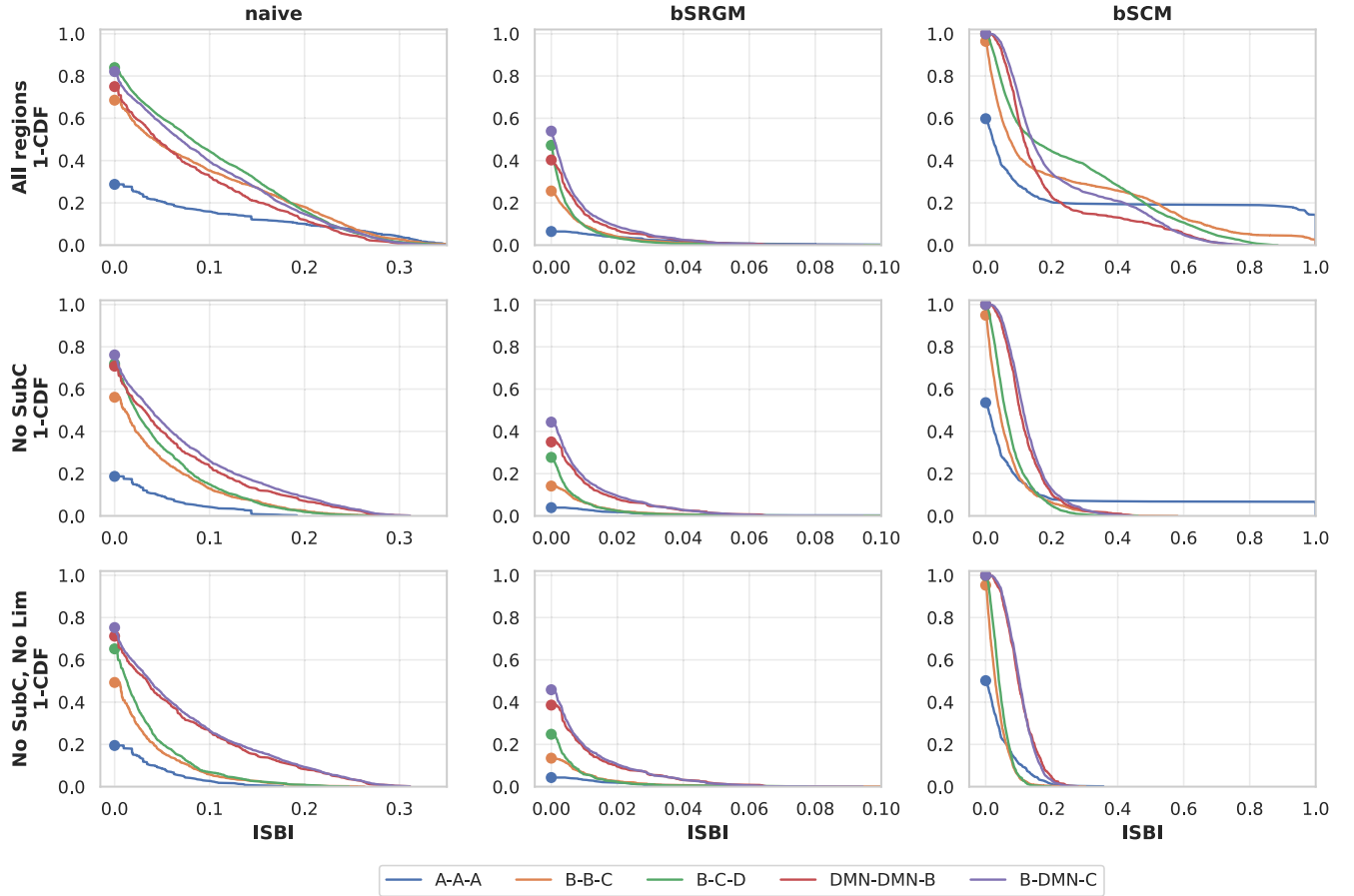


FIG. 9. Distributions of the ISBI values associated with triplets of brain regions. Cumulative distributions of the ISBI values across the three different projections of our 100 subjects: Top panels refer to all regions, middle panels refer to all regions without subcortical ones, and bottom panels refer to all regions without subcortical and limbic ones. The largest ISBI values characterize the triplets of subcortical regions: removing them leads to a different picture as the largest ISBI values, now, characterize the triplets within the LIM network; removing them as well reveals that the largest ISBI values characterize the triplets where the DMN is present once or twice—and, to a lesser extent, those within the DMN and CONT networks. The smallest ISBI values, instead, characterize those belonging to the same Yeo’s network (especially the DA, SM, and SVA networks).

and p_{rs}^- is induced by the empirical number of negative links between blocks r and s , i.e.,

$$p_{rs}^- = \frac{L_{rs}^-}{N_r N_s}. \quad (83)$$

1. A survey on individual projections

Let us start by considering individual partitions, stressing that a large link density does not prevent BIC minimization from detecting statistically significant mesoscopic structures: Its sensitivity to both the number and the sign of the links, in fact, makes it capable of partitioning both denser and sparser configurations.

Naive projections. As Fig. 10 shows, BIC minimization reveals the presence of eight communities: The vast majority of links *within* communities—if not all, as evident upon looking at some of the modules—is positive, a result implying that the series describing the activity of these regions are concordant most of the time. For what concerns the links *between* communities, instead, the picture becomes much more interesting: All communities in Fig. 10 are, in fact, connected by mostly

negative links with the eighth, an evidence suggesting that such a block of regions is “in counterphase” with every other block most of the time; the first and second communities, instead, are connected by mostly positive links, i.e., even if these blocks of regions are separated, they still seem to work “in phase” (i.e., in a *synchronous* fashion).

Validated projections. Filtering our projections with a homogeneous benchmark makes the picture even clearer, as each block is now constituted by links of practically the same sign: More precisely, while positive links constitute the majority of links, (few) negative links are found only between modules. Filtering our projections with a heterogeneous benchmark, instead, returns a picture that is halfway between the naive one and the bSRGM-induced one, as the blocks are sparser than the naive ones, less homogeneous than those induced by the bSRGM, but richer in negative connections than both. The overall message conveyed by such an analysis, however, does not change: The brain seems to be constituted by groups of regions working “in phase” most of the time and “in counterphase” (i.e., in an *asynchronous* fashion) with those of other groups; still, BIC minimization admits regions working “in

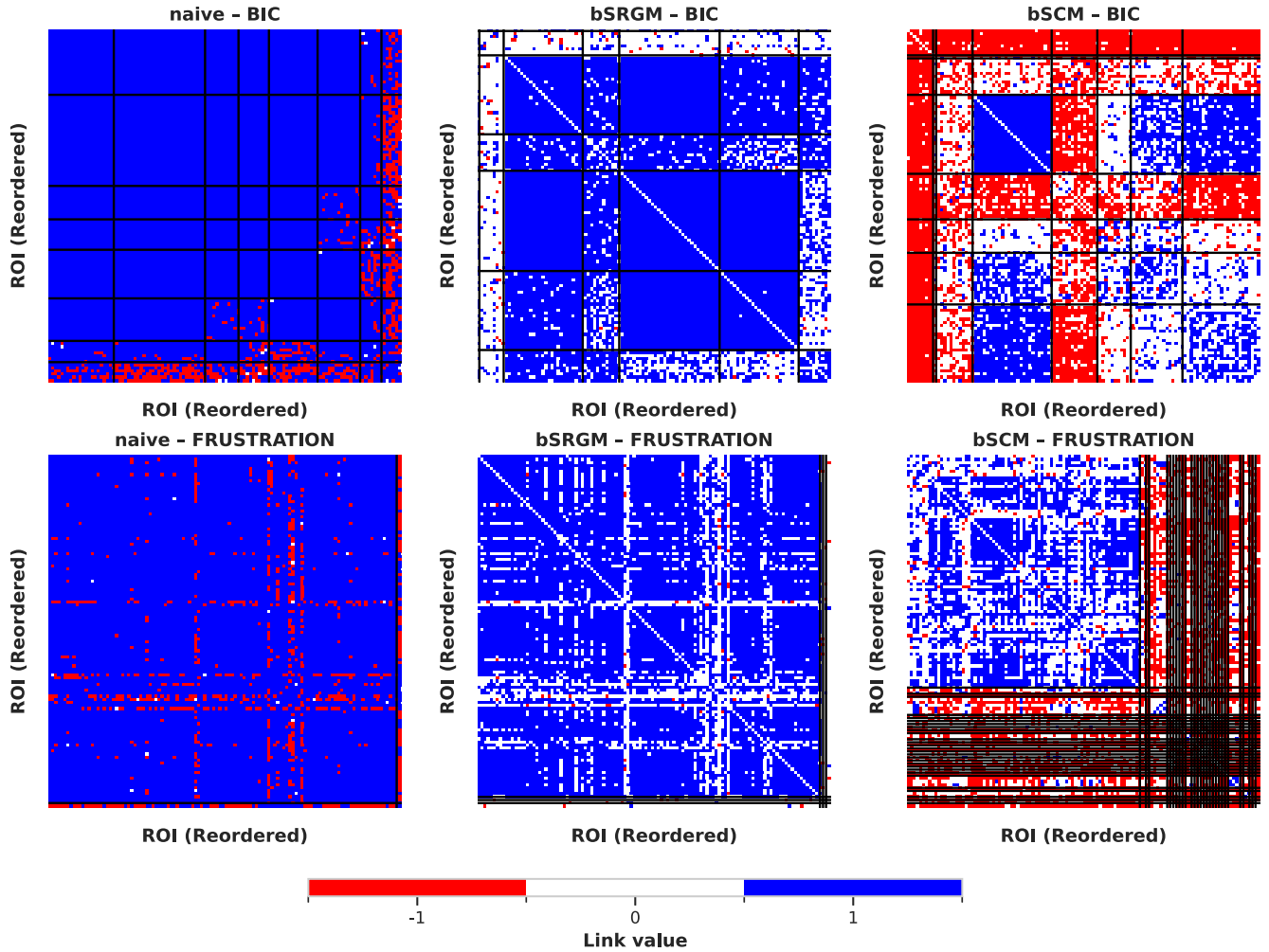


FIG. 10. Signed communities partitioning the brain of subject No. 106016. Top panels: Heatmaps illustrating the communities partitioning the projections of subject No. 106016, detected by minimizing the BIC. Bottom panels: Heatmaps illustrating the communities partitioning the projections of subject No. 106016, detected by minimizing the frustration (F). Although the number of detected communities varies, the overall message conveyed by the two methods can be summed up as follows: The brain seems to be constituted by groups of regions working “in phase” (i.e., in a *synchronous* fashion) most of the time and “in counterphase” (i.e., in an *asynchronous* fashion) with those of other groups; still, BIC minimization admits regions from the same group working “in counterphase”, as well as regions from different groups working “in phase”—while F minimization does not.

counterphase” even if belonging to the same group (see, for instance, the patterns highlighted in Figs. 7 and 8), as well as regions working “in phase” even if belonging to different groups.

2. Traditionally or relaxedly balanced?

Motivated by the last observation, we now employ our benchmarks to probe the patterns of structural (im)balance at the mesoscopic level. We also explicitly acknowledge that discussing balance theory for brain networks may seem somewhat inappropriate; nonetheless, this is the only existing framework to interpret our results.

Traditional theory of balance. According to the strong variant of the traditional balance theory (TBT), the optimal partition of a network set of nodes consists of two groups, whose internal (external) connections are positive (negative); in contrast, the weak variant of the TBT allows for any number

of groups. More formally, both versions of the TBT can be probed by finding the partition σ that minimizes the *frustration*, defined as

$$F(\sigma) = L_{\bullet}^{-} + L_{\circ}^{+}, \quad (84)$$

i.e., as the number of negative links within modules (indicated with a filled dot) plus the number of positive links between modules (indicated with an empty dot). Let us also remark that minimizing F leads to results that are very similar to the ones obtained upon maximizing the signed modularity [70].

The naive, F -induced partition in Fig. 10 illustrates a first way the optimization of F works: A large number of positive links leads F to collect them together, being more convenient to accommodate a small number of negative links within a larger, mostly positive, block than letting a large number of positive links be placed between blocks. The heterogeneous, F -induced partition in Fig. 10, instead, illustrates a second way the optimization of F works: A large number of negative

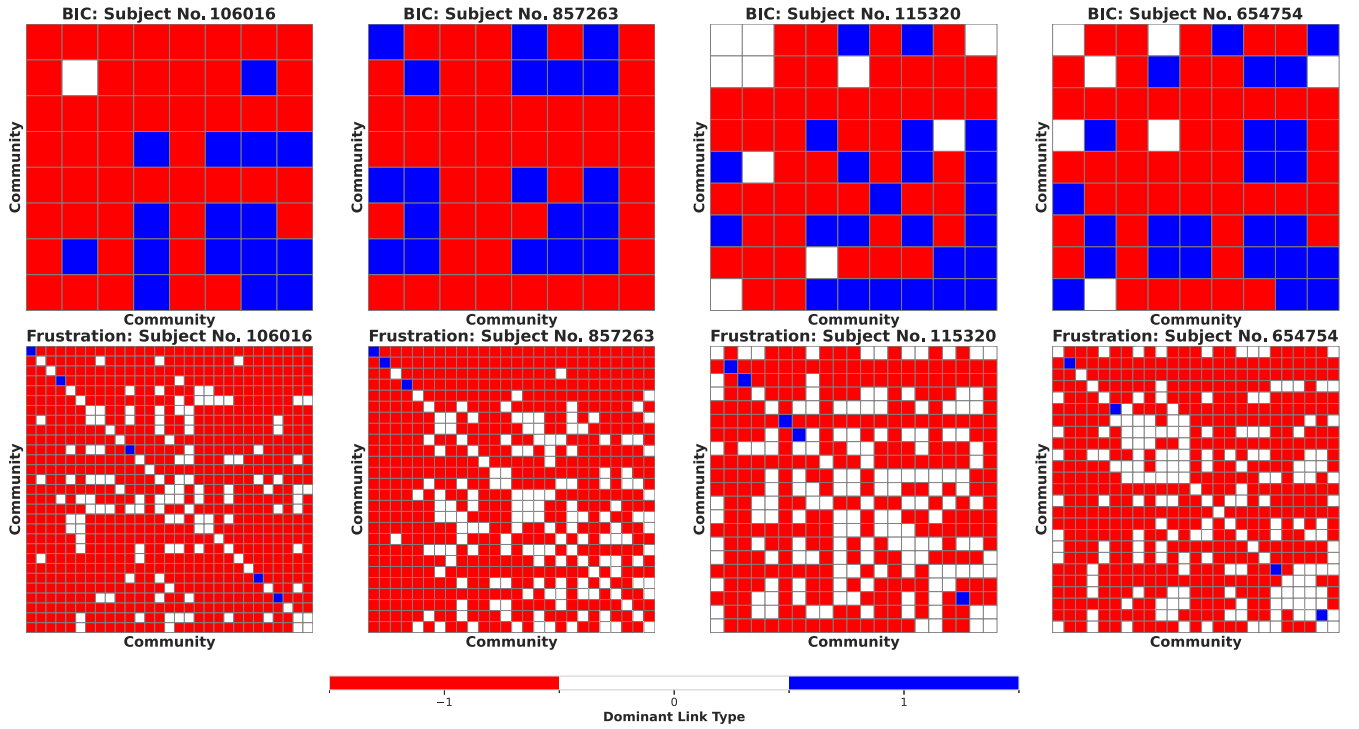


FIG. 11. Heatmaps of the communities partitioning the bSCM-induced projections of four subjects. Matrices are colored according to Eqs. (87) and (88), i.e., blocks are blue (red) if the majority of links populating them is positive (negative); white blocks on the diagonal indicate singletons and white blocks off the diagonal indicate that $p_{rs}^+ = p_{rs}^-$. Top panels: The communities partitioning the bSCM-induced projections of our subjects are detected by minimizing the BIC; all subjects exhibit diagonal blocks with a majority of negative links and off-diagonal blocks with a majority of positive links, thereby satisfying the statistical variant of the relaxed balance theory. Bottom panels: The communities partitioning the bSCM-induced projections of our subjects are detected by minimizing the frustration (F); (left aside the white ones) all subjects admit blocks with a majority of negative links only off the main diagonal and blocks with a majority of positive links only on the main diagonal, hence aligning better with the statistical variant of the traditional balance theory.

links leads F to split them, even at the price of creating many singletons. In any case, the minimization of F further confirms that none of our projections obeys the traditional formulation of the balance theory.

Statistical theory of balance. Such a result motivates us to inspect the theory of balance best accommodating our observations. To answer such a question, let us first notice that the TBT can be formalized upon writing

$$p_{rr}^- = 0, \quad \forall r \quad (85)$$

and

$$p_{rs}^+ = 0, \quad \forall r < s, \quad (86)$$

with p_{rr}^+ indicating the probability that any two nodes belonging to the same block r are connected by a positive link, p_{rs}^+ indicating the probability that any two nodes belonging to the different blocks r and s are connected by a positive link, and analogously for their negative counterparts. As noticed in Ref. [70], however, the positions defining the TBT can be replaced with the milder ones reading

$$\text{sgn}[p_{rr}^+ - p_{rr}^-] = +1, \quad \forall r, \quad (87)$$

which amounts to requiring $p_{rr}^+ > p_{rr}^-$, $\forall r$ and

$$\text{sgn}[p_{rs}^+ - p_{rs}^-] = -1, \quad \forall r < s, \quad (88)$$

which amounts to requiring $p_{rs}^+ < p_{rs}^-$, $\forall r < s$: In words, the deterministic rules first defined by Cartwright, Harary, and Davis are replaced by a set of probabilistic criteria individuating “a tendency” to obey, or not to obey, the TBT. More formally, a configuration satisfying Eqs. (87) and (88) will be claimed to support the *statistical variant* of the TBT—specifically, its strong variant if $k = 2$ and its weak variant if $k > 2$; if, instead, $p_{rr}^+ \leq p_{rr}^-$ for some diagonal blocks or $p_{rs}^+ \geq p_{rs}^-$ for some off-diagonal blocks, it will be claimed to support the *statistical variant of the relaxed balance theory* (RBT).

In order to evaluate the alignment of our projections with either statistical variants, let us take a look at the top panels of Fig. 10, depicting the BIC-induced partitions: Since $p_{rs}^+ \leq p_{rs}^-$ for several, diagonal blocks (i.e., with $r = s$) and $p_{rs}^+ \geq p_{rs}^-$ for several, off-diagonal blocks (i.e., with $r \neq s$), we are led to the conclusion that all projections obey the statistical variant of the RBT [70] as a non-negligible number of negative (positive) links is found within (between)—and not only between (within)—clusters. As Fig. 11 further confirms, no subject obeys the statistical variant of the TBT, regardless of the benchmark employed, when considering the BIC-induced partitions.

Let us, now, take a look at the bottom panels of Fig. 10, depicting the F -induced partitions: Although Eqs. (87) and (88) are not strictly obeyed, our subject, now, admits blocks

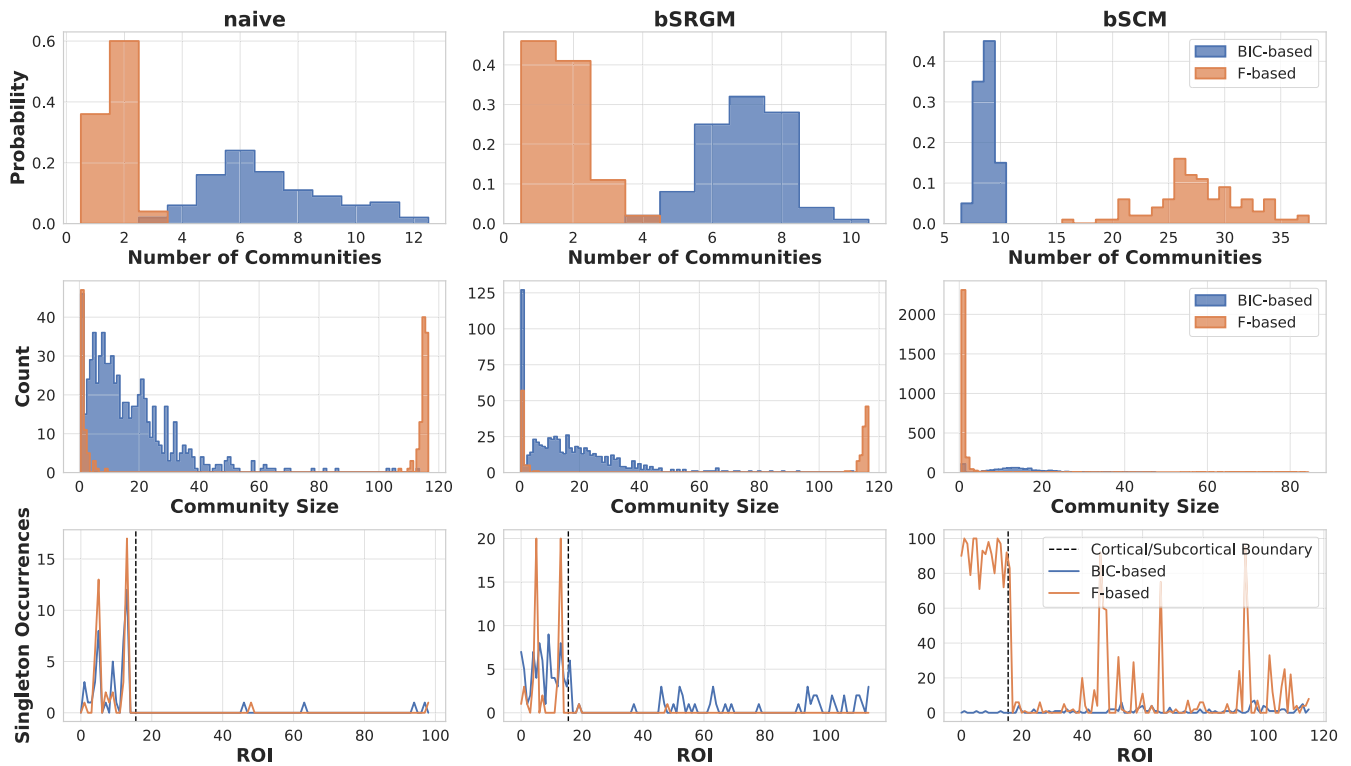


FIG. 12. Basic statistics of the partitions across subjects. Top panels: Histograms of the number of communities detected on the projections returned by each benchmark and optimization criterion. For what concerns naive and bSRGM-induced projections, F avoids creating frustrated patterns by individuating a smaller number of larger communities, while BIC singles out a larger number of smaller communities; for what concerns bSCM-induced projections, instead, F individuates a larger number of smaller communities than BIC. Middle panels: Histograms of the size of communities detected on the projections returned by each benchmark and optimization criterion. These panels refine the picture provided by the previous ones, clarifying that F tends to separate the singletons from the set of remaining nodes; BIC tends to individuate singletons as well, although the size of communities into which the set of remaining nodes is partitioned varies over a broader range. Bottom panels: Number of times each brain region is individuated as a singleton. As evident from the plot, this is almost solely observed for subcortical regions, especially on F -induced partitions; the most “integrated” picture is, instead, returned by the bSCM-induced projections, partitioned by minimizing BIC.

with a majority of negative links only off the main diagonal and blocks with a majority of positive links only on the main diagonal; we are, thus, led to the conclusion that all projections align better with the statistical variant of the TBT. As Fig. 11 further confirms, all subjects align better with the statistical variant of the TBT, regardless of the benchmark employed, when considering the F -induced partitions.

D. Toward a representative brain?

Let us now move from individual analyses to the aggregated one, the aim being to understand the extent to which the previous results can be generalized.

1. Distribution of the number of communities

First, let us plot the distribution of the number of communities per benchmark and optimization criterion. The results are shown in Fig. 12: For what concerns naive and bSRGM-induced projections, the number of communities detected by minimizing the frustration is practically always smaller than the number of communities detected by minimizing BIC—the F -induced mode being 2 and the BIC-induced modes being 6 and 7; moreover, the F -induced distributions are narrower than the BIC-induced ones, a result hinting at a more

pronounced intersubject variability, in the second case. Concerning the bSCM-induced projections, instead, F identifies a larger and more dispersed number of smaller communities than BIC.

2. The role of singletons

The histograms of the size of communities detected on the same projections refine the aforementioned picture, clarifying that F tends to individuate simpler structures than BIC, such as fewer modules accompanied by several singletons detached from the rest of the nodes—Although BIC may individuate singletons as well, the size of communities into which the rest of the nodes is partitioned varies over a broader range.

Interestingly, singletons often coincide with the so-called *subcortical regions*, i.e., regions serving as hubs “integrating” the *cortical regions*, among the other things: More quantitatively, the fraction of subcortical singletons amounts to (1) $\simeq 0.96$ according to F and $\simeq 0.91$ according to BIC, on the naive projections; (2) $\simeq 0.96$ according to F and $\simeq 0.60$ according to BIC, on the bSRGM-induced projections; and (3) $\simeq 0.63$ according to F and $\simeq 0.04$ according to BIC, on the bSCM-induced projections. In words, carrying out a BIC-based community detection on bSCM-induced projections

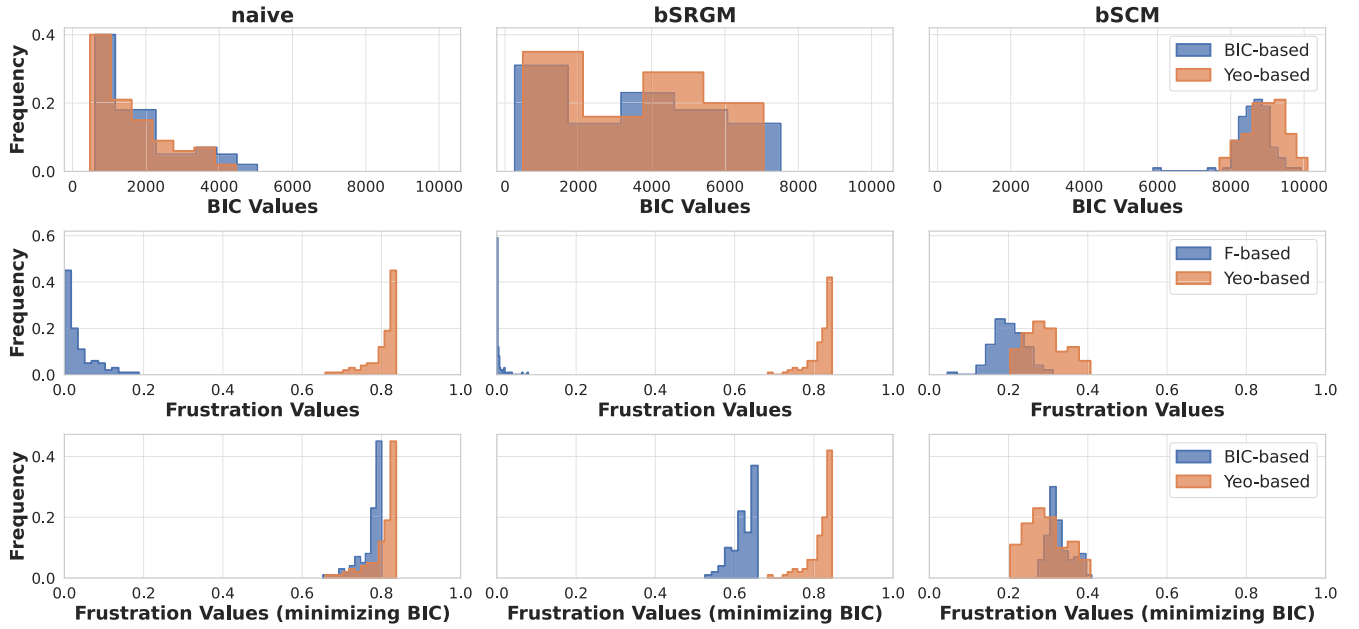


FIG. 13. Distributions of the BIC and F values across the partitions induced by their minimization. Top panels: Histograms of the BIC values characterizing the partitions determined by minimizing it (in blue) and Yeo’s partitions (in orange). As expected, the former are characterized by a slightly smaller value of such a score than Yeo’s partitions. Middle panels: Histograms of the F values characterizing the partitions determined by minimizing it (in blue) and Yeo’s partitions (in orange). In this second case, the former are characterized by a (much) smaller value of such a score than Yeo’s partitions. Bottom panels: Histograms of the F values characterizing the partitions determined by minimizing BIC (in blue) and Yeo’s partitions (in orange). In this third case, both are characterized by a quite large value of F . Bin widths are computed according to the Freedman-Diaconis rule. Taken together, our results indicate that (1) brain networks do not seek to minimize frustration—even when removing the subcortical regions; and (2) BIC minimization embodies a principle returning a structure that captures (at least) part of the features of the Yeo’s task-based one.

rarely leads to isolate subcortical regions, a result indicating that accounting for local properties such as the series- and the time-specific degrees leads to recover projections where all regions are (practically always) integrated.

3. Distribution of BIC and F values

A related analysis concerns the distributions of the BIC and F values across the partitions determined by minimizing each of them. As Fig. 13 confirms, the partitions retrieved by minimizing BIC are characterized by a smaller value of such a score than Yeo’s partitions; besides, both kinds of partitions are characterized by a similar, and quite large, value of F . Inspecting the distribution of the values of F , instead, reveals that the partitions retrieved by minimizing it are much less frustrated than Yeo’s partitions. Taken together, our results suggest that (1) brain networks do not seek to minimize frustration (even when removing subcortical regions—as in this case, to carry out a fair comparison with Yeo’s classification) and (2) BIC minimization is capable of capturing a structure that shares several similarities with the one defined by Yeo’s task-based networks.

4. The barycentric partition

Let us, now, ask ourselves whether a criterion to condense the information embodied by each individual partition can be devised. To this aim, one could try to single out the “most representative” brain by individuating the “most representative” mesoscale structure: Finding an answer to this question

ultimately amounts to finding a method to compare any two partitions, say, $X = \{x_1 \cdots x_n\}$ and $Y = \{y_1 \cdots y_m\}$.

Rand index. One of the simplest metrics is the Rand index (RI), reading

$$RI = \frac{TP + TN}{TP + FP + TN + FN} = \frac{TP + TN}{N(N-1)/2}, \quad (89)$$

and computing the ratio between the sum of *true positives* (TP; the number of nodes that are “together” in both partitions, i.e., in the same subset in X and in the same subset in Y) and *true negatives* (TN; the number of nodes that are “separated” in both partitions, i.e., in different subsets in X and in different subsets in Y) and the total number of pairs of nodes. Naturally, $0 \leq RI \leq 1$.

Other variants are the Wallace index (WI), defined as $WI = TP/(TP + FP)$, and the Jaccard index (JI), defined as $JI = TP/(TP + FP + FN)$: Notice, however, that both WI and JI measure the agreement between two partitions by solely considering the number of TP.

Normalized mutual information. A second metric is the normalized mutual information (NMI), reading

$$NMI(X, Y) = \frac{2MI(X, Y)}{S(X) + S(Y)}, \quad (90)$$

where

$$MI(X, Y) = \sum_{i=1}^N \sum_{j=1}^N h_{ij} \ln \left(\frac{h_{ij}}{f_i g_j} \right), \quad (91)$$

with $h_{ij} = |X_i \cap Y_j|/N$ being the fraction of nodes common to subsets X_i and Y_j , $S(X) = -\sum_{i=1}^N f_i \ln f_i$ being the Shannon entropy induced by the partition X and $f_i = |X_i|/N$ being the fraction of nodes within the cluster X_i , and $S(Y) = -\sum_{j=1}^N g_j \ln g_j$ being the Shannon entropy induced by the partition Y and $g_j = |Y_j|/N$ being the fraction of nodes within the cluster Y_j . Naturally, $0 \leq \text{NMI} \leq 1$.

Rajski distance. A third metric is the *variation of information* (VI), reading

$$\begin{aligned} \text{VI}(X, Y) &= S(X, Y) - \text{MI}(X, Y) \\ &= 2S(X, Y) - S(X) - S(Y) \\ &= S(X) + S(Y) - 2\text{MI}(X, Y) \end{aligned} \quad (92)$$

or, more explicitly,

$$\text{VI}(X, Y) = -\sum_{i=1}^N \sum_{j=1}^N h_{ij} \left[\ln \left(\frac{h_{ij}}{f_i} \right) + \ln \left(\frac{h_{ij}}{g_j} \right) \right]; \quad (93)$$

such an expression can be normalized by dividing it by the joint entropy $S(X, Y) = -\sum_{i=1}^N \sum_{j=1}^N h_{ij} \ln h_{ij}$, i.e.,

$$\text{RD}(X, Y) = \frac{\text{VI}(X, Y)}{S(X, Y)} = 1 - \frac{\text{MI}(X, Y)}{S(X, Y)}. \quad (94)$$

Naturally, $0 \leq \text{RD} \leq 1$.

Finding the barycentre. Any of the metrics above can be employed to assess how different the partitions characterizing the brains of any two subjects, say, u and u' , are. Once such a matrix of distances, say, $\mathbf{D} = \{D_{uu'}\}_{u,u'=1}^U$, with $U = 100$, has been obtained, a second method is needed to condense such a wealth of information. Our proposal is that of calculating the vector $\underline{\mathbf{D}} = \{\underline{D}_u\}$, where

$$\underline{D}_u = \frac{\sum_{u'=1}^U D_{uu'}}{U}, \quad (95)$$

and take the partition lying at distance

$$d = \min \{\underline{D}_u\}_{u=1}^U \quad (96)$$

from the others, i.e., at the minimum, average one. Such a partition will be referred to as the *barycentric partition* or *brain centroid*. Let us explicitly notice that while VI and Rajski distance (RD) are proper distances over the space of possible partitions of a system, hence inducing the positions $D_{uu'}^{\text{VI}} = \text{VI}_{uu'}$ and $D_{uu'}^{\text{RD}} = \text{RD}_{uu'}$, RI and NMI quantify the similarity of any two partitions [112], hence inducing the positions $D_{uu'}^{\text{RI}} = 1 - \text{RI}_{uu'}$ and $D_{uu'}^{\text{NMI}} = 1 - \text{NMI}_{uu'}$.

The observations about the importance of considering filtered projections and the prominent role played by heterogeneous benchmarks motivate us to focus on the partitions obtained by minimizing BIC on the projections induced by the bSCM.

Upon looking at Fig. 14, we realize that (1) RI individuates a barycentric projection with seven predominantly positive communities; (2) NMI and RD individuate the same barycentric projection with five predominantly positive communities; (3) WI and VI individuate a barycentric projection with four predominantly positive communities; and (4) JI individuates a barycentric projection with five major communities defined by predominantly positive connections. The explanation of

such a result may lie in the evidence that indices like WI and JI are solely defined in terms of true positives: As a consequence, identifying a smaller number of larger communities may enhance the detection of such events—although negative connections become underrepresented; indices like RI, on the other hand, account for true negatives as well, whence the larger number of smaller communities—connected by a vast majority of negative links.

Let us, now, compare the community structure of the different brain centroids with the partition defined by Yeo. As Fig. 14 shows, while all metrics tend to individuate communities capturing large portions of certain Yeo's communities, each one induces a different overlap between our structure-based partitions and Yeo's task-based one.

A simpler recipe to identify the representative brain could be that of averaging the individual projections in a benchmark-specific fashion; as the presence of a vast majority of positive links causes most of the negative ones populating the naive and bSRGM-induced projections to be deleted, here we limit ourselves to consider the bSCM-induced projections. As Fig. 15 shows, BIC minimization leads to single out 7 communities, while F minimization leads to single out 17 communities: In the second case, these modules appear as quite uninformative as 15 are very small groups of nodes fragmenting a single Yeo's network in a quite unrealistic fashion, while the first two ones gather all the remaining regions together; in the first case, instead, these modules share similarities with the ones induced by RI and those induced by NMI, and are capable to recover the LIM network in its entirety.

IX. DISCUSSION

The role played by negative connections within and between brain regions represents a topic that has recently gained attention within the neuroscientific literature [34,72,107]; still, they remain difficult to interpret because their estimation depends on preprocessing choices, correlation-based construction pipelines, and thresholding procedures. Here, we propose a statistically validated route to signed functional projections that does not require (constructing and thresholding) a correlation matrix: Instead, signed links are inferred by counting (concordant and discordant) temporal motifs and validating their abundance against two benchmarks, i.e., the bSRGM (which controls for positive and negative link densities) and the bSCM (which additionally preserves local heterogeneities).

Such a validation scheme allows us to address a number of neuroscientific questions while overcoming specific pitfalls characterizing traditional analyses. Crucially, however, our method goes beyond simply confirming known patterns: From a methodological perspective, the entire analysis represents an approach to network construction from time series, circumventing the intermediate, and problematic, step of handling a correlation matrix. Standard correlation-based analyses, in fact, face well-documented challenges: arbitrary threshold selection, sensitivity to GSR, and the difficulty of distinguishing genuine anticorrelations from preprocessing artifacts [97,113,114]. Our validation scheme sidesteps these issues by contrasting the occurrence of (concordant and

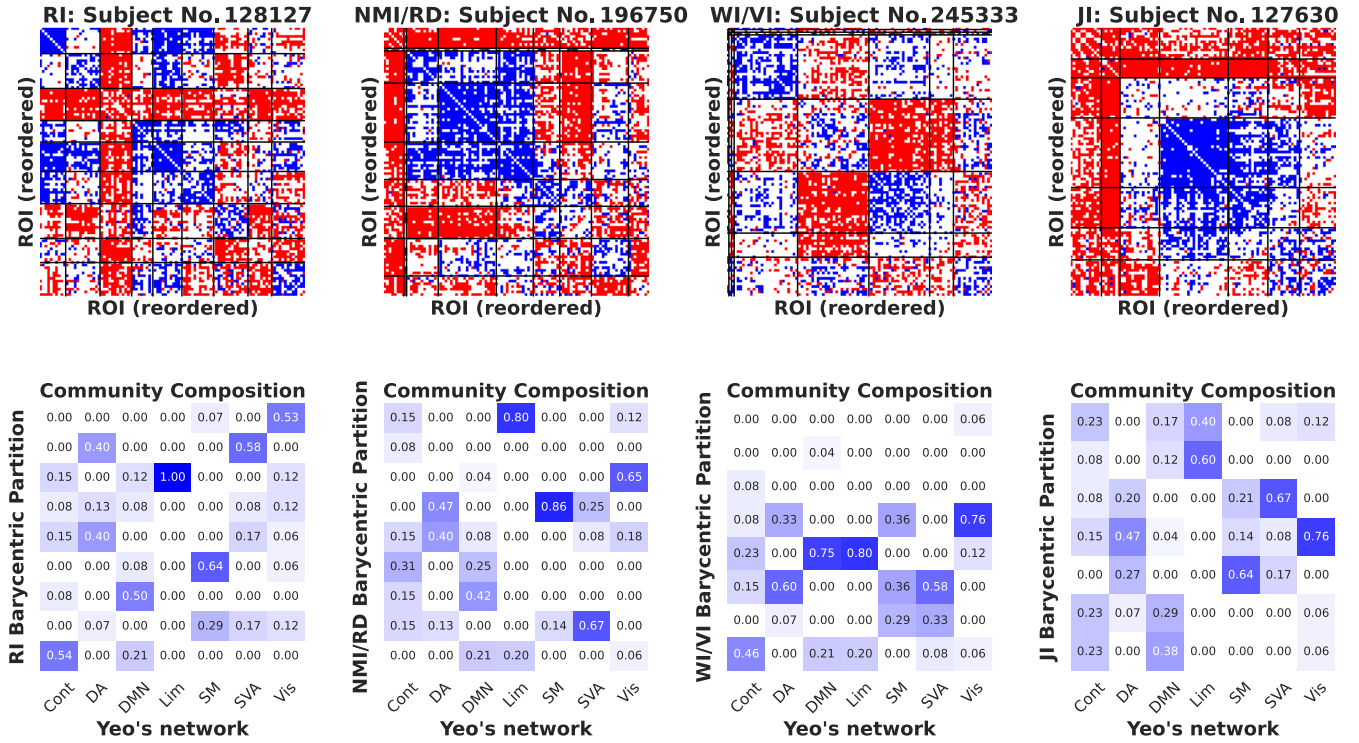


FIG. 14. Brain centroids individuated by the considered distances between partitions. Top panels: Brain centroids individuated by adopting the RI, the NMI, the RD, the WI, the JI, and the VI. Generally speaking, while RI, NMI, and RD individuate projections with a larger number of predominantly positive communities, WI, JI, and VI individuate projections with a smaller number of predominantly positive communities. Such a result may be explained by noticing that indices like WI and JI are solely defined in terms of true positives: As a consequence, identifying a smaller number of larger communities may enhance the detection of such events. Indices like RI, on the other hand, account for true negatives as well, whence the larger number of smaller communities. Bottom panels: Different metrics induce a different overlap between our structure-based partitions and Yeo's task-based one. Overall, the communities identified by WI, JI, VI, and RD are able to capture up to 86% of certain Yeo's networks, while RI recovers the LIM network in its entirety.

discordant) temporal motifs against properly defined benchmarks (bSRGM and bSCM), which, differently from traditional approaches, can account for the heterogeneity of each region. This approach provides several advantages that manifest in our results.

1. Negative links reveal brain (im)balance

While previous studies have noticed the presence of frustrated triads in brain networks, our IBI metrics provide a systematic quantification of this phenomenon: The evidence that the median IBI value is positive, regardless of the benchmark employed, allows us to conclude that brain networks are populated by a non-negligible number of $(+ - +)$ and $(- - -)$ triads.

However, the contrasting IBI values of bSRGM- and bSCM-induced projections illuminate a fundamental aspect of brain organization: The global filter (bSRGM), which levels out differences between ROIs, produces densely positive subgraphs with few negative connections; the local filter (bSCM), on the contrary, reveals a more balanced distribution of positive and negative links—a result demonstrating that not all regions participate equally in concordant and discordant dynamics.

As shown in Appendix E, our framework generalizes the one defined by Pearson's test of hypothesis, whose outcome overlaps to a very large extent with the one of the bSRGM: Left aside the particularly large value of the Jaccard similarity (amounting to 1, in some cases), such a result indicates that Pearson's test of hypothesis is a homogeneous one, i.e., ignores the peculiarities of the two ROIs whose empirical Pearson correlation coefficient is being tested; moreover, it returns projections with a majority of positive signs. From this perspective, the bSCM-induced projection scheme generalizes it by properly accounting for the role played by any two series asynchrony.

Interestingly, the negative connections constituting the frustrated triads are mostly found (1) within groups of SUBC and LIM regions and (2) between groups of nodes constituting the DMN and those constituting task-positive regions: These patterns are consistent with the established architecture of rs-networks [73,115–117], providing strong evidence that our method captures a genuine structure rather than a methodological artifact; similarly, the prominent role of subcortical regions in network frustration aligns with recent observations about their integrative function and contribution to network metastability [72,107]. These consistencies validate our approach: The statistically significant signed links emerging

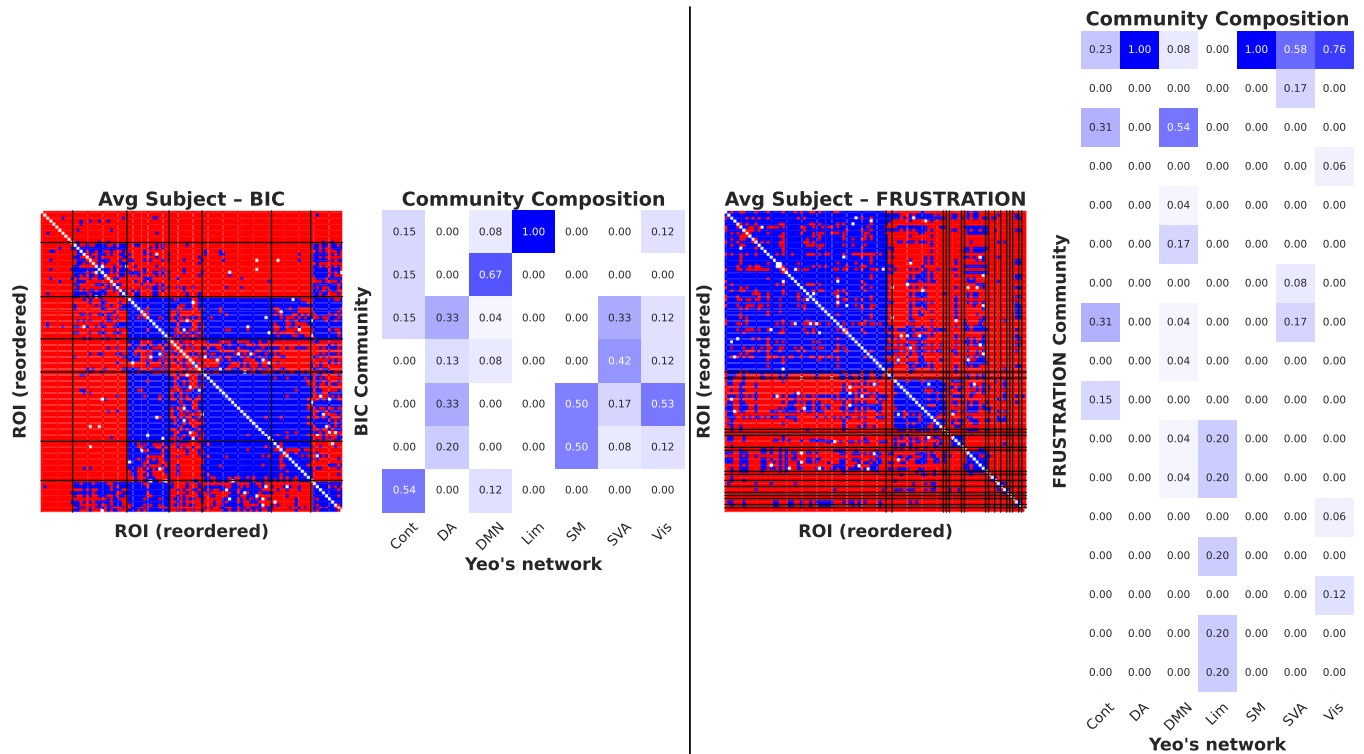


FIG. 15. Average brain, partitioned according to BIC minimization and F minimization. As the presence of a vast majority of positive links causes most of the negative ones populating the naive and bSRGM-induced projections to be deleted, only the bSCM-induced projections have been considered: While the average brain partitioned according to F minimization is characterized by 2 predominantly positive communities that gather the vast majority of Yeo's networks together and 13 predominantly negative modules that fragment the remaining ones, the average brain partitioned according to BIC minimization is characterized by modules sharing similarities with the ones induced by NMI and those induced by RI and NMI, and capable to recover the LIM network in its entirety.

from our benchmarks correspond to (functionally) meaningful neural connections.

2. Mesoscopic organization beyond traditional balance

At the mesoscale, our results indicate that the signed organization of rs-networks is not in accordance with the TBT: If traditional balance were the dominant organizing principle, one would expect predominantly positive links within communities and predominantly negative links between communities; here, instead, the observed partitions obtained through BIC minimization reveal that both negative (positive) connections can populate diagonal (off-diagonal)/within (between) blocks at the mesoscopic level.

While minimizing frustration leads to partitions that make the network as close as possible to a traditionally balanced state, BIC minimization identifies the block structure best describing the observed projection while penalizing model complexity; the two criteria, therefore, answer different questions and their disagreement suggests that the brain mesoscale is not primarily shaped by frustration reduction: Rather, the observed structure appears to preserve a mixture of cooperative and competitive interactions, both within and between functional systems.

Earlier studies reported that empirical brain networks are more balanced than suitably randomized configurations, thus

suggesting that functional brain organization may be biased toward relatively stable signed configurations [34]; more recently, Saberi *et al.* [118] provided dynamic evidence in favor of TBT by showing that balanced triads display longer lifetimes and larger peak energies than imbalanced ones—with a separation of triadic states consistent with the strong formulation of the TBT.

These findings may appear in tension with ours. However, they mainly assess *triadic* balance in correlation-based functional networks and in shorter time windows: Our analysis, instead, focuses on statistically validated signed projections and on their *mesoscopic* block organization. In this setting, the resulting partitions admit negative links within modules and positive links between modules, a result suggesting that the brain may operate in a functionally “excited” state rather than seeking a “ground” state of minimal frustration—The (statistical variant of the) RBT may, thus, represent a better framework than the TBT to interpret our results.

3. Intersubject variability and representative partitions

The analysis of Jaccard similarity to assess intersubject robustness of positive and negative connections reveals an interesting asymmetry on naive and bSRGM-induced projections: Positive links show a larger spatial coherence, while negative links are more variable. This pattern suggests

that positive connections, mirroring a concordant dynamics, follow a relatively stereotyped architecture across individuals, whereas negative connections, mirroring a discordant, or competitive, dynamics, may be more subject specific or state dependent. By accounting for local heterogeneity, the bSCM filter reconciles this asymmetry, a finding that may have implications in fingerprinting analysis and personalized neuroscience and that standard (i.e., correlation-based) analyses, which often threshold away negative connections, lose.

Given the high level of intersubject variability, the “most representative” brain network remains challenging to define, as different distance metrics yield different barycentric partitions: While those emphasizing true positives (WI, JI) favor a smaller number of larger communities, the ones incorporating true negatives (RI) identify a larger number of smaller communities. This sensitivity highlights a fundamental tension in network neuroscience, i.e., whether to prioritize the detection of strong and consistent connections (potentially missing sparse but important links) or to adopt a more fine-grained view (potentially fragmenting coherent functional systems).

Our averaged projections suggest a practical compromise: BIC minimization on bSCM-induced projections identifies communities that capture up to 80% (in one case, 100%) of certain Yeo’s networks, while maintaining an interpretable structure that does not discard negative interactions: This approach may prove valuable for defining reference networks in clinical applications or for generating prior hypotheses in connectivity studies.

4. Robustness and technical considerations

An important concern with any fMRI-based analysis is the potential confound of autocorrelations induced by the hemodynamic response. We have verified that our results are robust against standard cleaning procedures: Although a certain degree of autocorrelation is present, it primarily affects signals at small time lags—typically on the same order of magnitude as the hemodynamic response itself; running an autoregressive model of order up to 7 is sufficient to remove these effects without substantially altering the validated network structure.

The choice of temporal resolution and window length represents another important consideration: Our analysis used the full rs-scan duration but the framework naturally extends to sliding-window approaches for capturing dynamic FC. The key requirement is that the temporal window be sufficiently long to accumulate enough concordant and discordant motifs for statistical testing.

5. Limitations and caveats

While our method offers several advantages, certain limitations warrant discussion. First, the binary nature of our current framework, classifying temporal snapshots as simply concordant or discordant, discards information about the magnitude of co-fluctuations, which typically drives FC and encodes subject-specific information [119]: Future extensions incorporating weighted edges (as mentioned in Sec. X) would address this limitation; second, our analysis focuses on triadic motifs built upon pairwise relationships, while higher-order interactions beyond triads may play important roles in shaping

and characterizing (dis)concordant neural dynamics; third, the choice between bSRGM and bSCM represents a fundamental modeling decision: while the bSCM preserves heterogeneity, the bSRGM provides a simpler null model that may be preferable when strong prior hypotheses about degree distributions are lacking; and fourth, our benchmarks treat time steps as exchangeable—hence, the framework is insensitive to the temporal ordering of the original data: the resulting projection is identical for any permutation of time points that preserves the simultaneity structure across brain regions—although this renders the framework unsuitable for causal analyses, it is appropriate when focusing on (simultaneous) co-occurrences of ROIs; finally, we considered a single dataset (the HCP-Young Adults [91]) with a single parcellation (Schaefer + Tian [95,96]): additional results on different datasets and parcellations are needed to establish their robustness.

The interpretation of “frustration” in neural contexts also deserves caution: Unlike in social networks, where frustration may represent genuine conflict, in neural networks it may reflect more subtle physiological or functional mechanisms; the positive IBI values we observe may, thus, represent a more subtle signature of a complex, multiscale system that is maintained in an “excited” state through a balance of cooperative and competitive interactions.

6. Physiological mechanisms

For the physiological mechanisms underlying the emergence of negative connections, including potential contributions from neural inhibition, neurovascular coupling, and network-level oscillatory dynamics, we redirect the interested reader to more technical contributions [120]; the validation framework presented here is agnostic to these mechanisms, focusing on the statistical significance and topological organization of the observed patterns.

X. CONCLUSIONS

A critical, yet often overlooked, aspect of functional brain data is the role played by negative links: Frequently ignored due to the interpretative challenges brought along, emerging evidence indicates negative links to be fundamental in shaping resting-state networks—especially the properties affecting self-stabilization.

As excluding negative links, thus, means underestimating the brain’s ability to function properly, we have proposed a validation technique allowing statistically significant signed links to emerge out of the noise accompanying raw data; two, different benchmarks have been employed, filtering multivariate time series to a different extent and letting diverse patterns emerge out of the wealth of empirical observations: While the bSRGM levels out the differences between time series, the bSCM preserves their heterogeneity, thus inducing sparser projections that retain more information about the original structure.

In particular, our analysis points out that brain networks are, indeed, frustrated, the patterns induced by signed connections overlapping (in some cases, to a quite large extent) with Yeo’s task-based partitions and returning the picture of a system aligning with the statistical variant of the relaxed

balance theory rather than seeking to minimize frustration. Negative links are not evenly distributed, being concentrated between specific ROIs, especially those belonging to Yeo's SUBC and LIM networks.

A future research direction is that of employing our algorithm to obtain statistically validated *binary* and *weighted* projections of multivariate *weighted* time series. Such an extension would preserve the interpretability of the present framework, while enabling a more nuanced characterization of cooperative and competitive dynamics in functional brain networks.

ACKNOWLEDGMENTS

This work is supported by the PNRR-M4C2-Investimento 1.3, Partenariato Esteso PE00000013-“FAIR-Future Artificial Intelligence Research”-Spoke 1 “Human-centered AI”, funded by the European Commission under the NextGeneration EU programme. This work has also been supported by the European Union under the scheme HORIZON-INFRA-2021-DEV-02-01 – Preparatory phase of new ESFRI research infrastructure projects, Grant Agreement No. 101079043, “SoBigData RI PPP: SoBigData RI Preparatory Phase Project” by “Reconstruction, Resilience and Recovery of Socio-Economic Networks” RECON-NET EP_FAIR_005PE0000013 “FAIR,” Funded by the European Union NextGeneration EU, PNRR Mission 4 Component 2, Investment 3.1; and by the European Community programme under the funding schemes: ERC-2018-ADG G.A. 834756 “XAI: Science and technology for the explanation of AI decision making”.

Study conception and design: M.I.-B., M.D.V., T.G., and T.S. Literature review: E.A., M.D.V., and S.T. Data collection: E.A. Analysis and interpretation of results: E.A., M.D.V., and T.S. Draft manuscript preparation: E.A., M.D.V., and T.S. Draft manuscript revision: E.A., M.I.-B., M.D.V., S.T., T.G., and T.S.

The authors declare no competing interests.

DATA AVAILABILITY

Data concerning rs-fMRI from the HCP can be found in Ref. [121] for a description, see Ref. [122]. The Python package named SIBERIA (*Signed BEnchmarks foR tIme series Analysis*), implementing our benchmarks for the analysis of multivariate time series, is available on PyPI and in Ref. [123].

APPENDIX A: PROBABILISTIC MODELS FOR BINARY UNDIRECTED BIPARTITE SIGNED NETWORKS

The generalization of the ERG formalism for the analysis of binary undirected bipartite signed graphs rests upon the constrained maximization of Shannon entropy, i.e.,

$$\mathcal{L} = S[P] - \sum_{i=0}^M \theta_i [P(\mathbf{B}) C_i(\mathbf{B}) - \langle C_i \rangle], \quad (\text{A1})$$

where

$$S = - \sum_{\mathbf{B} \in \mathbb{B}} P(\mathbf{B}) \ln P(\mathbf{B}), \quad (\text{A2})$$

$C_0 = \langle C_0 \rangle = 1$ sums up the normalization condition, and the remaining $M - 1$ constraints represent proper, topological properties. Such an optimization procedure defines the expression

$$P(\mathbf{B}) = \frac{e^{-H(\mathbf{B})}}{Z} = \frac{e^{-H(\mathbf{B})}}{\sum_{\mathbf{B} \in \mathbb{B}} e^{-H(\mathbf{B})}} = \frac{e^{-\sum_{i=1}^M \theta_i C_i(\mathbf{B})}}{\sum_{\mathbf{B} \in \mathbb{B}} e^{-\sum_{i=1}^M \theta_i C_i(\mathbf{B})}} \quad (\text{A3})$$

that can be made explicit only after having chosen a specific set of constraints.

1. Signed random graph model

Let us consider the properties $L^+(\mathbf{B})$ and $L^-(\mathbf{B})$. The Hamiltonian describing such a problem reads

$$H(\mathbf{B}) = \alpha B^+(\mathbf{B}) + \gamma B^-(\mathbf{B}) = \sum_{i=1}^N \sum_{t=1}^T (\alpha b_{it}^+ + \gamma b_{it}^-) \quad (\text{A4})$$

and induces a partition function reading

$$\begin{aligned} Z &= \sum_{\mathbf{B} \in \mathbb{B}} e^{-H(\mathbf{B})} \\ &= \sum_{\mathbf{B} \in \mathbb{B}} e^{-\alpha B^+(\mathbf{B}) - \gamma B^-(\mathbf{B})} \\ &= \sum_{\mathbf{B} \in \mathbb{B}} e^{-\sum_{i=1}^N \sum_{t=1}^T (\alpha b_{it}^+ + \gamma b_{it}^-)} \\ &= \sum_{\mathbf{B} \in \mathbb{B}} \prod_{i=1}^N \prod_{t=1}^T e^{-(\alpha b_{it}^+ + \gamma b_{it}^-)} \\ &= \prod_{i=1}^N \prod_{t=1}^T \left[\sum_{b_{it} = -1, 1} e^{-(\alpha b_{it}^+ + \gamma b_{it}^-)} \right] \\ &= \prod_{i=1}^N \prod_{t=1}^T (e^{-\alpha} + e^{-\gamma}) \\ &= (e^{-\alpha} + e^{-\gamma})^{N \times T}. \end{aligned} \quad (\text{A5})$$

The expression above leads us to find

$$\begin{aligned} P_{\text{bSRGM}}(\mathbf{B}) &= \frac{e^{-\alpha B^+(\mathbf{B}) - \gamma B^-(\mathbf{B})}}{(e^{-\alpha} + e^{-\gamma})^{N \times T}} \\ &= \frac{x^{B^+(\mathbf{B})} y^{B^-(\mathbf{B})}}{(x + y)^{N \times T}} \\ &= (p^-)^{B^-(\mathbf{B})} (p^+)^{B^+(\mathbf{B})} \end{aligned} \quad (\text{A6})$$

having posed $p^- = \frac{e^{-\gamma}}{e^{-\alpha} + e^{-\gamma}} = \frac{y}{x+y}$ and $p^+ = \frac{e^{-\alpha}}{e^{-\alpha} + e^{-\gamma}} = \frac{x}{x+y}$, where p^+ is the probability that a generic series assumes a positive value at time t and p^- is the probability that a generic series assumes a negative value at time t .

In order to determine the parameters that define the bSRGM, let us maximize the likelihood function

$$\begin{aligned}\mathcal{L}_{\text{bSRGM}}(x, y) &= \ln P_{\text{bSRGM}}(\mathbf{B}^* | x, y) \\ &= B^+(\mathbf{B}^*) \ln x + B^-(\mathbf{B}^*) \ln y \\ &\quad - (N \times T) \ln(x + y),\end{aligned}\quad (\text{A7})$$

with respect to x and y . Upon doing so, we obtain the pair of equations

$$\frac{\partial \mathcal{L}_{\text{bSRGM}}(x, y)}{\partial x} = \frac{B^+(\mathbf{B}^*)}{x} - \frac{N \times T}{x + y}, \quad (\text{A8})$$

$$\frac{\partial \mathcal{L}_{\text{bSRGM}}(x, y)}{\partial y} = \frac{B^-(\mathbf{B}^*)}{y} - \frac{N \times T}{x + y}. \quad (\text{A9})$$

Equating them to zero leads us to find

$$p^+ = B^+(\mathbf{B}^*) / (N \times T), \quad (\text{A10})$$

$$p^- = B^-(\mathbf{B}^*) / (N \times T). \quad (\text{A11})$$

2. Signed configuration model

Let us consider the properties $\{k_i^+(\mathbf{B})\}_{i=1}^N$, $\{k_i^-(\mathbf{B})\}_{i=1}^N$, $\{\kappa_t^+(\mathbf{B})\}_{t=1}^T$, and $\{\kappa_t^-(\mathbf{B})\}_{t=1}^T$. The Hamiltonian describing such a problem reads

$$H(\mathbf{B}) = \sum_{i=1}^N [\alpha_i k_i^+(\mathbf{B}) + \gamma_i k_i^-(\mathbf{B})] + \sum_{t=1}^T [\delta_t \kappa_t^+(\mathbf{B}) + \eta_t \kappa_t^-(\mathbf{B})] \quad (\text{A12})$$

and induces a partition function reading

$$\begin{aligned}Z &= \sum_{\mathbf{B} \in \mathbb{B}} e^{-H(\mathbf{B})} \\ &= \sum_{\mathbf{B} \in \mathbb{B}} e^{-\sum_{i=1}^N [\alpha_i k_i^+(\mathbf{B}) + \gamma_i k_i^-(\mathbf{B})] - \sum_{t=1}^T [\delta_t \kappa_t^+(\mathbf{B}) + \eta_t \kappa_t^-(\mathbf{B})]} \\ &= \sum_{\mathbf{B} \in \mathbb{B}} e^{-\sum_{i=1}^N \sum_{t=1}^T [(\alpha_i + \delta_t) b_{it}^+ + (\gamma_i + \eta_t) b_{it}^-]} \\ &= \sum_{\mathbf{B} \in \mathbb{B}} \prod_{i=1}^N \prod_{t=1}^T e^{-(\alpha_i + \delta_t) b_{it}^+ - (\gamma_i + \eta_t) b_{it}^-} \\ &= \prod_{i=1}^N \prod_{t=1}^T \left[\sum_{b_{it} = -1, 1} e^{-(\alpha_i + \delta_t) b_{it}^+ - (\gamma_i + \eta_t) b_{it}^-} \right] \\ &= \prod_{i=1}^N \prod_{t=1}^T [e^{-(\alpha_i + \delta_t)} + e^{-(\gamma_i + \eta_t)}].\end{aligned}\quad (\text{A13})$$

The expression above leads us to find

$$\begin{aligned}P_{\text{bSCM}}(\mathbf{B}) &= \frac{e^{-\sum_{i=1}^N [\alpha_i k_i^+(\mathbf{B}) + \gamma_i k_i^-(\mathbf{B})] - \sum_{t=1}^T [\delta_t \kappa_t^+(\mathbf{B}) + \eta_t \kappa_t^-(\mathbf{B})]}}{\prod_{i=1}^N \prod_{t=1}^T [e^{-(\alpha_i + \delta_t)} + e^{-(\gamma_i + \eta_t)}]} \\ &= \prod_{i=1}^N \prod_{t=1}^T \frac{x_i^{k_i^+(\mathbf{B})} y_i^{k_i^-(\mathbf{B})} z_t^{\kappa_t^+(\mathbf{B})} w_t^{\kappa_t^-(\mathbf{B})}}{x_i z_t + y_i w_t} \\ &= \prod_{i=1}^N \prod_{t=1}^T (p_{it}^+)^{b_{it}^+} (p_{it}^-)^{b_{it}^-}\end{aligned}\quad (\text{A14})$$

having posed $p_{it}^- = \frac{e^{-(\gamma_i + \eta_t)}}{e^{-(\alpha_i + \delta_t)} + e^{-(\gamma_i + \eta_t)}} = \frac{y_i w_t}{x_i z_t + y_i w_t}$ and $p_{it}^+ = \frac{e^{-(\alpha_i + \delta_t)}}{e^{-(\alpha_i + \delta_t)} + e^{-(\gamma_i + \eta_t)}} = \frac{x_i z_t}{x_i z_t + y_i w_t}$, where p_{it}^+ is the probability that series i assumes a positive value at time t and p_{it}^- is the probability that series i assumes a negative value at time t .

In order to determine the parameters that define the bSCM, let us maximize the likelihood function

$$\begin{aligned}\mathcal{L}_{\text{bSCM}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}) &= \ln P_{\text{bSCM}}(\mathbf{B}^* | \mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w}) \\ &= \sum_{i=1}^N k_i^+(\mathbf{B}^*) \ln x_i + \sum_{i=1}^N k_i^-(\mathbf{B}^*) \ln y_i \\ &\quad + \sum_{t=1}^T \kappa_t^+(\mathbf{B}^*) \ln z_t + \sum_{t=1}^T \kappa_t^-(\mathbf{B}^*) \ln w_t \\ &\quad - \sum_{i=1}^N \sum_{t=1}^T \ln(x_i z_t + y_i w_t),\end{aligned}\quad (\text{A15})$$

with respect to x_i , y_i , z_t , and w_t , $\forall i, t$. Upon doing so, we obtain the system of equations

$$\frac{\partial \mathcal{L}_{\text{bSCM}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w})}{\partial x_i} = \frac{k_i^+(\mathbf{B}^*)}{x_i} - \sum_{t=1}^T \frac{z_t}{x_i z_t + y_i w_t}, \quad \forall i, \quad (\text{A16})$$

$$\frac{\partial \mathcal{L}_{\text{bSCM}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w})}{\partial y_i} = \frac{k_i^-(\mathbf{B}^*)}{y_i} - \sum_{t=1}^T \frac{w_t}{x_i z_t + y_i w_t}, \quad \forall i, \quad (\text{A17})$$

$$\frac{\partial \mathcal{L}_{\text{bSCM}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w})}{\partial z_t} = \frac{\kappa_t^+(\mathbf{B}^*)}{z_t} - \sum_{i=1}^N \frac{x_i}{x_i z_t + y_i w_t}, \quad \forall t, \quad (\text{A18})$$

$$\frac{\partial \mathcal{L}_{\text{bSCM}}(\mathbf{x}, \mathbf{y}, \mathbf{z}, \mathbf{w})}{\partial w_t} = \frac{\kappa_t^-(\mathbf{B}^*)}{w_t} - \sum_{i=1}^N \frac{y_i}{x_i z_t + y_i w_t}, \quad \forall t. \quad (\text{A19})$$

Equating them to zero leads us to find

$$k_i^+(\mathbf{B}^*) = \sum_{t=1}^T \frac{x_i z_t}{x_i z_t + y_i w_t} = \sum_{t=1}^T p_{it}^+ = \langle k_i^+ \rangle, \quad \forall i, \quad (\text{A20})$$

$$k_i^-(\mathbf{B}^*) = \sum_{t=1}^T \frac{y_i w_t}{x_i z_t + y_i w_t} = \sum_{t=1}^T p_{it}^- = \langle k_i^- \rangle, \quad \forall i, \quad (\text{A21})$$

$$\kappa_t^+(\mathbf{B}^*) = \sum_{i=1}^N \frac{x_i z_t}{x_i z_t + y_i w_t} = \sum_{i=1}^N p_{it}^+ = \langle \kappa_t^+ \rangle, \quad \forall t, \quad (\text{A22})$$

$$\kappa_t^-(\mathbf{B}^*) = \sum_{i=1}^N \frac{y_i w_t}{x_i z_t + y_i w_t} = \sum_{i=1}^N p_{it}^- = \langle \kappa_t^- \rangle, \quad \forall t. \quad (\text{A23})$$

The system above can be solved only numerically (see also Appendix B).

APPENDIX B: NUMERICAL OPTIMIZATION OF LIKELIHOOD FUNCTIONS

In order to numerically solve the system of equations defining the bSCM, we can follow the guidelines provided in Ref. [90]: More specifically, we will adapt the iterative recipe provided there to our setting. First, let us notice that, upon

TABLE I. Performance of our fixed-point algorithm to solve the system of equations defining the bSCM on five illustrative time series coming from the HCP-Young Adult dataset (each pedex refers to a subject): N is the number of brain regions, T is the duration of the related time series, B^+ is the overall number of positive values, B^- is the overall number of negative values, $c^+ = B^+/(N \times T)$ is the density of positive values, and $c^- = B^-/(N \times T)$ is the density of negative values. Our fixed-point algorithm leads to MAE values around 10^{-3} and MRE values on the order of 10^{-6} in less than 40 s.

	Basic statistics of time series						bSCM		
	N	T	B^+	B^-	c^+	c^-	MAE	MRE	Time (s)
HCP ₁₀₀₄₀₈	116	2.400	137.804	140.596	≈ 0.495	≈ 0.505	$\approx 1.30 \times 10^{-3}$	$\approx 1.15 \times 10^{-6}$	≈ 37.8
HCP ₁₀₆₀₁₆	116	2.400	142.407	135.993	≈ 0.512	≈ 0.488	$\approx 1.10 \times 10^{-3}$	$\approx 9.94 \times 10^{-7}$	≈ 28.8
HCP ₁₂₅₅₂₅	116	2.400	139.054	139.346	≈ 0.499	≈ 0.501	$\approx 1.26 \times 10^{-3}$	$\approx 1.17 \times 10^{-6}$	≈ 37.8
HCP ₁₃₀₃₁₆	116	2.400	137.196	141.204	≈ 0.493	≈ 0.507	$\approx 1.97 \times 10^{-3}$	$\approx 1.89 \times 10^{-6}$	≈ 37.7
HCP ₈₅₇₂₆₃	116	2.400	136.376	142.024	≈ 0.490	≈ 0.510	$\approx 1.10 \times 10^{-3}$	$\approx 9.90 \times 10^{-7}$	≈ 33.3

posing $x_i = e^{-\alpha_i}$, $y_i = e^{-\gamma_i}$, $\forall i$ and $z_t = e^{-\delta_t}$, $w_t = e^{-\eta_t}$, $\forall t$, the system of equations defining the bSCM can be rewritten as

$$x_i = \frac{k_i^+(\mathbf{B}^*)}{\sum_{t=1}^T \frac{z_t}{x_i z_t + y_i w_t}} \Rightarrow x_i^{(n)} = \frac{k_i^+(\mathbf{B}^*)}{\sum_{t=1}^T \frac{z_t^{(n-1)}}{x_i^{(n-1)} z_t^{(n-1)} + y_i^{(n-1)} w_t^{(n-1)}}}, \quad \forall i, \quad (\text{B1})$$

$$y_i = \frac{k_i^-(\mathbf{B}^*)}{\sum_{t=1}^T \frac{w_t}{x_i z_t + y_i w_t}} \Rightarrow y_i^{(n)} = \frac{k_i^-(\mathbf{B}^*)}{\sum_{t=1}^T \frac{w_t^{(n-1)}}{x_i^{(n-1)} z_t^{(n-1)} + y_i^{(n-1)} w_t^{(n-1)}}}, \quad \forall i, \quad (\text{B2})$$

$$z_t = \frac{\kappa_t^+(\mathbf{B}^*)}{\sum_{i=1}^N \frac{x_i}{x_i z_t + y_i w_t}} \Rightarrow z_t^{(n)} = \frac{\kappa_t^+(\mathbf{B}^*)}{\sum_{i=1}^N \frac{x_i^{(n-1)}}{x_i^{(n-1)} z_t^{(n-1)} + y_i^{(n-1)} w_t^{(n-1)}}}, \quad \forall t, \quad (\text{B3})$$

$$w_t = \frac{\kappa_t^-(\mathbf{B}^*)}{\sum_{i=1}^N \frac{y_i}{x_i z_t + y_i w_t}} \Rightarrow w_t^{(n)} = \frac{\kappa_t^-(\mathbf{B}^*)}{\sum_{i=1}^N \frac{y_i^{(n-1)}}{x_i^{(n-1)} z_t^{(n-1)} + y_i^{(n-1)} w_t^{(n-1)}}}, \quad \forall t. \quad (\text{B4})$$

In order for our iterative recipe to converge, an appropriate vector of initial conditions needs to be chosen; here, we have opted for the following one: $x_i = k_i^+(\mathbf{B}^*)/\sqrt{L^+(\mathbf{B}^*)}$, $\forall i$, $y_i = k_i^-(\mathbf{B}^*)/\sqrt{L^-(\mathbf{B}^*)}$, $\forall i$, $z_t = \kappa_t^+(\mathbf{B}^*)/\sqrt{L^+(\mathbf{B}^*)}$, $\forall t$, and $w_t = \kappa_t^-(\mathbf{B}^*)/\sqrt{L^-(\mathbf{B}^*)}$, $\forall t$.

Moreover, we have adopted two different stopping criteria: The first one is a condition on the Euclidean norm of the vector of differences between the values of the parameters at subsequent iterations, i.e., $\|\Delta\vec{\theta}\|_2 = \sqrt{\sum_{i=1}^N (\Delta\theta_i)^2} \leq 10^{-8}$; and the second one is a condition on the maximum number of iterations of our iterative algorithm, set to 10^3 .

The accuracy of our method in estimating the constraints has been evaluated by computing the *maximum absolute error*

(MAE), defined as

$$\text{bMAE} = \max_{i,t} \{ |k_i^+(\mathbf{B}^*) - \langle k_i^+ \rangle|, |k_i^-(\mathbf{B}^*) - \langle k_i^- \rangle|, |\kappa_t^+(\mathbf{B}^*) - \langle \kappa_t^+ \rangle|, |\kappa_t^-(\mathbf{B}^*) - \langle \kappa_t^- \rangle| \} \quad (\text{B5})$$

(i.e., as the infinite norm of the difference between the vector of the empirical values of the constraints and the vector of their expected values) and the *maximum relative error* (MRE), defined as

$$\text{bMRE} = \max_{i,t} \left\{ \frac{|k_i^+(\mathbf{B}^*) - \langle k_i^+ \rangle|}{k_i^+(\mathbf{B}^*)}, \frac{|k_i^-(\mathbf{B}^*) - \langle k_i^- \rangle|}{k_i^-(\mathbf{B}^*)}, \frac{|\kappa_t^+(\mathbf{B}^*) - \langle \kappa_t^+ \rangle|}{\kappa_t^+(\mathbf{B}^*)}, \frac{|\kappa_t^-(\mathbf{B}^*) - \langle \kappa_t^- \rangle|}{\kappa_t^-(\mathbf{B}^*)} \right\} \quad (\text{B6})$$

(i.e., as the infinite norm of the relative difference between the vector of the empirical values of the constraints and the vector of their expected values).

Table I, providing information on the basic statistics of five illustrative time series coming from the HCP-Young Adult dataset, shows the time employed by our algorithm to converge as well as its accuracy in reproducing the constraints defining the bSCM: Our method is, overall, fast and accurate, as the absolute error is steadily around 10^{-3} and the relative error is steadily around 10^{-6} , the time employed to achieve such an accuracy being always less than a minute.

APPENDIX C: TEMPORAL AUTOCORRELATION AND PREWHITENING OF BOLD TIME SERIES

A documented property of rs-fMRI signals concerns the presence of a non-negligible amount of temporal autocorrelations. Such a presence can be explicitly inspected by computing the autocorrelation function of the time series X_{it} :

$$r_{ii}^\tau = \frac{\sum_{t=1}^T (X_{it} - \bar{X}_i) \times (X_{it+\tau} - \bar{X}_i)}{\sum_{t=1}^T (X_{it} - \bar{X}_i)^2}. \quad (\text{C1})$$

As shown in Fig. 16, the selected series exhibits a large autocorrelation at small lags, decaying to zero over a timescale that is consistent with that of the hemodynamic response [98–100]. The prewhitening procedure described below is intended to remove it from each time series X_{it} .

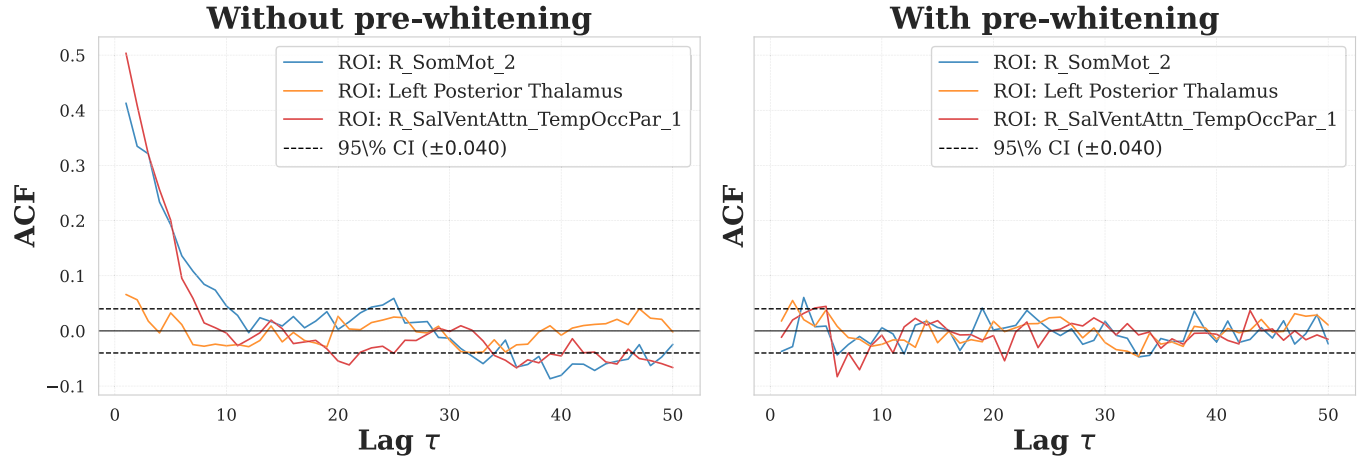


FIG. 16. Autocorrelation function of subject No. 106016, before and after prewhitening. Left panel: Autocorrelation function r_{ii}^T for three ROIs of subject No. 106016, as a function of lag τ ; without prewhitening, two of the three series exhibit positive autocorrelations at short lags—a signature of the hemodynamic response. Right panel: Autocorrelation function r_{ii}^T for the prewhitened binarized series; the autocorrelation function, now, lies within the 95% confidence interval induced by white noise and individuated by the values $\pm 1.96/\sqrt{T}$.

1. Bartlett's white noise test

For a given order p_i , Bartlett's white noise test [102–104] prescribes to compute the *periodogram* of residuals, i.e.,

$$I_i(\omega_k) = \left| \sum_{t=p_i+1}^T \varepsilon_{it} e^{-2\pi i \omega_k t} \right|^2, \quad \omega_k = \frac{k}{T - p_i}, \quad k = 1, \dots, m, \quad (C2)$$

where $m = \lfloor (T - p)/2 \rfloor$. Under the null hypothesis of *white noise*, i.e., $\varepsilon_{it} \sim \mathcal{N}(0, \sigma^2)$, $\forall i, t$, one finds that

$$\begin{aligned} I_i(\omega_k) &= \left(\sum_{t=p_i+1}^T \varepsilon_{it} e^{-2\pi i \omega_k t} \right) \left(\sum_{s=p_i+1}^T \varepsilon_{is} e^{2\pi i \omega_k s} \right) \\ &= \sum_{t=p_i+1}^T \sum_{s=p_i+1}^T \varepsilon_{it} \varepsilon_{is} e^{-2\pi i \omega_k (t-s)}, \end{aligned} \quad (C3)$$

an expression whose expected value reads $(T - p_i)\sigma^2$; as a consequence, the expected value of the normalized cumulative periodogram

$$F_i(\omega_k) = \frac{\sum_{j=1}^k I_i(\omega_j)}{\sum_{j=1}^m I_i(\omega_j)}, \quad k = 1, \dots, m \quad (C4)$$

reads k/m , i.e., grows linearly with k : In other words, the normalized periodogram obeys a uniform distribution on $[0, 1]$ [103]. Such a compatibility—any deviation from which would indicate residual autocorrelation—is assessed via a KS test [104] at the confidence level of 5% and the optimal order is selected as the minimum for which the KS test fails to reject the null hypothesis of white noise—to this aim, the upper bound $P = 30$ has been prespecified.

2. Correcting for multiple hypothesis testing

Since Bartlett's white noise test is carried out P times per series, a procedure accounting for the fact that multiple hypotheses are tested at the same time is needed. Here, we adopt the Benjamini-Yekutieli procedure [124], which controls the false discovery rate under arbitrary dependence of p -values. Operatively, the Pp -values are sorted in increasing order and the selected one is the minimum satisfying the condition

$$p\text{-value}_j \geq \frac{j\alpha}{P \times c(P)}, \quad (C5)$$

with $j = 1, \dots, P$ indexing the p -values associated with each ROI, $c(P) \simeq \ln P$ being the Euler-Mascheroni constant, $\alpha = 0.05$, and $P = 30$ —a choice that has been further validated *a posteriori*: Across all ROIs of our subjects, no value $\hat{p}_i = P$ has been observed, a result confirming that the results are independent of such a specific value.

Figure 17 summarizes the outcome of our prewhitening procedure: The distribution of $\hat{p}_i$ values, pooled across all ROIs of our subjects, is concentrated in the range $[2, 3]$, with a tail extending to, at most, $\hat{p}_i = 7$ —a result confirming that a low-order autoregressive model is powerful enough to whiten our preprocessed BOLD signals, consistently with previous findings [100,125].

APPENDIX D: DEGREE DISTRIBUTIONS OF CORTICAL AND SUBCORTICAL REGIONS

As anticipated in the main text, and shown in Fig. 18, SUBC regions tend to have a larger negative degree than cortical regions.

APPENDIX E: A COMPARISON WITH PEARSON'S TEST OF HYPOTHESIS

Since the signature S_{ij} retains the sign of the Pearson correlation coefficient between the time series i and j , across all

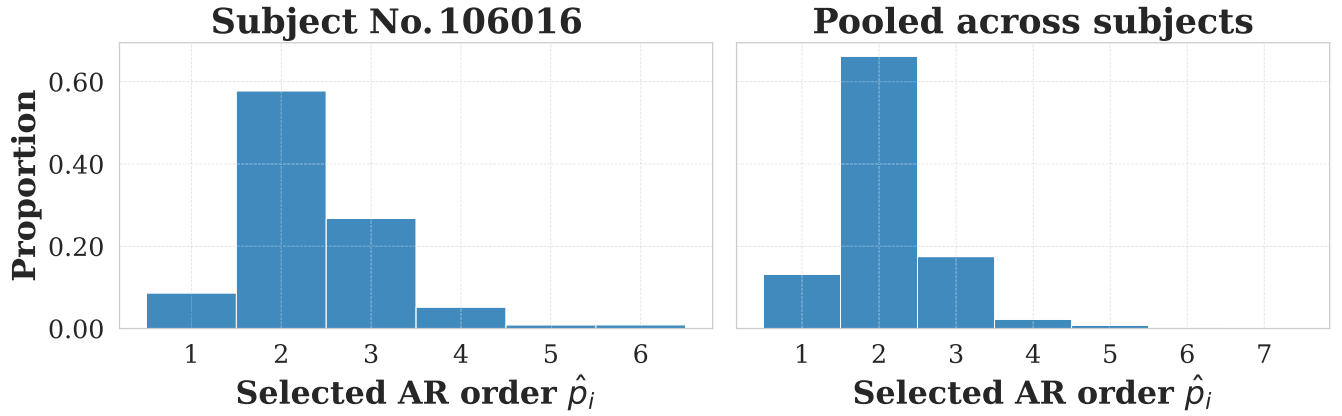


FIG. 17. Distribution of the order of the autoregressive model. Left panel: Distribution of the order of the autoregressive model across the $N = 116$ ROIs of subject No. 106016. Right panel: Distribution of the order of the autoregressive model, pooled across all ROIs of our subjects. Both distributions are concentrated in the range $[2, 3]$, with a tail extending to, at most, $\hat{p}_i = 7$ —a result confirming that a low-order autoregressive model is powerful enough to whiten our preprocessed BOLD signals, consistently with previous findings [100,125].

subjects, to a large extent, one may, now, wonder the extent to which our tests of hypothesis overlap with the Pearson's one. Let us remind what it states: In the case of zero correlation, T pairs of values from a bivariate normal distribution cause the

variable

$$t = r \sqrt{\frac{T-2}{1-r^2}} \quad (\text{E1})$$

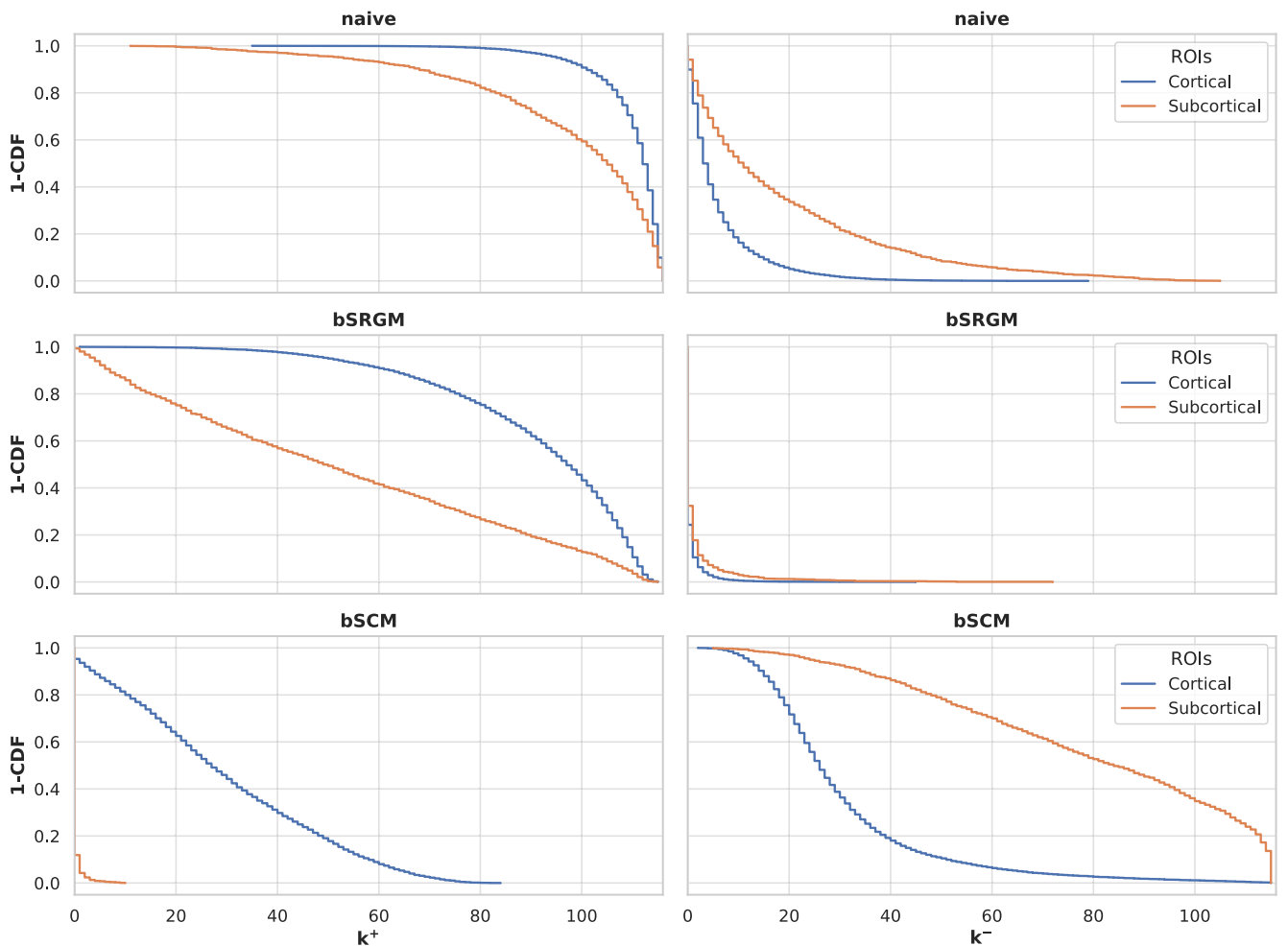


FIG. 18. Cumulative distributions of positive (k^+) and negative (k^-) degrees for cortical and subcortical regions, per benchmark. Analyses have been carried out in a pooled fashion (i.e., across the 100 subjects considered for the present analysis).

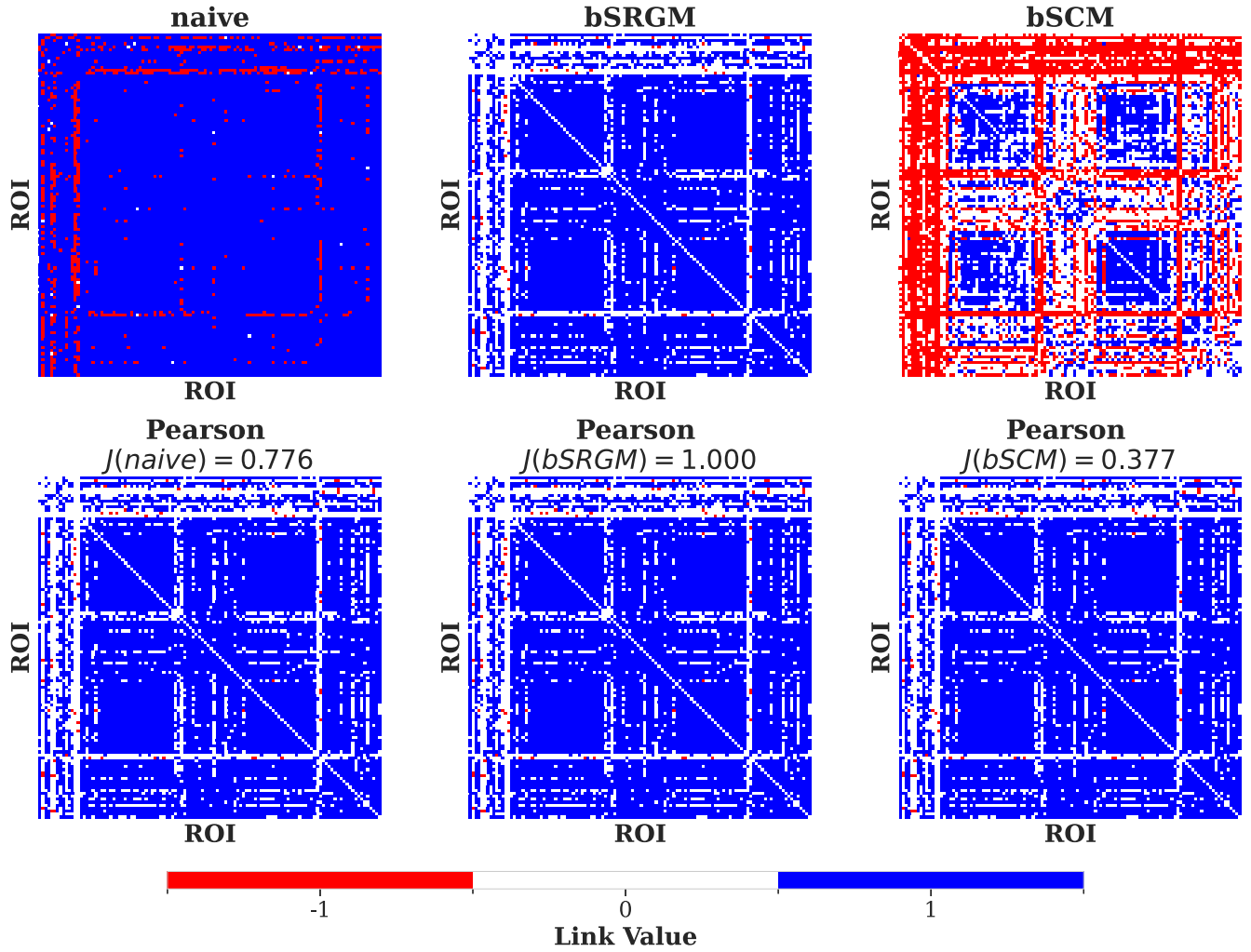


FIG. 19. A comparison between our tests of hypothesis and Pearson's one. As evident from the panels above, referring to subject No. 106016, and further confirmed by the value of the Jaccard similarity, employing Pearson's test of hypothesis to project our biadjacency matrices is equivalent to projecting them by adopting the bSRGM as a benchmark.

to obey a Student's t -distribution with $T - 2$ degrees of freedom; one can, thus, check whether the empirical value of the Pearson correlation coefficient, r , is significantly different from zero in a two-sided fashion. An alternative route is that of considering the Fisher transformation

$$F(r) = \frac{1}{2} \ln \left[\frac{1+r}{1-r} \right], \quad (\text{E2})$$

which obeys a normal distribution with $\mu = F(\rho)$ and $\sigma^2 = 1/(N - 3)$: The test concerning the value $\rho = 0$ can, thus, be

conducted on a single distribution (normal, in this case) as above.

Figure 19 summarizes the outcome of such a comparison on subject No. 106016. As evident, Pearson-induced projection overlaps to a large extent with the bSRGM-induced one: Such a result, probably due to the observed balance between the number of positive and negative signs populating our biadjacency matrices, holds true for all subjects.

- [1] J.-P. Onnela, K. Kaski, and J. Kertész, Clustering and information in correlation based financial networks, *Eur. Phys. J. B* **38**, 353 (2004).
- [2] S. Sinha, A. Chatterjee, A. Chakraborti, and B. K. Chakrabarti, *Econophysics: An Introduction* (Wiley-VCH, Berlin, Germany, 2010).
- [3] J.-P. Bouchaud and M. Potters, *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management* (Cambridge University Press, Cambridge, UK, 2003).

- [4] S. Kulkarni and D. S. Bassett, Toward principles of brain network organization and function, *Annu. Rev. Biophys.* **54**, 353 (2025).
- [5] D. S. Bassett and O. Sporns, Network neuroscience, *Nat. Neurosci.* **20**, 353 (2017).
- [6] We redirect the interested reader to the reviews [126,127], where hundreds of proximity measures (such as the dynamic time warping, the mutual information, and the transfer entropy) are discussed and compared.

- [7] M. Hampson, B. S. Peterson, P. Skudlarski, J. C. Gatenby, and J. C. Gore, Detection of functional connectivity using temporal correlations in MR images, *Hum. Brain Mapp.* **15**, 247 (2002).
- [8] P. Fransson and G. Marrelec, The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: Evidence from a partial correlation network analysis, *NeuroImage* **42**, 1178 (2008).
- [9] M. R. Brier, A. Mitra, J. E. McCarthy, B. M. Ances, and A. Z. Snyder, Partial covariance based functional connectivity computation using Ledoit–Wolf covariance regularization, *NeuroImage* **121**, 29 (2015).
- [10] M. Rahim, B. Thirion, and G. Varoquaux, Population shrinkage of covariance (POSCE) for better individual brain functional-connectivity estimation, *Med. Image Anal.* **54**, 138 (2019).
- [11] S. Huang, J. Li, L. Sun, J. Ye, A. Fleisher, T. Wu, K. Chen, and E. Reiman, Alzheimer’s Disease NeuroImaging Initiative, Learning brain connectivity of Alzheimer’s disease by sparse inverse covariance estimation, *NeuroImage* **50**, 935 (2010).
- [12] R. Salvador, J. Suckling, M. R. Coleman, J. D. Pickard, D. Menon, and E. Bullmore, Neurophysiological architecture of functional magnetic resonance images of human brain, *Cereb. Cortex* **15**, 1332 (2005).
- [13] H. Lee, D. Soo. Lee, H. Kang, B.-N. Kim, and M. K. Chung, Sparse brain network recovery under compressed sensing, *IEEE Trans. Med. Imaging* **30**, 1154 (2011).
- [14] F. Santucci, A. Jimenez-Marin, A. Gabrielli, P. Bonifazi, M. Ibáñez-Berganza, T. Gili, and J. M. Cortes, Partial correlation as a tool for mapping functional-structural correspondence in human brain connectivity, *Network Neurosci.* **9**, 1065 (2025).
- [15] E. Bullmore and O. Sporns, Complex brain networks: Graph theoretical analysis of structural and functional systems, *Nat. Rev. Neurosci.* **10**, 186 (2009).
- [16] M. Rubinov and O. Sporns, Complex network measures of brain connectivity: Uses and interpretations, *NeuroImage* **52**, 1059 (2010).
- [17] M. MacMahon and D. Garlaschelli, Community detection for correlation matrices, *Phys. Rev. X* **5**, 021006 (2015).
- [18] T. Heimo, K. Kaski, and J. Saramäki, Maximal spanning trees, asset graphs and random matrix denoising in the analysis of dynamics of financial networks, *Physica A* **388**, 145 (2009).
- [19] Aaron F. Alexander-Bloch, N. Gogtay, D. Meunier, R. Birn, L. Clasen, F. Lalonde, R. Lenroot, J. Giedd, and E. T. Bullmore, Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia, *Front. Syst. Neurosci.* **4** (2010).
- [20] L. K. Gallos, H. A. Makse, and M. Sigman, A small world of weak ties provides optimal global integration of self-similar modules in functional brain networks, *Proc. Natl. Acad. Sci. USA* **109**, 2825 (2012).
- [21] A dendrogram is constructed by clustering objects in a hierarchical fashion: To this aim, a distance between clusters is needed. Some of these “linkage” algorithms induce an ultrametric distance on the data [27] that has been argued to induce distances progressively less similar to the original ones as higher levels of the taxonomic tree are climbed. According to this intuition, such a method is more reliable to determine the low-level structure of the tree (induced by the strongest correlations) than the high-level one (induced by the weakest correlations) [17,128].
- [22] R. N. Mantegna and H. E. Stanley, *Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, Cambridge, UK, 1999).
- [23] V. Tola, F. Lillo, M. Gallegati, and R. N. Mantegna, Cluster analysis for portfolio optimization, *J. Econ. Dyn. Control* **32**, 235 (2008).
- [24] M. E. J. Newman, Modularity and community structure in networks, *Proc. Natl. Acad. Sci. USA* **103**, 8577 (2006).
- [25] C. Bordier, C. Nicolini, and A. Bifone, Graph analysis and modularity of brain functional connectivity networks: Searching for the optimal threshold, *Front. Neurosci.* **11**, 441 (2017).
- [26] E. Marchese, G. Caldarelli, and T. Squartini, Detecting mesoscale structures by surprise, *Commun. Phys.* **5**, 132 (2022).
- [27] R. Rammal, G. Toulouse, and M. A. Virasoro, Ultrametricity for physicists, *Rev. Mod. Phys.* **58**, 765 (1986).
- [28] G. Caldarelli, M. Cristelli, A. Gabrielli, L. Pietronero, A. Scala, and A. Tacchella, A network analysis of countries’ export flows: Firm grounds for the building blocks of the economy, *PLoS One* **7**, e47278 (2012).
- [29] R. Mastrandrea, A. Gabrielli, F. Piras, G. Spalletta, G. Caldarelli, and T. Gili, Organization and hierarchy of the human functional brain network lead to a chain-like core, *Sci. Rep.* **7**, 4888 (2017).
- [30] R. Mastrandrea, F. Piras, A. Gabrielli, N. Banaj, G. Caldarelli, G. Spalletta, and T. Gili, The unbalanced reorganization of weaker functional connections induces the altered brain network topology in schizophrenia, *Sci. Rep.* **11**, 15400 (2021).
- [31] G. Bardella, A. Bifone, A. Gabrielli, A. Gozzi, and T. Squartini, Hierarchical organization of functional connectivity in the mouse brain: A complex network approach, *Sci. Rep.* **6**, 32060 (2016).
- [32] T. Aste, T. Di Matteo, and S. T. Hyde, Complex networks on hyperbolic surfaces, *Physica A* **346**, 20 (2005).
- [33] M. Tumminello, T. Aste, T. Di Matteo, and R. N. Mantegna, A tool for filtering information in complex systems, *Proc. Natl. Acad. Sci. USA* **102**, 10421 (2005).
- [34] M. Saberi, R. Khosrowabadi, A. Khatibi, B. Misic, and G. Jafari, Topological impact of negative links on the stability of resting-state brain network, *Sci. Rep.* **11**, 2176 (2021).
- [35] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. N. Amaral, T. Guhr, and H. E. Stanley, Random matrix approach to cross correlations in financial data, *Phys. Rev. E* **65**, 066126 (2002).
- [36] A. Utsugi, K. Ino, and M. Oshikawa, Random matrix theory analysis of cross correlations in financial markets, *Phys. Rev. E* **70**, 026110 (2004).
- [37] M. Potters, J.-P. Bouchaud, and L. Laloux, Financial applications of random matrix theory: Old laces and new pieces, *Acta Phys. Pol. B* **36**, 2767 (2005).
- [38] S. M. Zema, G. Fagiolo, T. Squartini, and D. Garlaschelli, Mesoscopic structure of the stock market and portfolio optimization, *J. Econ. Interact. Coord.* **20**, 307 (2025).
- [39] N. Masuda, S. Kojaku, and Y. Sano, Configuration model for correlation matrices preserving the node strength, *Phys. Rev. E* **98**, 012312 (2018).

- [40] S. Kojaku and N. Masuda, Constructing networks by filtering correlation matrices: A null model approach, *Proc. R. Soc. A* **475**, 20190578 (2019).
- [41] We redirect the interested reader to the reviews [129,130], discussing the differences between the existing (classes of) algorithms and comparing their performances.
- [42] L. Lacasa, B. Luque, F. Ballesteros, J. Luque, and J. C. Nuño, From time series to complex networks: The visibility graph, *Proc. Natl. Acad. Sci. USA* **105**, 4972 (2008).
- [43] B. Luque, L. Lacasa, F. Ballesteros, and J. Luque, Horizontal visibility graphs: Exact results for random time series, *Phys. Rev. E* **80**, 046103 (2009).
- [44] L. Lacasa, A. Nunez, É. Roldán, J. MR. Parrondo, and B. Luque, Time series irreversibility: A visibility graph approach, *Eur. Phys. J. B* **85**, 217 (2012).
- [45] T.-T. Zhou, N.-D. Jin, Z.-K. Gao, and Y.-B. Luo, Limited penetrable visibility graph for establishing complex network from time series, *Acta Phys. Sin.* **61**, 030506 (2012).
- [46] S. Supriya, S. Siuly, H. Wang, J. Cao, and Y. Zhang, Weighted visibility graph with complex network features in the detection of epilepsy, *IEEE Access* **4**, 6554 (2016).
- [47] G. Zhu, Y. Li, and P. Wen, Analysis and classification of sleep stages based on difference visibility graphs from a single-channel EEG signal, *IEEE J. Biomed. Health Inf.* **18**, 1813 (2014).
- [48] A. S. L. O. Campanharo, M. I. Sirer, R. D. Malmgren, F. M. Ramos, and L. A. N. Amaral, Duality between time series and networks, *PLoS One* **6**, e23378 (2011).
- [49] A. H. Shirazi, G. Reza Jafari, J. Davoudi, J. Peinke, M. Reza Rahimi Tabar, and M. Sahimi, Mapping stochastic processes onto complex networks, *J. Stat. Mech.: Theory Exp.* (2009) P07046.
- [50] Z. Gao and N. Jin, Complex network from time series based on phase space reconstruction, *Chaos* **19**, 033137 (2009).
- [51] M. Stephen, C. Gu, and H. Yang, Visibility graph based time series analysis, *PLoS One* **10**, e0143015 (2015).
- [52] J. Zhang, X. Luo, and M. Small, Detecting chaos in pseudoperiodic time series without embedding, *Phys. Rev. E* **73**, 016216 (2006).
- [53] M. Small, J. Zhang, and X. Xu, Transforming time series into complex networks, in *Complex Sciences*, edited by J. Zhou, Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, Vol 5 (Springer, Berlin, Heidelberg, 2009), pp. 2078–2089.
- [54] X. Xu, J. Zhang, and M. Small, Superfamily phenomena and motifs of networks induced from time series, *Proc. Natl. Acad. Sci. USA* **105**, 19601 (2008).
- [55] R. V. Donner, Y. Zou, J. F. Donges, N. Marwan, and J. Kurths, Recurrence networks—A novel paradigm for nonlinear time series analysis, *New J. Phys.* **12**, 033025 (2010).
- [56] J. H. Feldhoff, R. V. Donner, J. F. Donges, N. Marwan, and J. Kurths, Geometric detection of coupling directions by means of inter-system recurrence networks, *Phys. Lett. A* **376**, 3504 (2012).
- [57] J. H. Feldhoff, R. V. Donner, J. F. Donges, N. Marwan, and J. Kurths, Geometric signature of complex synchronisation scenarios, *Europhys. Lett.* **102**, 30007 (2013).
- [58] M. C. Romano, M. Thiel, J. Kurths, and W. von Bloh, Multivariate recurrence plots, *Phys. Lett. A* **330**, 214 (2004).
- [59] N. Marwan, N. Wessel, U. Meyerfeldt, A. Schirdewan, and J. Kurths, Recurrence-plot-based measures of complexity and their application to heart-rate-variability data, *Phys. Rev. E* **66**, 026702 (2002).
- [60] J. P. Zbilut, A. Giuliani, and C. L. Webber, Detecting deterministic signals in exceptionally noisy environments using cross-recurrence quantification, *Phys. Lett. A* **246**, 122 (1998).
- [61] L. Lacasa, V. Nicosia, and V. Latora, Network structure of multivariate time series, *Sci. Rep.* **5**, 15508 (2015).
- [62] D. Eroglu, N. Marwan, M. Stebich, and J. Kurths, Multiplex recurrence networks, *Phys. Rev. E* **97**, 012312 (2018).
- [63] G. Varoquaux and R. C. Craddock, Learning and comparing functional connectomes across subjects, *NeuroImage* **80**, 405 (2013).
- [64] F. Váša and B. Mišić, Null models in network neuroscience, *Nat. Rev. Neurosci.* **23**, 493 (2022).
- [65] More explicitly, the practice of employing graph-theoretical estimations for intersubject analyses without comparing them with a proper benchmark gives “[...] a fairly unspecific characterisation of the brain, being fragile to noise [...] Another caveat is that [...] correlation matrices display small-world properties [...] by construction [...] This observation highlights the need for well-defined null hypotheses [...] but also for controlled recovery of brain functional connectivity going beyond empirical correlation matrices [...]” [63].
- [66] J. Park and M. E. J. Newman, Statistical mechanics of networks, *Phys. Rev. E* **70**, 066117 (2004).
- [67] D. Garlaschelli and M. I. Loffredo, Maximum likelihood: Extracting unbiased information from complex networks, *Phys. Rev. E* **78**, 015101(R) (2008).
- [68] T. Squartini and D. Garlaschelli, Analytical maximum-likelihood method to detect patterns in real networks, *New J. Phys.* **13**, 083001 (2011).
- [69] A. Gallo, D. Garlaschelli, R. Lambiotte, F. Saracco, and T. Squartini, Testing structural balance theories in heterogeneous signed networks, *Commun. Phys.* **7**, 154 (2024).
- [70] A. Gallo, D. Garlaschelli, and T. Squartini, Assessing frustration in real-world signed networks: A statistical theory of balance, *Phys. Rev. Res.* **6**, L042065 (2024).
- [71] A. Gallo, F. Saracco, and T. Squartini, Statistically validated projection of bipartite signed networks, *npj Complexity* **2**, 22 (2025).
- [72] M. Saberi, R. Khosrowabadi, A. Khatibi, B. Misic, and G. Jafari, Pattern of frustration formation in the functional brain network, *Network Neurosci.* **6**, 1334 (2022).
- [73] A. Demertzi, A. Kucyi, A. Ponce-Alvarez, G. A. Keliris, S. Whitfield-Gabrieli, and G. Deco, Functional network antagonism and consciousness, *Network Neurosci.* **6**, 998 (2022).
- [74] In what follows, we will precisely consider time series of standardized *residuals* (see Sec. VIII).
- [75] T. Watanabe, S. Hirose, H. Wada, Y. Imai, T. Machida, I. Shirouzu, S. Konishi, Y. Miyashita, and N. Masuda, A pairwise maximum entropy model accurately describes resting-state human brain networks, *Nat. Commun.* **4**, 1370 (2013).
- [76] T. Ezaki, E. Fonseca dos Reis, T. Watanabe, M. Sakaki, and N. Masuda, Closer to critical resting-state neural dynamics in individuals with higher fluid intelligence, *Commun. Biol.* **3**, 52 (2020).

- [77] S. Larivière, C. Paquola, B.-Y. Park, J. Royer, Y. Wang, O. Benkarim, R. Vos de Wael, S. L. Valk, S. I. Thomopoulos, M. Kirschner, *et al.*, The ENIGMA toolbox: Multiscale neural contextualization of multisite neuroimaging datasets, *Nat. Methods* **18**, 698 (2021).
- [78] A. Fronczak, Exponential random graph models, in *Encyclopedia of Social Network Analysis and Mining*, edited by R. Alhajj and J. Rokne (Springer, New York, 2018), p. 810.
- [79] T. Squartini and D. Garlaschelli, *Maximum-Entropy Networks: Pattern Detection, Network Reconstruction and Graph Combinatorics*, *SpringerBriefs in Complexity*, 1st ed. (Springer, Cham, 2017), pp. XII, 116.
- [80] A mathematical construction consisting of the list of elementary events and the associated probabilities.
- [81] The finite scheme obeyed by $C_{ijt} - D_{ijt}$ is the same finite scheme obeyed by $b_{it}b_{jt}$.
- [82] The value p_{th} will be determined later, at the third step of our algorithm: Hereby, a merely discursive introduction to the proper validation step is provided.
- [83] A simple argument can, indeed, be provided. The probability that, solely by chance, at least one out of $|H|$ hypotheses is rejected, and the corresponding link validated, amounts to $\text{FWER} = 1 - (1 - t)^{|H|}$. Hence, $\text{FWER} \simeq 1$ for (just) $|H| = 100$ tests conducted at the significance level of $t = 0.05$.
- [84] C. E. Bonferroni, *Teoria Statistica delle Classi e Calcolo delle Probabilità*, *Pubblicazioni del R. Istituto superiore di scienze economiche e commerciali di Firenze* **8** (Seeber, Firenze, 1936).
- [85] M. Tumminello, S. Miccichè, F. Lillo, J. Piilo, and R. N. Mantegna, Statistically validated networks in bipartite complex systems, *PLoS One* **6**, 1 (2011).
- [86] S. Gualdi, G. Cimini, K. Primicerio, R. Di Clemente, and D. Challet, Statistically validated network of portfolio overlaps and systemic risk, *Sci. Rep.* **6**, 39467 (2016).
- [87] Indeed, the stricter condition $\text{FWER} = 0.05$ leads to the threshold $p_{th} = t/|H|$, rapidly vanishing as $|H|$ grows: As a consequence, very sparse (if not empty) projections are obtained.
- [88] Y. Benjamini and Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing, *J. R. Stat. Soc. Ser. B Methodol.* **57**, 289 (1995).
- [89] R. Marcaccioli and G. Livan, Maximum entropy approach to multivariate time series randomization, *Sci. Rep.* **10**, 10656 (2020).
- [90] N. Vallarano, M. Bruno, E. Marchese, G. Trapani, F. Saracco, G. Cimini, M. Zanon, and T. Squartini, Fast and scalable likelihood maximization for exponential random graph models with local constraints, *Sci. Rep.* **11**, 15227 (2021).
- [91] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, and K. W.-M. H. C. Ugurbil, The WU-Minn Human Connectome Project: An overview, *NeuroImage* **80**, 62 (2013).
- [92] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. Van Essen, and M. Jenkinson, The minimal preprocessing pipelines for the Human Connectome Project, *NeuroImage* **80**, 105 (2013).
- [93] L. Griffanti, G. Salimi-Khorshidi, C. F. Beckmann, E. J. Auerbach, G. Douaud, C. E. Sexton, E. Zsoldos, K. P. Ebmeier, N. Filippini, C. E. Mackay, *et al.*, ICA-based artefact removal and accelerated fMRI acquisition for improved resting state network imaging, *NeuroImage* **95**, 232 (2014).
- [94] T. D. Satterthwaite, M. A. Elliott, R. T. Gerraty, K. Ruparel, J. Loughead, M. E. Calkins, S. B. Eickhoff, H. Hakonarson, R. C. Gur, R. E. Gur, *et al.*, An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data, *NeuroImage* **64**, 240 (2013).
- [95] A. Schaefer, R. Kong, E. M. Gordon, T. O. Laumann, X.-N. Zuo, A. J. Holmes, S. B. Eickhoff, and B. T. T. Yeo, Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI, *Cereb. Cortex* **28**, 3095 (2018).
- [96] Y. Tian, D. S. Margulies, M. Breakspear, and A. Zalesky, Topographic organization of the human subcortex unveiled with functional connectivity gradients, *Nat. Neurosci.* **23**, 1421 (2020).
- [97] K. Murphy, R. M. Birn, D. A. Handwerker, T. B. Jones, and P. A. Bandettini, The impact of global signal regression on resting state correlations: Are anti-correlated networks introduced? *NeuroImage* **44**, 893 (2009).
- [98] K. J. Friston, P. Jezzard, and R. Turner, Analysis of functional MRI time-series, *Hum. Brain Mapp.* **1**, 153 (1994).
- [99] G. H. Glover, Deconvolution of impulse response in event-related BOLD fMRI, *NeuroImage* **9**, 416 (1999).
- [100] M. W. Woolrich, B. D. Ripley, M. Brady, and S. M. Smith, Temporal autocorrelation in univariate linear modeling of fMRI data, *NeuroImage* **14**, 1370 (2001).
- [101] To this aim, we have employed the AutoReg estimator from the statsmodels Python library.
- [102] M. S. Bartlett, *An Introduction to Stochastic Processes* (Cambridge University Press, Cambridge, 1955).
- [103] M. B. Priestley, *Spectral Analysis and Time Series* (Academic Press, London, 1981).
- [104] J. Durbin, Tests of serial independence based on the cumulated periodogram, *Bull. Int. Stat. Inst.* **42**, 1040 (1967).
- [105] F. Saracco, M. J. Straka, R. Di Clemente, A. Gabrielli, G. Caldarelli, and T. Squartini, Inferring monopartite projections of bipartite networks: An entropy-based approach, *New J. Phys.* **19**, 053022 (2017).
- [106] As highlighted in Ref. [107], “[...] whereas cooperative (positive-valued) interactions in the effective connectivity tend to be strong, modular and relatively short-range, competitive interactions are weaker but more long-range, more diffuse and less clustered [...]”
- [107] A. I. Luppi, Y. Sanz Perl, J. Vohryzek, P. A. M. Mediano, F. E. Rosas, F. Milisav, L. E. Suarez, S. Gini, D. Gutierrez-Barragan, A. Gozzi, B. Misic, G. Deco, and M. L. Kringelbach, Competitive interactions shape mammalian brain network dynamics and computation, *Nat. Neurosci.* **29**, 915 (2026).
- [108] Speaking of cortical regions, there it is reported that “[...] it is less likely that competitively-interacting neighbours of a node will themselves be interacting competitively [...]”
- [109] As confirmed by a one-sided Wilcoxon signed-rank test at the confidence level of 5%.
- [110] B. T. Thomas Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni, B. Fischl, H. Liu, and R. L. Buckner,

- The organization of the human cerebral cortex estimated by intrinsic functional connectivity, *J. Neurophysiol.* **106**, 1125 (2011).
- [111] As highlighted in Ref. [72], “Our results suggest that SUBC regions have a prominent role in frustration formation in the brain network at both nodal and connectional levels. So they can bring instability and alter properties that facilitate systemic level neural changes, providing adaptive characteristics to the brain.”
- [112] In fact, both RI and NMI could be directly maximized.
- [113] Z. S. Saad, S. J. Gotts, K. Murphy, G. Chen, H. J. Jo, A. Martin, and R. W. Cox, Trouble at rest: How correlation patterns and group differences become distorted after global signal regression, *Brain Connect.* **2**, 25 (2012).
- [114] M. P. van den Heuvel, S. C. de Lange, A. Zalesky, C. Seguin, B. T. Yeo, and R. Schmidt, Proportional thresholding in resting-state fMRI functional connectivity networks and consequences for patient-control connectome studies: Issues and recommendations, *NeuroImage* **152**, 437 (2017).
- [115] M. D. Greicius, B. Krasnow, A. L. Reiss, and V. Menon, Functional connectivity in the resting brain: A network analysis of the default mode hypothesis, *Proc. Natl. Acad. Sci. USA* **100**, 253 (2003).
- [116] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. V. Essen, and M. E. Raichle, The human brain is intrinsically organized into dynamic, anticorrelated functional networks, *Proc. Natl. Acad. Sci. USA* **102**, 9673 (2005).
- [117] M. Hampson, N. Driesen, J. K. Roth, J. C. Gore, and R. T. Constable, Functional connectivity between task-positive and task-negative brain areas and its relation to working memory performance, *Magn. Reson. Imaging* **28**, 1051 (2010).
- [118] M. Saberi, A. HaqiqiFar, A. Abdolalizadeh, B. Misic, and A. Khatibi, Empirical evidence for structural balance theory in functional brain networks, *Front. Network Physiol.* **5**, 1681597 (2025).
- [119] F. Zamani Esfahlani, Y. Jo, J. Faskowitz, L. Byrge, D. P. Kennedy, O. Sporns, and R. F. Betzel, High-amplitude cofluctuations in cortical activity drive functional connectivity, *Proc. Natl. Acad. Sci. USA* **117**, 28393 (2020).
- [120] S. Harita, D. Momi, Z. Wang, S. P. Bastiaens, and J. D. Griffiths, The role of inhibition in resting-state fMRI negative correlations, *bioRxiv*, doi:10.1101/2024.03.01.583030.
- [121] <https://www.humanconnectome.org/study/hcp-young-adult>.
- [122] <https://www.humanconnectome.org/storage/>.
- [123] SIBERIA v0.1.1 is available both on github at <https://github.com/MarsMDK/SIBERIA>, on pypi at <https://pypi.org/project/siberia/> and on Zenodo with the DOI: <https://doi.org/10.5281/zenodo.21036013>.
- [124] Y. Benjamini and D. Yekutieli, The control of the false discovery rate in multiple testing under dependency, *Ann. Stat.* **29**, 1165 (2001).
- [125] M. R. Arbabshirani, E. Damaraju, R. Phlypo, S. Plis, E. Allen, S. Ma, D. Mathalon, A. Preda, J. G. Vaidya, J. T. Voyvodic, and V. D. Calhoun, Impact of autocorrelation on functional connectivity, *NeuroImage* **102**, 294 (2014).
- [126] O. M. Cliff, A. G. Bryant, J. T. Lizier, N. Tsuchiya, and B. D. Fulcher, Unifying pairwise interactions in complex dynamics, *Nat. Comput. Sci.* **3**, 883 (2023).
- [127] Z.-Q. Liu, A. I. Luppi, J. Y. Hansen, Y. E. Tian, A. Zalesky, B. T. Yeo, B. D. Fulcher, and B. Misic, Benchmarking methods for mapping functional connectivity in the brain, *Nat. Methods* **22**, 1593 (2025).
- [128] G. W. Milligan, Ultrametric hierarchical clustering algorithms, *Psychometrika* **44**, 343 (1979).
- [129] V. F. Silva, M. E. Silva, P. Ribeiro, and F. Silva, Time series analysis via network science: Concepts and algorithms, *WIREs Data Min. Knowl. Discovery* **11**, e1404 (2021).
- [130] Y. Zou, Reik V. Donner, N. Marwan, J. F. Donges, and J. Kurths, Complex network approaches to nonlinear time series analysis, *Phys. Rep.* **787**, 1 (2019).