Learning under constraints: inference, selection and validation with maximum entropy models

Giuffrida, F.

CHAPTER 5

Validation of higher-order interactions: application to brain data

This chapter is based on: F. Giuffrida, E. Agrimi, M. Marzi, A. Gallo, T. Squartini and T. Gili. Signatures of higher-order coordination in the human brain. *In preparation.*

Abstract

Null models and statistical tests play a key role in validating empirical patterns in data, including relational structures that may involve interactions of arbitrary order. This perspective is particularly relevant in integrative neuroscience, where cognitive processes are understood to emerge from the coordinated activity of multiple brain regions. Here, we propose an algorithm to detect triadic hyperedges of cooperating brain regions, rooted in entropy-maximization. Within such a statistical framework, it is possible to constrain the number of lower-order co-activation patterns, thus enabling the identification of higher-order, over-expressed interactions. Our method generalizes the dyadic validation scheme by returning a matrix of hyperedge-specific p-values, from which a validated hypergraph of brain interactions can be constructed after correcting for multiple hypothesis testing. When applied to human resting-state fMRI time series, our approach suggests that subcortical, limbic, and transmodal regions are more frequently involved in triadic interactions that lack full dyadic support, whereas somatomotor and visual regions are comparatively more associated with triples consistent with dyadic closure (i.e., 2-simplices). Moreover, these purely higher-order interactions tend to link more spatially distant regions and exhibit lower redundancy than 2-simplicial configurations. By combining an information-theoretic perspective with hypergraph modeling, this framework provides a principled way to characterize higher-order functional coordination in the brain.

5.1 Introduction

The human brain is a complex system constituted of interacting neurons whose organized activity allows properties such as behaviors and cognition to emerge [96, 97]. Functional connectivity (FC) analyses represent the most popular approach to quantifying these interactions from fMRI data, typically proxied by correlations between (the activity of) different regions of the brain [98, 99, 100]. Pairwise analyses of this kind have been instrumental in mapping the large-scale architecture of the human brain, revealing a small-world, modular, and hierarchical organization [101]. Still, the neuroscientific literature has mainly focused on two-body interactions so far, thus shedding light on only a subset of the interactions that can take place in the brain, while disregarding those involving more than two regions, i.e., the so-called *higher-order interactions* (HOIs). Recent literature has successfully introduced HOIs in various contexts, including biological [102], ecological [103], financial [33], and social systems [104]; recent evidence has shown that HOIs are also key ingredients for the emergence of higher-level brain functions, such as cognition and consciousness [32, 33, 34, 35].

Several measures of HOIs have been introduced in neuroscience, the definitions of which originate from multiple theoretical contexts, including statistical inference [105], topological data analysis [33], multivariate information theory and partial information decomposition [106, 107, 108, 109], and machine learning [110].

Information-theoretic studies have characterized HOIs between brain regions in terms of *redundancy* and *synergy*: while redundancy emerges when neural populations convey overlapping information, synergy dominates when the information coming from multiple regions is integrated. The contributions of redundancy and synergy have been quantified by introducing the *S-information* [111] and *O-information* [106], quantities which have been shown to provide an exhaustive description of the higher-order interactions between three bodies [112]. These two modalities differ in their spatial distribution: while redundancy dominates in regions that process information from a single, perceptual or motor, domain – like the primary sensory and the motor cortices – synergy becomes more prominent within association networks - like the frontoparietal, default-mode and dorsal-attention networks [32, 108].

At the same time, the topological data analysis of FC data [113, 114, 115] often identifies the presence of higher-order interactions together with that of all lower-order counterparts [116]. This simplicial closure condition aligns naturally with redundant interactions, but precludes the identification of collective effects that arise without lower-order support. Models based on the concept of hypergraph, instead, allow for higher-order structures that do not require the presence of the corresponding lower-order hyperedges [114]; yet, even within this framework, applications have lacked a statistically principled procedure to determine whether empirical group-wise interactions are a consequence of simpler lower-order constraints. As a result, the extent to which higher-order functional motifs rely upon lower-order structures, thus satisfying the simplicial closure condition, remains unknown: answering this question requires a statistical approach capable of testing whether higher-order behaviors are consistent with (expectations based upon) lower-order dependencies or, instead, exceed them to

a significantly large extent, thereby reflecting a genuinely collective organization.

The framework based on entropy maximization provides a rigorous solution to this problem [31, 6], as it allows one to test whether a given pattern carries information beyond that embodied by the enforced constraints. To apply such a framework to brain activity data, we represent resting-state fMRI time series as a binary, undirected, bipartite network: a link between a region of interest (ROI) and a (discrete) time point indicates that the region is *active* at that specific time; we then turn such a representation into a functional hypergraph, where nodes are connected only if they are found to be active simultaneously a significantly large amount of time. To assess the statistical significance of dyadic interactions, we employ the Bipartite Configuration Model (BiCM) [117, 55], a linear Exponential Random Graph Model (ERGM) controlling for both the regional and the temporal heterogeneity; to assess the statistical significance of triadic interactions, we introduce the Bipartite Partial Local Overlap Model (BiPLOM), a non-linear ERGM that, in addition to controlling for regional and temporal heterogeneity, explicitly accounts for each region's aggregated propensity to co-activate in pairs. Used in combination with the BiCM, the BiPLOM allows us to assess which statistically validated higher-order interactions are better represented as simplicial complexes, in which the underlying pairwise interactions are validated as well (via the BiCM) or, conversely, if they require the hypergraphs formalism, in which the underlying pairwise interactions are not necessarily statistically significant.

We show that introducing second-order constraints reshapes the landscape of validated higher-order interactions, revealing distinct regimes of coordination across brain systems. Triads supported by validated pairwise interactions are spatially compact and display positive O-information, consistent with redundancy. In contrast, triads lacking dyadic support are more spatially dispersed and exhibit reduced or null redundancy. Importantly, the distribution of these regimes is functionally heterogeneous: limbic, subcortical, and transmodal regions show a higher prevalence of triadic configurations that would remain undetected in a simplicial framework relying exclusively on dyadic validation.

5.2 Results

5.2.1 From time series to bipartite networks

FC studies are typically carried out by exploiting the statistical (ir)regularities of the BOLD fMRI time series [118]: the BOLD signal reflects local changes in blood oxygenation, and large deviations are, in fact, believed to reflect episodes of neural activation. In this respect, a foundational method is the so-called *point process analysis* (PPA) [119], which converts normalized BOLD signals into (likely sparse) sequences of discrete events by retaining only values exceeding a predefined threshold. Despite the drastic reduction of dimensionality, PPA recovers a resting-state network (RSN) structure comparable to that obtained via standard correlation-based methods. Extensions of this approach include the so-called *co-activation pattern* (CAP) analysis [120, 121, 122], which clusters high-amplitude BOLD frames into recurring

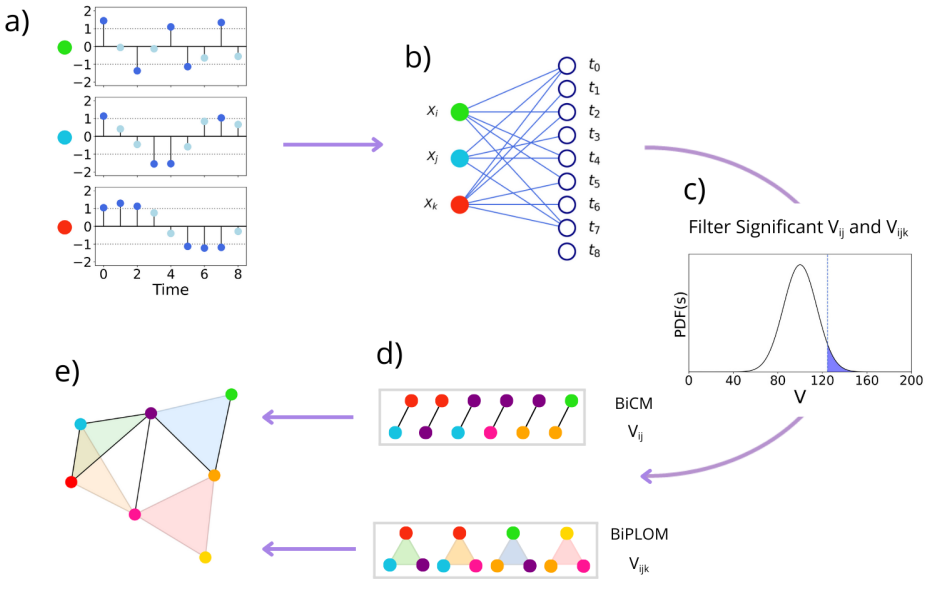


Figure 5.1: ***Infographics illustrating the pipeline of our analysis.*** *Pictorial representation of the pipeline we follow in the present contribution, from the registration of brain activity to the definition of a hypergraph the nodes of which are brain regions:* **a)** *resting-state fMRI time-series are standardized and binarized;* **b)** *time series are mapped to a binary, undirected, bipartite graph;* **c)** *the empirical abundances of dyads, V_{ij}, are validated against the benchmark named Bipartite Configuration Model (BiCM) [117, 55], while the empirical abundances of triads, V_{ijk}, are validated against the benchmark named Bipartite Partial Local Overlap Model (BiPLOM);* **d)** *the sets of validated dyads and triads are identified by the two models;* **e)** *a binary, undirected hypergraph is populated with the statistically validated interactions.*

network states, and event-based analyses in EEG data [123, 124, 125], which reveal activation patterns on sub-second timescales.

As these findings support the intuition that large fluctuations carry key information about the collective dynamics of brain, here we represent a subject's activation data as a binary, rectangular matrix $\mathbf{B} = \{b_{it}\}_{\substack{i\in\{1\dots N\}\\ t\in\{1\dots T\}}}$, where the rows correspond to the N brain regions of interest (ROIs) and the columns correspond to the T time points (see Section 5.3.1). The generic entry of such a matrix is then defined as

$$b_{it} = \begin{cases} 1, & \text{if } |z_i(t)| > \tau, \\ 0, & \text{otherwise.} \end{cases} \tag{5.1}$$

where $z_i(t) = (x_i(t) - \mu_i)/\sigma_i$, with μ_i and σ_i being the average and standard deviation of the time series $\{x_i(t)\}_{t\in\{1\dots T\}}$ associated with region i [119, 123, 126]. In other words, region i is flagged as 'active' whenever its standardized signal deviates from the average by more than a fixed threshold τ: a typical choice is $\tau = 1$, corresponding to one standard deviation. The resulting matrix $\mathbf{B}$ can thus be interpreted as a

biadjacency matrix.

Although associations between regions and time points can be represented via a bipartite network that retains the information about the amplitude and sign of signals [127, 128], here we restrict our attention to the occurrences of extreme events, a simpler representation that offers a valid basis for statistical validation via maximum entropy null models.

5.2.2 Inferring dyadic and triadic interactions

Once a multivariate time series has been represented as a bipartite network, it is possible to leverage the information encoded in this representation to construct a network connecting brain regions that exhibit a similar behavior: what is obtained is known as *monopartite projection* [129, 55, 130, 131, 127].

To this aim, we adopt and extend the framework introduced in [55] to identify significantly similar interactions between pairs and triplets of brain regions. Specifically, two brain regions are considered (functionally) similar whenever the number of co-activations exceeds the expectation from a null model that constrains the activity of each brain region and the number of regions active at each time; analogously, three brain regions are considered as similar whenever they are joint active more frequently than expected under a null model additionally constraining the tendency of each node to co-activate in pairs.

More formally, the number of co-activations for the pair of regions (i, j) is defined as

$$V_{ij} = \sum_{t=1}^{T} b_{it} b_{jt}, \tag{5.2}$$

i.e., the number of time points at which both regions i and j are found to be active; analogously, the number of co-activations for the triplet of regions (i, j, k) is defined as

$$V_{ijk} = \sum_{t=1}^{T} b_{it} b_{jt} b_{kt}. \tag{5.3}$$

For each n-tuple of regions ($n = 2, 3$), our algorithm quantifies the statistical significance of the empirical number of co-activations. To this aim, we employ the BiCM [117, 55] to assess the statistical significance of the empirical abundances of dyadic interactions, say $\{V_{ij}\}_{i,j\in\{1\dots N\}}$, and the BiPLOM to assess the statistical significance of the empirical abundances of triadic interactions, say $\{V_{ijk}\}_{i,j,k\in\{1\dots N\}}$ (see Section 5.3.2).

Once the null model has been specified, each empirical value, say V^*_{ij} and V^*_{ijk}, is compared to the ensemble distribution of the corresponding statistics, yielding a p-value (i.e. the probability of observing a number of interactions greater than, or equal to, the empirical one) per tested pair and triplet of areas, reading

$$p\text{-value} = \Pr[V \geq V^* \mid \text{null model}] \tag{5.4}$$

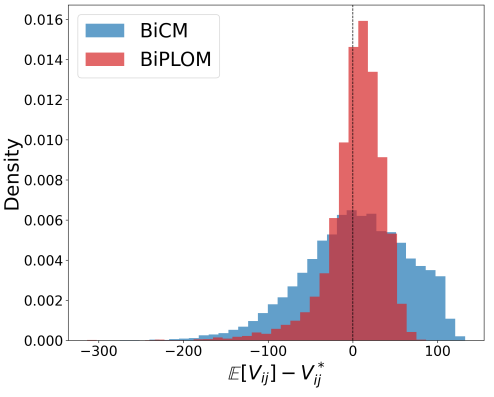


Figure 5.2: ***Co-activations residuals under BiCM and BiPLOM.*** *For a representative subject (id: 366446), histogram of residuals $\mathbb{E}[V_{ij}] - V^*_{ij}$ under BiCM (blue) and BiPLOM (red): BiPLOM yields a sharper, higher peak around zero, indicating a closer agreement between expected and empirical pairwise co-activations once pairwise tendencies are controlled for.*

(where, for the sake of generality, we have dropped the indices).

Since thousands of tests are performed across pairs of regions and millions across triplets of regions, the resulting p-values are corrected for multiple hypothesis testing. Hereby, the Benjamini-Hochberg procedure controlling for the percentage of 'false discoveries' (equivalently, 'incorrectly rejected null hypotheses'), i.e., the False Discovery Rate (FDR) [132], has been applied. The correction for dyads is performed separately from the correction for triads, thereby controlling the FDR within each order of interactions (as already done, for example, in [133]). Interactions whose p-value falls below the adjusted threshold (the single-test significance level of which has been set to $\alpha = 0.05$) are statistically validated, hence retained (see Section 5.3.3).

The complete workflow, from data pre-processing to validation and from validation to the analysis of the resulting higher-order network, is depicted in fig. 5.1.

5.2.3 Impact of second-order constraints on triplet validation

In principle, triadic interactions could be validated under both BiCM and BiPLOM. However, the two null models differ in the constraints they enforce. The BiCM imposes only first-order constraints, namely the activity of each region and the number of active regions at each time point. As such, it does not account for second-order co-activation tendencies.

The BiPLOM extends this framework by additionally constraining, for each region, its aggregated propensity to co-activate in pairs. Importantly, this does not mean that BiPLOM constrains each individual pairwise co-activations V_{ij}. Imposing such pairwise constraints would require $O(N^2)$ parameters and would render the model substantially less tractable. Instead, BiPLOM constrains node-level aggregated co-activations V_i, which encode each region's overall tendency to co-activate in pairs, without imposing constraints on each individual pairwise motif V_{ij}. The formal

definition of the model and its constraints is provided in Section 5.3.2. Notably, although V_{ij} are not directly constrained, enforcing V_i indirectly improves the reproduction of pairwise co-activations V_{ij} as well; this effect is illustrated in Figure 5.2 for a representative subject.

Because of their structural difference, the two models yield different outcomes in triplet validation. Since BiCM does not account for aggregated second-order tendencies, it is more likely to flag as significant triplets whose apparent higher-order behavior is largely driven by pairwise co-activations. In other words, part of the signal detected by BiCM may reflect dyadic structure that is not controlled for in its null model.

Empirical evidence for this interpretation comes from the analysis of closed triangles in the validated pairwise network. We found that BiCM systematically validates all such triangles also at the triadic level, indicating that, under BiCM, dyadic closure is automatically promoted to triadic significance. In contrast, BiPLOM validates only a subset of these configurations, with the retained fraction varying substantially across subjects (Figure 5.3a).

To further quantify the difference between the two models, we compared, for each subject, the sets of triplets validated under BiCM and BiPLOM, denoted by $\mathcal{T}^{\mathrm{BiCM}}$ and $\mathcal{T}^{\mathrm{BiPLOM}}$. Over their union, we computed the proportion of triplets validated exclusively by BiCM, exclusively by BiPLOM, or by both models (Figure 5.3b). Across subjects, BiPLOM validates fewer triplets overall, consistent with a stricter null model that accounts for aggregated pairwise tendencies. At the same time, the two validation sets are not simply nested: although a substantial fraction of triplets is validated by both models, BiPLOM identifies a non-negligible subset of configurations that BiCM does not recover.

5.2.4 Geometric and informational signatures of validated triplets

To characterize the triadic interaction between regions of interest, we considered three specific classes of validated triplets: *connected triplets*, in which all three constituent pairs are validated by BiCM and can therefore be modeled as simplicial complexes, *mixed triplets*, in which one or two of the constituent pairs are validated by BiCM (Fig.5.4a), and *unconnected triplets*, in which none of the constituent pairs are validated by BiCM.

To test whether these classes exhibit distinct spatial footprints, spanning across the brain's lobes or staying local, we quantified the geometric area of each validated triplet using the centroids of its constituent ROIs. To aggregate the results, we first calculated the area of each validated triplet. Then, for each individual, the areas of connected, mixed, and unconnected triplets were averaged separately, yielding a mean value per category per subject. Finally, a relative mean area was obtained for each subject by dividing its mean area by the median area of all possible triplets of regions in our atlas. This procedure yielded three sets of relative average areas, one for each subject in each set, allowing comparison of the typical spatial footprints of connected, unconnected, and mixed triplets across the entire cohort (Fig. 5.5). Unconnected and mixed triplets consistently exhibit larger spatial areas, suggesting a more dispersed

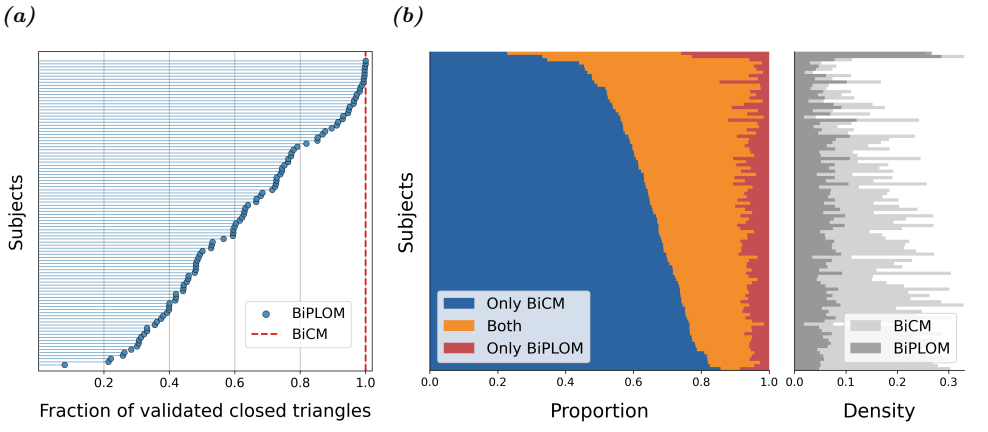


Figure 5.3: ***Comparison of BiCM and BiPLOM triplets validations.*** ***(a)*** *Fraction of closed triangles (i.e., triads of nodes whose three pairs are validated by BiCM) that are further validated as statistically significant triplets by BiPLOM (blue dots) and by BiCM (red dashed line), for all subjects ranked by their BiPLOM value. While BiCM consistently validates all these triangles across subjects, BiPLOM exhibits significant variability across subjects.* ***(b)*** *The left panel decomposes, for each subject, the union of validated triplets into three disjoint sets: triplets validated exclusively by BiCM (blue), exclusively by BiPLOM (red), and by both models (orange). Subjects are ordered according to the proportion of triplets validated exclusively by BiCM. The right panel shows the across-subject validation densities (validated triplets over the total number of possible triplets) for both models. Overall, BiPLOM yields a more conservative set of validated triplets than BiCM, while still detecting configurations that BiCM does not.*

anatomical footprint that may support integrative processes across different areas. In contrast, connected triplets tend to remain more spatially localized, supporting the coordination of activity across more homogeneous, spatially closer areas.

In addition, we investigated whether connected, mixed, and unconnected triplets also differ in their O-information values, which quantify the balance between redundancy and synergy in multivariate interactions [106]. For each subject and each validated triplet (i, j, k), we considered the corresponding time series array $\mathbf{X}_{lt}$, with $l \in \{i, j, k\}$ and $t \in \{1, \ldots, T\}$, and calculated the O-information $\Omega(\mathbf{X}_{lt})$ [134]. For each subject, the O-information values were averaged across all validated triplets. Unconnected triplets show O-information values concentrated close to zero, indicating little net redundancy. Mixed triplets display moderately positive values, while connected triplets exhibit the largest positive O-information, reflecting increasing redundancy as the number of pairwise-supported interactions within the triplet grows (Fig. 5.6).

5.2.5 Mesoscale-level characterization

Finally, we assessed mesoscale patterns by categorizing validated triplets according to the Resting State Networks (RSNs) defined by Yeo et al. [135], to which their con-

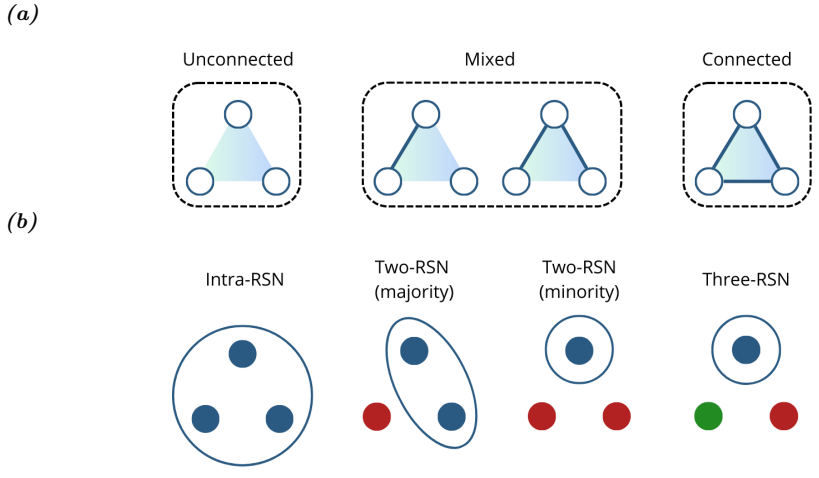


Figure 5.4: ***Triplets classification*** ***(a)*** *Validated triplets were classified according to the number of validated pairwise interactions among their underlying pairs. We distinguish between connected triplets, where all internal pairs are validated, mixed triplets, where one or two internal pairs are validated, and unconnected triplets, where no internal pair is validated.* ***(b)*** *Validated triplets were also classified based on the distribution of their nodes across resting-state networks. Specifically, for a given node in a resting-state network R, a validated triplet involving the node is said to be intra-RSN if it involves only nodes in R; two-RSN (majority) if it involves two distinct RSN and the majority of nodes belong to R; two-RSN (minority) if it involves two distinct RSN and only one node belongs to R; three RSN if it involves three distinct RSN, including R.*

stituent regions are assigned. This classification enables the analysis of co-activation motifs both within and across RSNs, thereby facilitating the interpretation of higher-order interactions in terms of known large-scale functional systems. The validated triplets were grouped into connected, mixed, and unconnected triplets. Moreover, for each RSN R, triplets involving at least one node in R were further distinguished into four categories (Fig. 5.4b): (i) intra-RSN triplets, in which all three nodes belong to R; (ii) *two-RSN (majority)* triplets, involving exactly two RSNs with two nodes in R; (iii) *two-RSN (minority)* triplets, involving exactly two RSNs with a single node in R; (iv) *three-RSN* triplets, spanning three distinct RSNs, with only one node in R.

To visualize the spatial distribution of brain regions across these motif categories, we computed regional density maps for each configuration separately. Specifically, we counted the number of validated triplets of category c that include region r in each subject s, yielding $N_s(r, c)$. These counts were normalized within each subject and configuration to obtain proportional densities expressed by the higher order regional density index $HORD_s(r, c)$. Group-level maps, one for each configuration c, were generated by averaging $HORD_s(r, c)$ across subjects.

In addition to the regional density maps, we derived regional connectedness maps to summarize whether each region tended to participate more often in connected or unconnected motifs. For each configuration type, we computed the higher-order regional flexibility index $HORF_s(r, c)$ for every subject and region, and then averaged

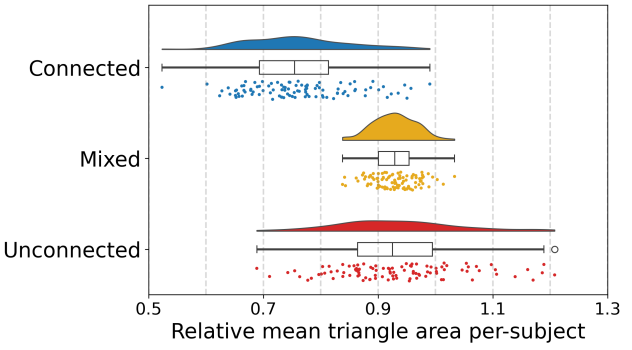


Figure 5.5: ***Comparison of relative triangle areas for validated connected, unconnected, and mixed triplets.*** *Rain plots illustrate the distribution of relative mean triangle areas across subjects (each of which is represented here by a dot). Triplets are classified according to the validation of their constituent pairs: connected triplets (all pairs validated by BiCM), unconnected triplets (no pairs validated), and mixed triplets (one or two pairs validated by BiCM). For each subject, triangle areas were averaged across all validated triplets and divided by the median of all the possible triangle areas in our atlas. Unconnected and mixed triplets consistently exhibit larger spatial areas, suggesting a more dispersed anatomical footprint that may support integrative processes across different areas. In contrast, connected triplets tend to remain more spatially localized, supporting the coordination of activity across more functionally homogeneous and spatially closer areas.*

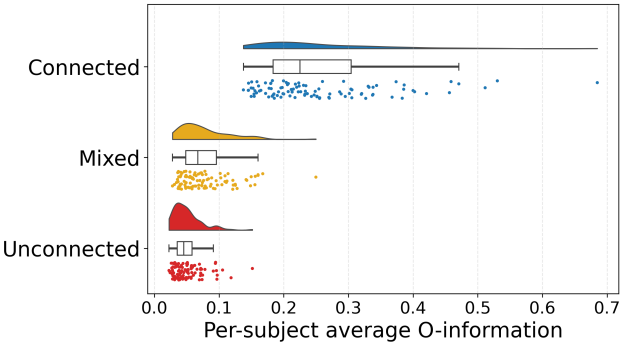


Figure 5.6: ***Comparison of O-information for validated connected, unconnected, and mixed triplets.*** *Rain plots illustrate the distribution of mean O-information values across subjects (each represented by a dot). Triplets are classified according to the validation of their constituent pairs: connected triplets (all pairs validated by BiCM), unconnected triplets (no pairs validated), and mixed triplets (one or two pairs validated by BiCM). For each subject, O-information values were averaged across all validated triplets. Connected triplets consistently show larger positive O-information, reflecting redundancy and stronger lower-order dependencies. Unconnected triplets reveal informational signatures more consistent with genuine higher-order interdependencies. Mixed triplets show an intermediate regime of coexistence between lower and higher order interactions.*

these values across subjects to obtain group-level connectedness maps. The resulting maps illustrate, for each configuration, the relative predominance of connected versus unconnected triplets across the cortex (Fig.5.7).

To aggregate these results at the RSN level, we considered, for each RSN R, all triplets including R. We then computed the proportion of connected, mixed, and unconnected triplets with a given configuration, normalizing by the total number of validated triplets. Fig. 5.8 summarizes the aggregated results, showing whether connected, mixed or unconnected triplets better describe interactions between and within specific RSNs.

5.3 Methods

5.3.1 Dataset description

We analyzed 3T resting-state fMRI data from the unrelated subjects subset of the Young Adult Human Connectome Project (HCP) [136]. The complete data set comprises approximately 1,200 participants aged 22 to 35 years, each scanned with a high-resolution multiband EPI BOLD sequence (TR = 720 ms, voxel size = 2 mm isotropic, 72 slices) optimized for whole-brain coverage and functional coherence. During acquisition, subjects were at rest with their eyes open and focused on a central cross. The imaging parameters included a repetition time (TR) of 720 ms, an echo time (TE) of 33.1 ms, a flip angle of 52°, and a spatial resolution of 2 mm isotropic across 72 slices. Each scanning session comprised two 14.5-minute runs with opposite phase-encoding directions (RL and LR), yielding 1,200 frames per run.

Data preprocessing followed the HCP minimal pipeline [137], which includes motion correction, intensity normalization, and spatial alignment to standard space. Structured noise components were further removed using the FIX denoising procedure [138], which regresses the motion-related confounding values estimated from rigid-body parameters and their temporal derivatives [139]. Temporal high-pass filtering was applied prior to regression. To maximize temporal coverage, RL and LR runs were concatenated after removing the mean signal across time.

The functional time series were divided into 116 regions of interest (ROIs), combining the 100 cortical parcels of the Schaefer atlas [140] with the 16 subcortical regions defined by the Tian atlas [141] (see Appendix 5.C). The final dataset comprised 100 multivariate time series with 116 dimensions and 2400 time points.

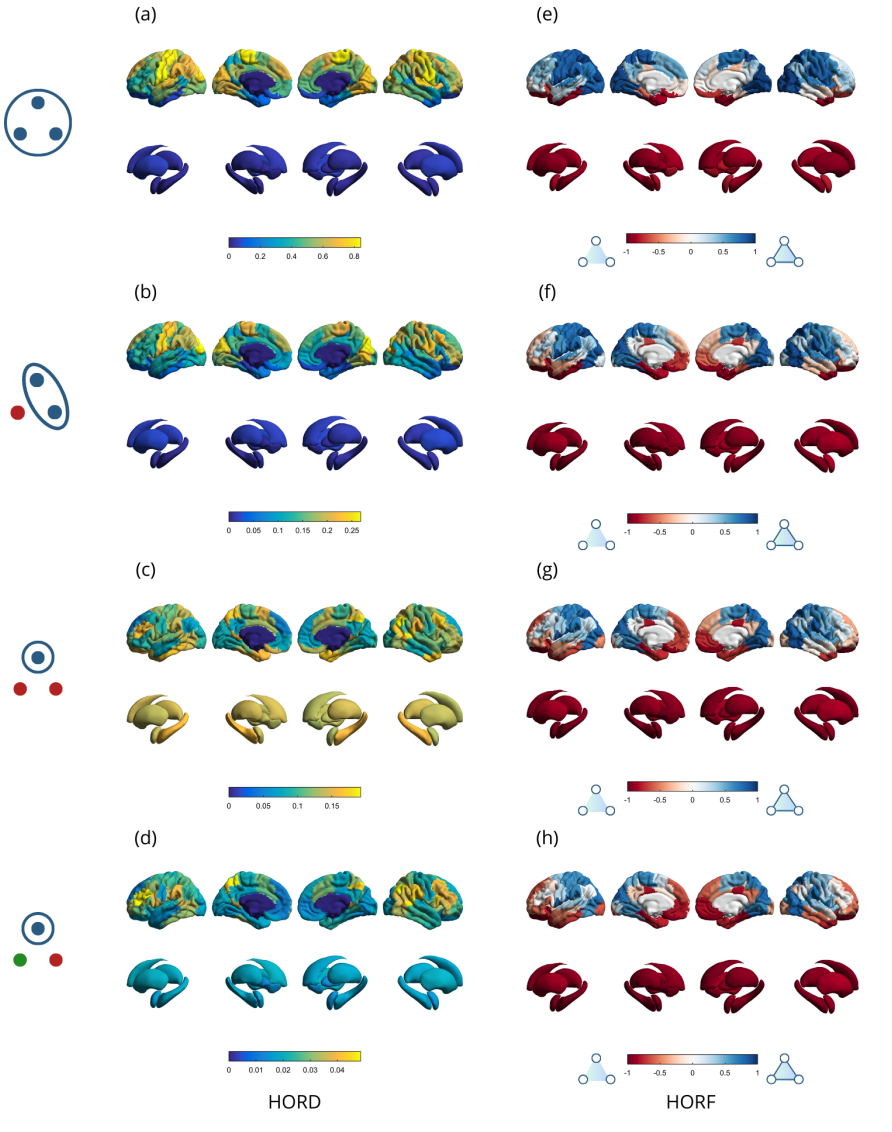


Figure 5.7: ***Regional density and connectedness of triplet configurations.*** *Panels (a–d) and (e–h) show cortical and subcortical maps of regional density and regional connectedness, respectively, for different classes of validated triplets. Each row corresponds to a distinct triplet configuration (intra-RSN, two-RSN majority, two-RSN minority, and three-RSN triplets). Panels (a–d) report regional density maps (yellow–blue scale), quantified by the higher-order regional density index (HORD), where warmer colors indicate more frequent participation in validated triplets, averaged across subjects. Panels (e–h) report regional connectedness maps (red–white–blue scale), quantified by the higher-order regional flexibility index (HORF), with blue (red) indicating a relative predominance of connected (unconnected) triplets, averaged across subjects. Notably, limbic and subcortical regions show increasing relative involvement in triplets spanning multiple RSNs and a relative predominance of unconnected configurations. In contrast, somatomotor and visual regions are more prominently associated with connected triplets.*

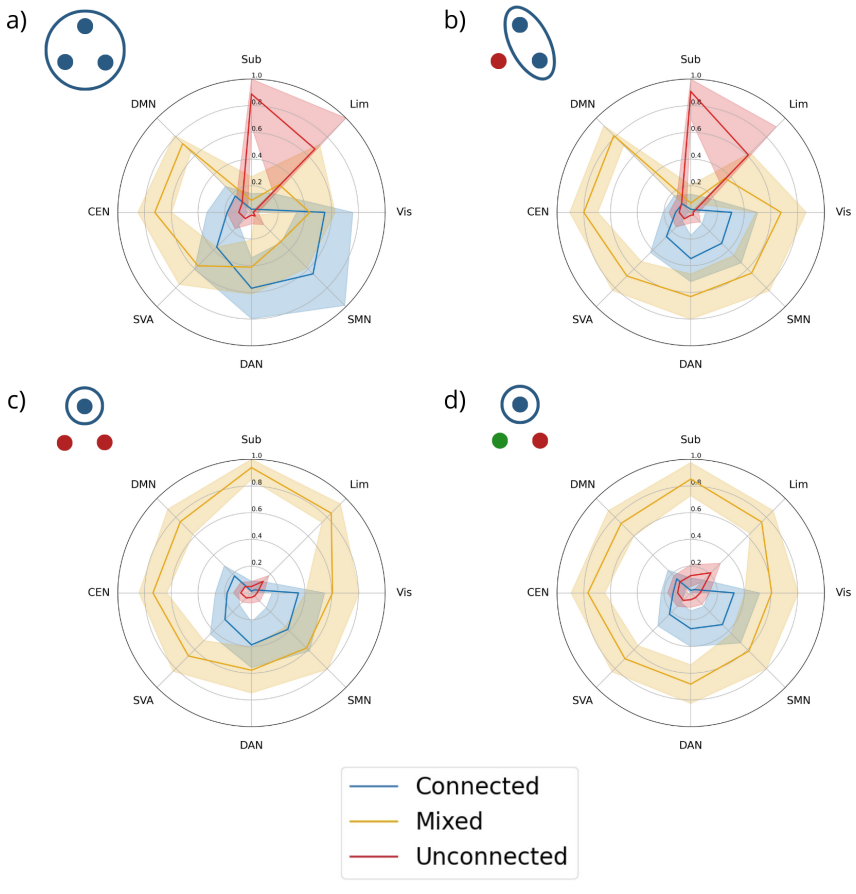


Figure 5.8: ***Mesoscale characterization of validated triplets across RSNs.*** *For each RSN R, validated triplets involving at least one node in R are classified into (a) intra-RSN triplets, (b) two-RSN (majority), (c) two-RSN triplets (minority), and (d) three-RSN triplets spanning three distinct RSNs. For each class and RSN, values are normalized to the total number of validated triplets. Shaded bands represent the standard deviation across subjects. This analysis reveals how connected, unconnected, and mixed triplets distribute across canonical functional systems, highlighting which RSNs are predominantly characterized by redundant lower-order dependencies versus genuinely higher-order interactions. A clear pattern emerges: sensorimotor and visual networks show limited involvement in unconnected triplets, whereas attentional and cognitive networks display a balance between connected and unconnected triplets, with a clear predominance of mixed triplets. At the same time, subcortical and limbic systems are often characterized by a higher prevalence of unconnected triplets than connected ones, underscoring their role as hubs of higher-order integration.*

5.3.2 Null models for validation

To assess whether the observed co-activations V_{ij}^* and V_{ijk}^* arise beyond expectations from random fluctuations, we employ canonical maximum-entropy null models, also known, in the context of network science, as *Exponential Random Graph Models*

[31, 4, 5] (see Appendix A). These models define probability ensembles of bipartite graphs that reproduce selected empirical properties of the data while randomizing all others, thereby providing rigorous statistical baselines. As anticipated, two null models are implemented in this work: the Bipartite Configuration Model for pairwise validation and the Bipartite Partial Local Overlap Model for triadic validation.

Pairwise validation: the BiCM

The BiCM [55] is a first–order[1] maximum–entropy model that preserves, in expectation, the degree sequence of both layers of the observed bipartite network $\mathbf{B}^*$. In formulas, it enforces the following constraints:

$$\mathbb{E}[k_i] = k_i^*, \qquad \forall\, i \in 1, \dots N, \tag{5.5}$$

$$\mathbb{E}[h_t] = h_t^*, \qquad \forall\, t \in 1, \dots T, \tag{5.6}$$

where $\mathbb{E}[x]$ represents the expected value of x in the chosen model, and k_i^* and h_t^* are the observed degrees of nodes i and t or, in other words, the empirical number of activations in region i and the number of active regions at time t. In this model, the entries b_{it} are *independent Bernoulli variables* by construction, with heterogeneous success probabilities p_{it} determined by imposing the above constraints. The probability of a bipartite configuration factorizes as

$$P(\mathbf{B}) = \prod_{i,t} p_{it}^{b_{it}} (1 - p_{it})^{1 - b_{it}}, \tag{5.7}$$

where

$$p_{it} = \frac{e^{-(\beta_i + \gamma_t)}}{1 + e^{-(\beta_i + \gamma_t)}}, \tag{5.8}$$

and β_i, γ_t are Lagrange multipliers estimated either by maximizing the likelihood of the model or by solving the constraint equations. This independence property implies that, for any pair of regions (i, j), the random variable $V_{ij} = \sum_t b_{it} b_{jt}$ is the sum of T independent Bernoulli trials with heterogeneous probabilities $p_{it} p_{jt}$, hence it follows a Poisson–Binomial distribution (see Appendix 5.B). Empirical values V_{ij}^* are compared with this distribution to compute p-values for pairwise validation. A detailed derivation of the Poisson-Binomial distribution and its computation can be found in [55].

Triadic validation: the BiPLOM

BiCM controls first-order heterogeneity by reproducing, in expectation, the activity level of each region and the number of active regions per time point. This allows the identification of pairwise interactions, quantities of the second order that cannot be explained solely by local differences in activation frequency. When moving to the validation of triplets, however, this model becomes insufficient: to test for genuine third–order interactions, one must account not only for local heterogeneity but also for

[1] Here, "first-order" indicates that the model constrains only quantities that are linear in the entries of $\mathbf{B}$.

the tendency of regions to co-activate at the pair level. For this reason, we introduce a second–order maximum–entropy model, the *Bipartite Partial Local Overlap Model* (BiPLOM), which constrains both first- and second-order statistics. By controlling for lower–order dependencies, BiPLOM provides a stricter statistical baseline, reducing the number of spurious triplets that would otherwise appear significant under the BiCM, as shown in the Results section. A full derivation of the BiPLOM and a detailed discussion of its properties are provided in Appendix 5.A.

For each region i, in addition to constraining its degree, as performed by the BiCM, the BiPLOM preserves the expected number of co-activations the region shares with others:

$$\langle V_i \rangle = \sum_{j \neq i} \sum_{t} p_{it} p_{jt} = V_i^* \qquad \forall i \in 1, \ldots, N. \tag{5.9}$$

This additional set of constraints ensures that the model reproduces the empirical tendency of each region to co-activate with others, effectively incorporating second–order correlations into the null ensemble. Unlike the BiCM, these additional constraints introduce dependencies among the variables b_{it}, making them non-independent Bernoulli variables. To obtain a tractable formulation, we adopt a mean–field approximation (see Appendix 5.A), which replaces the full joint distribution with a product of effective Bernoulli marginals. Under this approximation, each b_{it} becomes effectively independent, and the model regains a factorized form analogous to the BiCM, with success probabilities which read:

$$p_{it} = \frac{e^{-(\alpha_i + \beta_t + \delta_i h_t^*)}}{1 + e^{-(\alpha_i + \beta_t + \delta_i h_t^*)}}, \tag{5.10}$$

where the parameters α_i, β_t and δ_i are estimated by maximizing the likelihood. In the same mean-field approximation, the number of triplet co-activations

$$V_{ijk} = \sum_{t} b_{it} b_{jt} b_{kt} \tag{5.11}$$

is again the sum of T independent Bernoulli trials with heterogeneous probabilities $p_{it}p_{jt}p_{kt}$ and follows a Poisson–Binomial distribution. Comparison between empirical and expected values yields the p-values used for triadic validation.

Computation of p-values. Exact evaluation of Poisson–Binomial distributions is computationally demanding, especially when repeated for a large number of pairs, triplets, and subjects. For this reason, while the null distributions of V_{ij} and V_{ijk} are defined exactly by the corresponding Poisson–Binomial laws, in practice we rely on analytical approximations to compute p-values efficiently. A detailed comparison of different approximations and their performance on our dataset is provided in Appendix 5.B. Based on this analysis, we adopt the refined normal approximation throughout this work, which offers an excellent compromise between computational efficiency and accuracy.

5.3.3 Multiple hypothesis testing

The validation of dyadic and triadic interactions involves testing a large number of statistical hypotheses simultaneously. Specifically, for each subject, hypotheses are tested for all pairs and all triplets of brain regions, resulting in thousands of tests at the dyadic level and millions at the triadic level. To control for the inflation of false positives arising from multiple comparisons, we adopt a False Discovery Rate (FDR) control procedure [132].

Given a set of $|H|$ hypotheses $\{H_1, H_2, \ldots, H_{|H|}\}$, each associated with a p-value, the Benjamini–Hochberg procedure prescribes sorting the p-values in non-decreasing order,

$$\text{p-value}_1 \leq \cdots \leq \text{p-value}_{|H|}, \tag{5.12}$$

and identifying the largest index $\hat{i}$ such that

$$\text{p-value}_{\hat{i}} \leq \frac{\hat{i}}{|H|}\, t, \tag{5.13}$$

where t denotes the nominal single-test significance level, set here to $t = 0.05$.

All hypotheses H_i with $i \leq \hat{i}$ are then rejected, ensuring control of the expected proportion of false discoveries among the rejected hypotheses. In our context, each hypothesis tests whether the empirical number of co-activations for a given pair or triplet of regions is consistent with the distribution predicted by the corresponding null model (BiCM or BiPLOM).

Rejecting a dyadic hypothesis amounts to retaining the corresponding pair of regions as a statistically validated link in the monopartite projection [55]. Analogously, rejecting a triadic hypothesis results in the validation of the corresponding triplet as a higher-order interaction. FDR correction is applied separately to dyadic and triadic hypotheses, ensuring control of the FDR within each family of tests.

5.3.4 Classification and characterization of validated triplets

To characterize the validated triplets identified by the application of BiPLOM, we distinguished them based on the number of their constituent pairs validated by BiCM. Triplets in which all three pairwise interactions were validated by the BiCM were labeled as *connected triplets*, whereas triplets for which none of the three pairs were validated were labeled *unconnected triplets*. These two extreme cases provide the clearest contrast between higher-order structure strongly supported by pairwise statistics and structure that is essentially "invisible" at the pairwise level. Intermediate cases in which exactly one or two edges were validated are also possible and were labeled *mixed triplets* (see Fig. 5.4a for a schematic overview of all configurations).

Global geometric characterization of validated triplets

Having defined connected, mixed, and unconnected triplets, we next examined whether these classes differed in their spatial footprint. To this end, we associated each

ROI with a centroid in standard space, using the combined Schaefer–Tian atlas aligned to the MNI152NLin2009cAsym space, which removes inter-subject differences in overall brain size. For every validated triplet (i, j, k), with centroid coordinates $\mathbf{x}_i, \mathbf{x}_j, \mathbf{x}_k \in \mathbb{R}^3$, we computed the area of the triangle between the three centroids,

$$A_{ijk} = \frac{1}{2} \left\| (\mathbf{x}_j - \mathbf{x}_i) \times (\mathbf{x}_k - \mathbf{x}_i) \right\|, \tag{5.14}$$

which provides a simple Euclidean distance-based geometric measure of how spatially extended each triplet is. This also allows the use of the same measure for cortical (both ipsilateral and contralateral) and subcortical distances, whereas geodesic distances would not.

O-information analysis of higher-order dependencies

To complement the geometric characterization with an information-theoretic perspective, we quantified the nature of higher-order statistical dependencies within validated triplets using the O-information [106]. In this work, the O-information is computed on the original fMRI time series associated with each validated triplet.

Specifically, for each validated triplet (i, j, k) and for each subject, we considered the corresponding time series array

$$\mathbf{X}_{lt} = \{x_l(t)\}_{\substack{l \in \{i,j,k\} \\ t \in \{1,\dots,T\}}}, \tag{5.15}$$

where $x_l(t)$ denotes the fMRI signal of region l at time t. The time index t is treated as indexing repeated observations of the joint random variable (X_i, X_j, X_k), allowing the estimation of the associated entropic quantities from the empirical joint distribution of the signals.

The O-information $\Omega(\mathbf{X}_{lt})$ is a signed measure that quantifies whether the joint interaction among the three time series is dominated by lower-order, redundant dependencies (positive values) or by synergistic contributions (negative values). For three variables, it can be expressed as

$$\Omega(X_i, X_j, X_k) = TC(X_i, X_j, X_k) - DTC(X_i, X_j, X_k), \tag{5.16}$$

where TC and DTC denote the Total Correlation and Dual Total Correlation, respectively. The Total Correlation is defined as

$$TC(X_i, X_j, X_k) = H(X_i) + H(X_j) + H(X_k) - H(X_i, X_j, X_k), \tag{5.17}$$

while the Dual Total Correlation reads

$$\begin{aligned} DTC(X_i, X_j, X_k) = H(X_i, X_j, X_k) \; &- \big[H(X_i \mid X_j, X_k) \\ &+ H(X_j \mid X_i, X_k) + H(X_k \mid X_i, X_j)\big], \end{aligned} \tag{5.18}$$

with $H(X)$ denoting Shannon entropy of variable X and $H(X \mid Y)$ denoting the conditional entropy of X with respect to Y.

For each subject, $\Omega(\mathbf{X}_{lt})$ was computed for all validated triplets and subsequently averaged within each class (connected, mixed, and unconnected), yielding subject-level estimates of the prevailing informational regime associated with each class.

Triplet configuration types and higher-order regional density

To relate validated higher-order interactions to their spatial embedding across the cortex and subcortex, we further classified triplets according to the distribution of their constituent regions across resting-state networks (RSNs). We distinguished three main configuration types: *intra-RSN* triplets, in which all three regions belong to the same RSN; *two-RSN* triplets, involving exactly two distinct RSNs; and *three-RSN* triplets, spanning three different RSNs. For two-RSN triplets, we additionally distinguished between *majority* and *minority* configurations with respect to a given RSN R, depending on whether R contributed two or only one region to the triplet (see Fig. 5.4b for a schematic overview).

This classification allows the characterization of individual brain regions by how often they participate in higher-order motifs associated with different RSN-based configurations. Specifically, for each subject s, region r, and configuration type c (e.g., intra-RSN, two-RSN majority, two-RSN minority, three-RSN), we counted the number of validated triplets of type c that include region r, denoted by $n_s(r, c)$.

To enable comparison across subjects, these counts were normalized within each subject and configuration, defining the higher-order regional density (HORD) as

$$HORD_s(r, c) = \frac{n_s(r, c)}{\sum_{r'} n_s(r', c)}. \tag{5.19}$$

Accordingly, $HORD_s(r, c)$ represents the fraction of validated triplets of configuration c that involve region r for subject s. Group-level regional density maps were then obtained by averaging $HORD_s(r, c)$ across subjects.

Regional connectedness index

In addition to quantifying how frequently regions participate in different triplet configurations, we assessed whether regions tend to be involved preferentially in triplets supported by validated pairwise interactions or in triplets that emerge only at the higher-order level. To this end, we retained the same triplet classification described above and further distinguished, for each subject s, region r, and configuration c, how many of the contributing triplets were *connected* or *unconnected*. We denote these counts by $n_s^{\text{conn}}(r, c)$ and $n_s^{\text{unconn}}(r, c)$, respectively.

We then defined a regional connectedness index, termed higher-order regional flexibility (HORF), which contrasts these two contributions:

$$HORF_s(r, c) = \frac{n_s^{\text{conn}}(r, c) - n_s^{\text{unconn}}(r, c)}{n_s^{\text{conn}}(r, c) + n_s^{\text{unconn}}(r, c)}. \tag{5.20}$$

This index takes values in the interval $[-1, 1]$. Positive values indicate that region r participates more frequently in connected triplets of type c, whereas negative values indicate a predominance of unconnected triplets. Values close to zero reflect a balanced involvement in both types of higher-order interactions. The index is defined only for regions participating in at least one validated triplet of configuration c.

For each configuration type, $HORF_s(r, c)$ was averaged across subjects to obtain group-level maps that summarize the relative involvement of each region in lower-order-supported versus genuinely higher-order interaction patterns.

To aggregate these results at the RSN level, we considered, for each RSN R, all triplets including R. We then computed the proportion of connected, mixed, and unconnected triplets with a given configuration, normalizing by the total number of validated triplets. Fig. 5.8 summarizes the aggregated results, showing whether interactions between and within specific RSNs are better described by connected, mixed, or unconnected triplets.

5.4 Discussion and conclusion

Understanding how brain regions coordinate their activity beyond pairwise interactions is a central challenge in network neuroscience. While pairwise functional connectivity has provided valuable insights into large-scale brain organization, it is increasingly recognized that collective interactions involving more than two regions may capture additional aspects of neural coordination that are not evident at the dyadic level. In this work, we addressed this problem by introducing a statistically grounded framework to identify and characterize higher-order co-activation patterns directly from multivariate brain time series.

By representing binarized fMRI signals as a bipartite network between brain regions and time points, and by validating co-activation motifs through maximum-entropy null models, we obtained a principled mapping from temporal data to statistically validated higher-order interactions. Crucially, the use of null models that explicitly control for lower-order statistics allows us to distinguish between triadic co-activations that are consistent with pairwise structure and those that exceed it. This separation provides an operational definition of higher-order interactions in statistical terms, independent of any *a priori* functional interpretation.

The resulting validated triplets reveal a non-trivial organization of higher-order co-activation patterns in resting-state brain activity. Rather than forming a homogeneous class, triadic interactions span a spectrum of statistical configurations, reflecting different degrees of support from lower-order dependencies. This motivates a systematic comparison of triplets based on their statistical, spatial, and informational properties.

A central outcome of our analysis is the emergence of distinct regimes of triadic coordination, defined by the extent to which higher-order co-activations are supported by validated pairwise interactions. *Connected* triplets are fully supported by dyadic structure, indicating that their higher-order organization can be largely explained in terms of redundant lower-order dependencies. *Unconnected* triplets lack any statistically validated pairwise support and therefore constitute genuinely higher-order interactions in a strict statistical sense. In addition, a substantial fraction of validated triplets falls into an intermediate *mixed* category, in which only a subset of the underlying pairs is statistically supported. Importantly, these regimes are not defined heuristically, but arise directly from comparisons between null models encoding different levels of constraints. An important methodological feature of our analysis is

that unconnected and mixed triplets do not satisfy the simplicial closure condition. As a consequence, they would remain undetected in frameworks based on simplicial complexes or clique expansions of pairwise networks [113, 114].

These different regimes exhibit systematic differences in both spatial organization and informational signatures. Connected triplets tend to be spatially compact and are associated with positive, large O-information values, consistent with interaction patterns dominated by redundant, lower-order dependencies. Unconnected triplets, in contrast, are typically more spatially extended and display attenuated or near-zero O-information values. Consistently, mixed triplets occupy an intermediate regime in O-information values, reflecting the coexistence of lower-order and higher-order interdependencies within the same motif. However, the spatial extension of mixed triplets is compatible with that of unconnected ones, showing that both categories help the integration of information across the brain.

At the mesoscale, these higher-order interaction regimes exhibit structured relationships with the organization of resting-state networks (RSNs). Across RSNs, connected, unconnected, and mixed triplets are observed in all configuration types (intra-, two-, and three-RSN), but their relative prevalence varies systematically across functional systems. Unimodal networks, such as visual and somatomotor systems, are predominantly associated with connected and within-RSN triplets, consistent with their tightly coupled, locally and functionally redundant dynamics [32].

In contrast, limbic and subcortical areas display a marked predominance of unconnected and mixed configurations. These areas are primarily involved in information routing across RSN and integrating, relaying, and modulating signals between cortical areas [142, 143, 144]. Our findings suggest that these functions are predominantly supported by spatially sparse, purely higher-order interactions. Notably, these interactions favor integration between subcortical/limbic regions and functionally diverse areas, rather than forming isolated clusters within the limbic or subcortical systems themselves.

A particularly coherent mesoscale profile is observed for Control- and association-related networks, most notably the Central Executive Network (CEN), the Default Mode Network (DMN), and the Salience/Ventral Attention Network (SVA), where mixed triplets dominate in number over both connected and unconnected ones across all the types of configurations: within-RSN, two-RSN, and three-RSN. This shared signature suggests that their functional interplay relies on a balance between lower-order supported interactions and genuinely higher-order coordination.

This is consistent with the *Triple Network Model* [145, 146, 147], which conceptualizes these systems as a tightly integrated control ensemble supporting adaptive cognition. Our results extend this framework by showing that interactions involving these networks are supported by statistically validated higher-order co-activation patterns that cannot be fully reduced to dyadic structures.

More generally, the prominence of higher-order configurations, especially in cross-network triplets, highlights the role of higher-order dependencies as a flexible substrate for integration across functional systems. These interactions capture coordination patterns that cannot be reduced to purely dyadic ones, yet recur robustly across subjects and spatial scales. Our results, therefore, identify a class of statistically

validated higher-order interactions that are invisible to simplicial representations by construction, despite playing a non-marginal role in the empirical organization of brain activity. This finding suggests that restricting the analysis to cliques or simplices may overlook important dynamics such as those related to large-scale functional integration.

Beyond methodological considerations, the observed distinction establishes a statistically grounded link between topological and information-theoretic descriptions of multivariate interactions. By combining topological validation with the O-information, we show that triplets lacking full pairwise support are not only non-simplicial from a structural standpoint but are also associated with distinct informational regimes, characterized by near-zero redundancy. Consistently, connected triplets exhibit a markedly higher level of redundancy. Crucially, this correspondence does not follow from the construction of the analysis: despite being obtained through independent procedures, topological validation on binarized data and information-theoretic quantification on the full time series, the resulting distinction persists, indicating that topological and informational notions of higher-order interaction converge at the level of empirically validated brain dynamics.

Despite the novelty and robustness of the methods and the reported findings, some limitations must be acknowledged. First, the present formulation relies on binary and unsigned representations of the data obtained through thresholding. Although this choice facilitates the analysis, it inevitably removes information about the magnitude and polarity of fluctuations, which may still be relevant, especially in the context of higher-order synergistic interactions. Recent extensions of pairwise validation techniques to signed bipartite networks [127] and to signed brain time series [128] provide promising avenues for future research. Generalizing these approaches to the higher-order setting discussed here could enable the explicit inclusion of sign and weight information in the validation process.

Second, the current framework remains invariant under temporal reshuffling. Although rs-fMRI dynamics are typically modeled under stationarity assumptions, such simplifications inherently overlook the full complexity of brain activity over time. While some aspects of temporal heterogeneity are indirectly captured via degree-based constraints, the explicit temporal ordering of events is ignored. As a result, the method cannot account for time-dependent correlations or causal interactions. Future developments could address this by introducing second-order null models that retain temporal structure or constrain sequential motifs, thereby enabling a more direct investigation of co-activation dynamics.

Third, although computationally efficient formulations, such as canonical models and mean-field approximations, have been introduced, the scalability of higher-order validation remains a practical challenge, especially when testing all possible triplets or larger combinations. Optimizing sampling procedures or developing analytical approximations for motif distributions may be necessary to ensure applicability to large-scale datasets.

Despite the methodological limitations discussed above, our results demonstrate that statistically validated higher-order interactions provide a meaningful and interpretable level of description for brain activity that complements and extends pairwise

connectivity analyses. By revealing robust triadic coordination patterns that cannot be reduced to dyadic structure, this work shows that collective neural organization cannot be fully captured within purely pairwise or simplicial frameworks.

More broadly, the proposed approach illustrates how null-model–based validation can be used to define higher-order interactions in operational statistical terms, enabling principled comparisons between structural, spatial, and informational signatures of multivariate coordination. In this sense, higher-order co-activation patterns emerge not merely as descriptive constructs, but as statistically grounded objects that can inform our understanding of large-scale integration in complex systems, in the brain and beyond.

5.5 Data availability

Data concerning rs-fMRI from the HCP can be found on the Human Connectome Project website; for a detailed description, see the HCP S1200 Release Reference Manual.

Code availability

The Python package named `TRIDENT` (*TRIplets DEtectioN via staTistical validation*), implementing the algorithms described in the main text, is available on PyPI and at the URL `https://github.com/mattiamarzi/TRIDENT`.

Appendix 5.A Canonical null models for hypothesis testing

The canonical null models employed in this work belong to the general family of Exponential Random Graph Models (ERGMs) [4, 148, 5, 6]. Such models can be derived as maximum-entropy ensembles under *soft constraints*, that is, by requiring that selected topological quantities match their observed expectations while maximizing randomness in every other respect. In this work, we are interested in implementing null models for undirected, binary, and bipartite networks. As in [117], we follow the analytical approach described in [148] and further developed in [31]. This approach rests upon an inference procedure allowing us to find a probability distribution which is maximally non-committal with respect to the missing data; quantitatively, it is based on the maximization of the so called *Shannon entropy* $S[P] = -\sum_{\mathbf{B}\in\mathbb{B}} P(\mathbf{B}) \ln P(\mathbf{B})$, where $\mathbb{B}$ is the set of all undirected, binary, bipartite graphs having N nodes in one layer and T nodes in the other, and $P(\mathbf{B})$ is a probability distribution defined over $\mathbb{B}$. The entropy maximization is carried out while constraining the expected values of a chosen vector of proper topological quantities, the constraints vector $\mathbf{c}(\mathbf{B})$. Solving such a maximization problem leads us to find the exponential probability distribution:

$$P(\mathbf{B};\boldsymbol{\theta}) = \frac{e^{-H(\mathbf{B},\boldsymbol{\theta})}}{Z(\boldsymbol{\theta})} = \frac{e^{-\boldsymbol{\theta}\cdot\mathbf{c}(\mathbf{B})}}{Z(\boldsymbol{\theta})}, \tag{5.21}$$

where $\boldsymbol{\theta}$ is the vector of Lagrangian multipliers associated with the vector of constraints $\mathbf{c}(\mathbf{B})$, $H(\mathbf{B},\boldsymbol{\theta}) = \boldsymbol{\theta}\cdot\mathbf{c}(\mathbf{B})$ is the *Hamiltonian* of the model and $Z(\boldsymbol{\theta}) =$

$\sum_{\mathbf{B}\in\mathbb{B}} e^{-\boldsymbol{\theta}\cdot\mathbf{c}(\mathbf{B})}$ is the *partition function* ensuring normalization. To tune the parameters defining the entropy-based null model, we require that the expected values of the constrained quantities $\mathbb{E}[\mathbf{c}(\mathbf{B})]$ equate the corresponding values $\mathbf{c}^* = \mathbf{c}(\mathbf{B}^*)$ computed on the observed data $\mathbf{B}$. Interestingly, it can be shown that in this framework the maximum-likelihood estimation of the Lagrangian multipliers leads precisely to satisfy the condition $\mathbb{E}[\mathbf{c}(\mathbf{B})] = \mathbf{c}(\mathbf{B}^*)$ [149].

In what follows, we focus on models in which the probability of each link can be written as an independent Bernoulli random variable. This assumption holds exactly when the imposed constraints are linear functions of the adjacency matrix, as in the case of node degrees (see the BiCM below). When non-linear quantities are constrained, independence no longer holds exactly; in such cases, we adopt a mean-field approximation, as discussed later for the BiPLOM. Crucially, the independent-link formulation allows the distribution of the number of shared neighbors for any n-tuple of nodes to be expressed as a Poisson–Binomial law, providing a tractable and analytically consistent basis for hypothesis testing. From a modeling perspective, this independence structure is essential: it ensures that the significance of higher-order co-occurrences can be computed analytically rather than estimated by computationally expensive simulations. While here we apply these models to validate interactions between time series, the proposed framework is far more general: it applies to all binary, two-mode datasets that can be represented as bipartite networks. For this reason, in this section, while retaining the same notation as in the main text, we do not explicitly talk about brain or time nodes, but, respectively, the bottom and top layers.

5.A.1 Bipartite Configuration Model

The Bipartite Configuration Model [55] is a linear ERGM for binary, bipartite networks, defined by two sets of local constraints, the degree sequences $\boldsymbol{k}(\mathbf{B}) = (k_1(\mathbf{B}), \cdots, k_N(\mathbf{B}))$ and $\boldsymbol{h}(\mathbf{B}) = (h_1(\mathbf{B}), \cdots, h_T(\mathbf{B}))$. Its Hamiltonian reads

$$H(\mathbf{B};\boldsymbol{\beta},\boldsymbol{\gamma}) = \sum_{i=1}^{N}\beta_i k_i(\mathbf{B}) + \sum_{t=1}^{T}\gamma_t h_t(\mathbf{B}) = \sum_{i=1}^{N}\sum_{t=1}^{T}(\beta_i+\gamma_t)b_{it}. \tag{5.22}$$

where the N-dimensional vector $\boldsymbol{\beta} = (\beta_1, \cdots \beta_N)$ and the T-dimensional vector $\boldsymbol{\gamma} = (\gamma_1, \cdots, \gamma_T)$ are the parameter vectors corresponding to the top and bottom degrees. The partition function can be computed analytically, and the probability distribution reads

$$P_{\text{BiCM}}(\mathbf{B};\boldsymbol{\beta},\boldsymbol{\gamma}) = \frac{e^{-\sum_{i=1}^{N}\sum_{t=1}^{T}(\beta_i+\gamma_t)b_{it}}}{\prod_{i=1}^{N}\prod_{t=1}^{T}(1+e^{-\beta_i-\gamma_t})} = \prod_{i=1}^{N}\prod_{t=1}^{T} p_{it}^{b_{it}}(1-p_{it})^{1-b_{it}}, \tag{5.23}$$

where the parameters p_{it} in the above expression are defined as

$$p_{it} \equiv \frac{e^{-\beta_i-\gamma_t}}{1+e^{-\beta_i-\gamma_t}} \qquad \forall i,t. \tag{5.24}$$

The formula above is often written in terms of the exponential parameters $\boldsymbol{x}, \boldsymbol{y}$, obtained by posing $x_i = e^{-\beta_i}$ and $y_t = e^{-\gamma_t}$ $\forall i, t$:

$$p_{it} \equiv \frac{x_i y_t}{1 + x_i y_t} \quad \forall i, t, \tag{5.25}$$

p_{it} represents the probability that nodes i and t are linked. Thus, under the BiCM, each entry b_{it} of a biadjacency matrix $\mathbf{B} \in \mathbb{B}$ is distributed as a Bernoulli random variable with probability coefficient p_{it}. In turn, this implies that $\mathbf{B}$ is a collection of $N \times T$ independent, non-identically distributed Bernoulli variables.

As mentioned, to tune the parameters defining an ERGM, we can either maximize the likelihood function or solve the constraint equations. We now explicitly show this, as it underscores the differences with the BiPLOM (defined later), where this equivalence is partially lost. The BiCM log-likelihood function reads

$$\begin{aligned}
\mathcal{L}_{\text{BiCM}}(\boldsymbol{\beta}, \boldsymbol{\gamma}) &\equiv \ln P_{\text{BiCM}}(\mathbf{B}^*; \boldsymbol{\beta}_i, \boldsymbol{\gamma}_t) \\
&= \sum_{i=1}^{N} b_{it}^* \ln \left(\frac{e^{-\beta_i - \gamma_t}}{1 + e^{-\beta_i - \gamma_t}} \right) - \sum_{i=1}^{N} \sum_{t=1}^{T} (1 - b_{it}^*) \ln(1 + e^{-\beta_i - \gamma_t}) \\
&= \sum_{i=1}^{N} b_{it}^* \ln(e^{-\beta_i - \gamma_t}) - \sum_{i=1}^{N} \sum_{t=1}^{T} \ln(1 + e^{-\beta_i - \gamma_t}) \\
&= - \sum_{i=1}^{N} \sum_{t=1}^{T} b_{it}^* (\beta_i + \gamma_t) - \sum_{i=1}^{N} \sum_{t=1}^{T} \ln(1 + e^{-\beta_i - \gamma_t}).
\end{aligned} \tag{5.26}$$

In order to maximize it with respect to $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, we need to solve the system of equations

$$\frac{\partial \mathcal{L}_{\text{BiCM}}(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \beta_i} = - \sum_{t=1}^{T} b_{it}^* + \sum_{t=1}^{T} \frac{e^{-\beta_i - \gamma_t}}{1 + e^{-\beta_i - \gamma_t}} \quad \forall i, \tag{5.27}$$

$$\frac{\partial \mathcal{L}_{\text{BiCM}}(\boldsymbol{\beta}, \boldsymbol{\gamma})}{\partial \gamma_t} = - \sum_{i=1}^{N} b_{it}^* + \sum_{i=1}^{N} \frac{e^{-\beta_i - \gamma_t}}{1 + e^{-\beta_i - \gamma_t}} \quad \forall t, \tag{5.28}$$

and, upon equating them to zero, we find

$$k_i^* = \sum_{t=1}^{T} \frac{e^{-\beta_i - \gamma_t}}{1 + e^{-\beta_i - \gamma_t}} = \sum_{t=1}^{T} p_{it} = \mathbb{E}[k_i] \quad \forall i, \tag{5.29}$$

$$h_t^* = \sum_{i=1}^{N} \frac{e^{-\beta_i - \gamma_t}}{1 + e^{-\beta_i - \gamma_t}} = \sum_{i=1}^{N} p_{it} = \mathbb{E}[h_t] \quad \forall t. \tag{5.30}$$

The BiCM thus provides a first-order baseline that fully accounts for heterogeneity in node activation patterns but not for correlations between nodes. As discussed in the main text, this limitation becomes crucial when validating structures of order three

or higher. Notice that the system above can be solved only numerically. A ready-to-use implementation of the Bipartite Configuration Model, along with several other Maximum Entropy models, is available in the open-source Python package NEMtropy [150].

5.A.2 From first to second-order null models

As explained in the main text, in order to validate triplets, first-order constraints might be insufficient. This motivates the introduction of second-order null models in the validation pipeline described in this work. The main limitation of introducing second-order constraints in a canonical model is the loss of link independence, and consequently of the analytical simplicity that stems from it, an aspect that is central to the validation framework introduced by Saracco et al. [55] and that we aim to extend here. In addition, second-order models are typically not analytically solvable.

To overcome these issues, a mean-field approach is employed [148, 151, 152]. The mean-field approximation builds on the concept of *separable Hamiltonians*, in which the Hamiltonian is expressed as a sum of local contributions that may depend on global or mesoscopic properties of the network. In contrast to linear ERGMs, whose Hamiltonians take the general form

$$H(\mathbf{B}, \boldsymbol{\theta}) = \sum_{i=1}^{N} \sum_{t=1}^{T} \theta_{it}\, b_{it}, \tag{5.31}$$

and yield independent link probabilities, separable Hamiltonians adopt a more general expression:

$$H(\mathbf{B}, \boldsymbol{\theta}) = \sum_{i=1}^{N} \sum_{t=1}^{T} b_{it}\, f_{it}(\mathbf{B}, \boldsymbol{\theta}), \tag{5.32}$$

where the functions f_{it} introduce non-linear dependencies (e.g., on k_i, h_t, or motif statistics), possibly involving all entries of $\mathbf{B}$.

The mean-field approximation consists of replacing the adjacency matrix $\mathbf{B}$ with its expectation $\mathbf{P} = \mathbb{E}[\mathbf{B}]$, i.e., the matrix of link probabilities, inside the functional terms f_{it}. Under this approximation, the expression becomes separable, and links become independent:

$$P(\mathbf{B}) = \prod_{i=1}^{N} \prod_{t=1}^{T} p_{it}^{b_{it}} (1 - p_{it})^{1 - b_{it}}, \tag{5.33}$$

where now

$$p_{it} = \frac{e^{-f_{it}(\mathbf{P}, \boldsymbol{\theta})}}{1 + e^{-f_{it}(\mathbf{P}, \boldsymbol{\theta})}}. \tag{5.34}$$

This approximation restores independence at the cost of only approximate constraint satisfaction, which is the price paid to maintain analytic tractability.

It is important to note that, once mean-field is applied, solving the constraint equations and maximizing the likelihood are no longer guaranteed to be equivalent. In linear ERGMs (such as the BiCM), the two procedures coincide, but in non-linear models, the mean-field approximation effectively modifies the model class. This subtlety will play a key role in the construction of the BiPLOM in the next section.

5.A.3 The Bipartite Partial Local Overlap Model

Here, we introduce the Bipartite Partial Local Overlap Model, a second-order ERGM which we solve in the mean-field approximation.

Ideally, in order to validate co-occurrences between triplets, a model designed to reproduce, beyond degree sequences, local second-order statistics is preferred. A direct way to impose second-order information would be to constrain all pairwise overlaps V_{ij}. However, since there are $O(N^2)$ of them, this approach is computationally prohibitive and would make the model intractable. For this reason, we seek summary statistics that encode local second-order structure while keeping the number of constraints linear in N. To this aim, for each node i in the bottom layer, define the aggregate overlap

$$V_i = \sum_{t=1}^{T} \sum_{\substack{j=1 \\ (j\neq i)}}^{N} b_{it} b_{jt} = \sum_{t=1}^{T} b_{it}(h_t - b_{it}) \qquad \forall i. \tag{5.35}$$

V_i counts the total number of V_{ij} motifs in which node i participates, or equivalently, the total number of bottom-layer nodes that share at least one neighbor with i. By constraining both degree sequences and the V_i patterns, the resulting null model accounts for (i) the local tendency of each node to form links (through degrees) and (ii) its local tendency to share neighbors with others. This provides a suitable baseline for statistically validating triplets whose co-occurrences cannot be explained by lower-order tendencies. Such constraints induce the Hamiltonian

$$H(\mathbf{B}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}) = \sum_{i=1}^{N} \beta_i k_i(\mathbf{B}) + \sum_{t=1}^{T} \gamma_t h_t(\mathbf{B}) + \sum_{i=1}^{N} \delta_i V_i(\mathbf{B}) \tag{5.36}$$

$$= \sum_{i=1}^{N} \sum_{t=1}^{T} \beta_i b_{it} + \sum_{t=1}^{T} \sum_{i=1}^{N} \gamma_t b_{it} + \sum_{i=1}^{N} \sum_{t=1}^{T} \delta_i b_{it}(h_t(\mathbf{B}) - b_{it}) \tag{5.37}$$

$$= \sum_{i=1}^{N} \sum_{t=1}^{T} \left[\beta_i + \gamma_t + \delta_i \big(h_t(\mathbf{B}) - b_{it}\big) \right] b_{it}. \tag{5.38}$$

The mean field approximation requires replacing $\left[\beta_i + \gamma_t + \delta_i(h_t(\mathbf{B}) - b_{it})\right]$ with its expectation under the model, $\left[\beta_i + \gamma_t + \delta_i(\mathbb{E}[h_t] - p_{it})\right]$. The induced mean-field link probabilities read

$$p_{it} = \frac{e^{-\beta_i - \gamma_t - \delta_i(\mathbb{E}[h_t] - p_{it})}}{1 + e^{-\beta_i - \gamma_t - \delta_i(\mathbb{E}[h_t] - p_{it})}}. \tag{5.39}$$

Upon assuming that the constraints on the degrees are met, one gets:

$$p_{it} = \frac{e^{-\beta_i - \gamma_t - \delta_i(h_t^* - p_{it})}}{1 + e^{-\beta_i - \gamma_t - \delta_i(h_t^* - p_{it})}}. \tag{5.40}$$

However, it is easily shown that the solution to the constraint equation is degenerate

in the mean-field approximation. In fact, for each i the expectation of V_i reads:

$$\begin{aligned}\mathbb{E}[V_i] &= \mathbb{E}\left[\sum_{t=1}^{T}\sum_{\substack{j=1\\j\neq i}} b_{it}b_{jt}\right] = \sum_{t=1}^{T}\sum_{\substack{j=1\\j\neq i}} p_{it}p_{jt} \\ &= \sum_{t=1}^{T}\sum_{j=1}^{N} p_{it}p_{jt} - \sum_{t=1}^{T} p_{it}^2 \qquad (5.41)\\ &= \sum_{t=1}^{T} p_{it}(h_t^* - 1) + \sum_{t=1}^{T} p_{it}(1-p_{it}) \qquad (5.42)\\ &= \sum_{t=1}^{T} p_{it}(h_t^* - 1) + \mathrm{Var}[k_i], \qquad (5.43)\end{aligned}$$

where we assumed again that the constraints on the degrees are met. Requiring $\mathbb{E}[V_i] = V_i^*$ then implies

$$\sum_{t=1}^{T} p_{it}(h_t^* - 1) + \mathrm{Var}[k_i] = \sum_{t=1}^{T} b_{it}^*(h_t^* - 1) \qquad \forall i. \tag{5.44}$$

Summing both sides over i, one gets:

$$\sum_{t=1}^{T} h_t^*(h_t^*-1) + \sum_{i=1}^{N}\mathrm{Var}[k_i] = \sum_{t=1}^{T} h_t^*(h_t^*-1) \implies \sum_{i=1}^{N}\mathrm{Var}[k_i] = 0 \implies \mathrm{Var}[k_i] = 0\ \forall i. \tag{5.45}$$

Moreover, in the independent-links approximation,

$$\mathrm{Var}[k_i] = \sum_{t=1}^{T} p_{it}(1-p_{it}) = 0 \qquad \forall i, \tag{5.46}$$

which is obtained if and only if $p_{it} \in \{0, 1\}$. Thus, the model collapses to a deterministic configuration: no canonical ensemble exists.

Since exact constraint satisfaction is impossible, one may attempt to maximize likelihood. However, in this case, the expression for p_{it} in (5.40) does not yield a closed form depending only on model parameters and observed data.

To solve this problem, we introduce a modified version of V_i, including the *self term* $j = i$. Formally, we constrain, together with the degrees, the quantities

$$\widetilde{V}_i = \sum_{t=1}^{T}\sum_{j=1}^{N} b_{it}b_{jt} = \sum_{t=1}^{T} b_{it}\, h_t \qquad \forall i. \tag{5.47}$$

This modification does not remove the degeneracy that would arise from enforcing the constraints exactly: if one attempts to solve the constraint equations, the system

still collapses to a deterministic configuration:

$$\mathbb{E}_{\mathrm{MF}}[\widetilde{V}_i] = \mathbb{E}_{\mathrm{MF}}\left[\sum_{t=1}^{T} b_{it}h_t\right] = \sum_{t=1}^{T}\left(\sum_{\substack{j=1\\ j\neq i}}^{N} p_{it}p_{jt} + p_{it}\right)$$
$$= \sum_{t=1}^{T}\left(\sum_{j=1}^{N} p_{it}p_{jt} - p_{it}(p_{it}-1)\right) = \sum_{t=1}^{T} p_{it}h_t^* - \mathrm{Var}_{\mathrm{MF}}[k_i] \qquad \forall i, \quad (5.48)$$

where we used the fact that $b_{it}^2 = b_{it}$. Equating this expression to $\widetilde{V}_i$ and summing over i, one ends up again with condition (5.46). The role of $\widetilde{V}_i$ is, then, different and purely operational: by including the self-term, the link probabilities p_{it} acquire a closed-form expression, which makes maximum-likelihood estimation feasible even though exact constraint satisfaction remains impossible.

The modified vector of patterns $\widetilde{\boldsymbol{V}} = (\widetilde{V}_1, \cdots, \widetilde{V}_N)$, together with the degree sequences $\boldsymbol{k}$ and $\boldsymbol{h}$, define the constraints of the Bipartite Partial Local Overlap Model. The model Hamiltonian reads:

$$H(\mathbf{B}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}) = \sum_{i=1}^{N}\sum_{t=1}^{T}\left[\beta_i + \gamma_t + \delta_i\, h_t(\mathbf{B})\right] b_{it}. \quad (5.49)$$

Applying the mean-field approximation yields:

$$H_{\mathrm{MF}}(\mathbf{B}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}) = \sum_{i=1}^{N}\sum_{t=1}^{T} b_{it}\left[\beta_i + \gamma_t + \delta_i\, \mathbb{E}[h_t]\right] = \sum_{i=1}^{N}\sum_{t=1}^{T} b_{it}\left[\beta_i + \gamma_t + \delta_i\, h_t^*\right], \quad (5.50)$$

where we assume that the constraints on the degrees $\boldsymbol{h}$ are satisfied. Thus, in mean-field, the distribution factorizes:

$$P_{\mathrm{MF}}(\mathbf{B}; \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}) = \prod_{i=1}^{N}\prod_{t=1}^{T} p_{it}^{b_{it}}(1-p_{it})^{1-b_{it}}, \quad (5.51)$$

with link probabilities admitting the closed form

$$p_{it} = \frac{e^{-\beta_i - \gamma_t - \delta_i h_t^*}}{1 + e^{-\beta_i - \gamma_t - \delta_i h_t^*}}. \quad (5.52)$$

The log-likelihood reads:

$$\mathcal{L}_{\mathrm{MF}}(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\delta}) = -\sum_{i=1}^{N}\sum_{t=1}^{T} b_{it}^*(\beta_i + \gamma_t + \delta_i h_t^*) - \sum_{i=1}^{N}\sum_{t=1}^{T} \ln\left(1 + e^{-\beta_i - \gamma_t - \delta_i h_t^*}\right). \quad (5.53)$$

Setting its derivatives to zero, we get for the degree constraints:

$$\frac{\partial \mathcal{L}}{\partial \beta_i} = -\sum_{t=1}^{T} b_{it}^* + \sum_{t=1}^{T} p_{it} = 0 \quad \Longrightarrow \quad k_i^* = \sum_{t=1}^{T} p_{it} = \mathbb{E}_{\mathrm{MF}}[k_i], \quad (5.54)$$

$$\frac{\partial \mathcal{L}}{\partial \gamma_t} = -\sum_{i=1}^{N} b_{it}^* + \sum_{i=1}^{N} p_{it} = 0 \quad \Longrightarrow \quad h_t^* = \sum_{i=1}^{N} p_{it} = \mathbb{E}_{\mathrm{MF}}[h_t]. \quad (5.55)$$

For the $\widetilde{V}_i$ constraints:

$$\frac{\partial \mathcal{L}}{\partial \delta_i} = -\sum_{t=1}^{T} h_t^* b_{it}^* + \sum_{t=1}^{T} h_t^* p_{it} = 0 \qquad \Longrightarrow \qquad \widetilde{V}_i^* = \sum_{t=1}^{T} h_t^* \, p_{it} \tag{5.56}$$

while the expected value is:

$$\mathbb{E}_{\mathrm{MF}}[\widetilde{V}_i] = \sum_{t=1}^{T} p_{it} h_t^* - \mathrm{Var}_{\mathrm{MF}}[k_i]. \tag{5.57}$$

Therefore:

$$\widetilde{V}_i^* = \mathbb{E}_{\mathrm{MF}}[\widetilde{V}_i] + \mathrm{Var}_{\mathrm{MF}}[k_i] \qquad \forall i. \tag{5.58}$$

The model thus reproduces, in expected value, first-order patterns exactly, and second-order patterns as closely as possible while preserving fluctuations. By doing so, it captures each node's local tendency to co-activate with others, without collapsing onto the observed configuration.

Figures 5.9 and 5.10 summarize how BiPLOM enforces the constraints in practice. The parity plots in Figure 5.9 compare the constraint implementation of BiCM (top row) and BiPLOM (bottom row) for a representative subject. Both models accurately reproduce the degree sequences. In particular, the BiPLOM mean-field estimates $\mathbb{E}\mathrm{MF}[k_i]$ and $\mathbb{E}\mathrm{MF}[h_t]$ lie essentially on the identity line when plotted against the empirical degrees, indicating that first-order patterns are matched up to the numerical tolerance of the solver. As expected, BiCM does not reproduce the individual tendency to activate in pairs, as reflected in the aggregated overlaps V_i. By construction, BiPLOM captures these second-order aggregated tendencies much more accurately, albeit at the cost of the mean-field approximation. The bottom-right panel reports the aggregated overlaps $\widetilde{V}_i$, comparing the empirical values with both the plain mean-field prediction $\mathbb{E}_{\mathrm{MF}}[\widetilde{V}i]$ and the variance-corrected estimate $\mathbb{E}_{\mathrm{MF}}[\widetilde{V}_i] + \mathrm{Var}_{\mathrm{MF}}[k_i]$.

A complementary error diagnostic for $\widetilde{V}_i$ across subjects is shown in Figure 5.10: the histogram displays the distribution of relative errors across all nodes and subjects; the violin plot reports the per-subject median relative error; and the hexbin plot compares the empirical discrepancy $\widetilde{V}_i^* - \mathbb{E}_{\mathrm{MF}}[\widetilde{V}_i]$ with the corresponding variance-correction term. Taken together, these panels show that including the variance term systematically compensates for the bias inherent in the plain mean-field estimate of $\widetilde{V}_i$.

Although this systematic bias is intrinsic to our approximation, it is predictable and therefore controllable: a direct check on the variances provides an a posteriori diagnostic of the magnitude of the correction. In this sense, the residual discrepancy represents the price to pay for retaining a model that remains analytically tractable and computationally efficient, and that can therefore be implemented simply and efficiently.

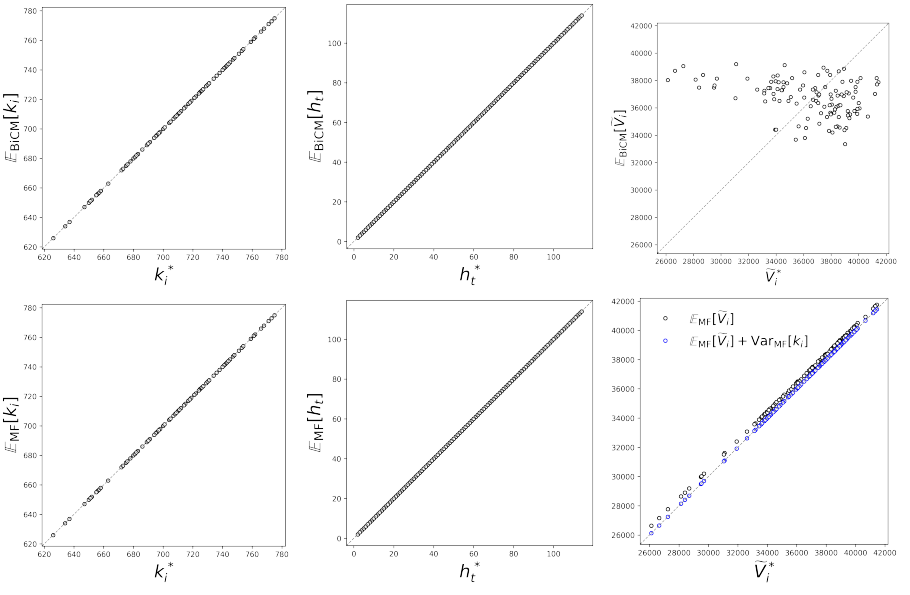


Figure 5.9: ***Constraint implementation with BiCM and BiPLOM.*** *Top row: parity plots for a random subject (id: 366466) for bottom layer degrees k_i (left), top layer degrees h_t (centre), and aggregate overlaps $\widetilde{V}_i$ (right). Each marker reports an empirical value on the horizontal axis and the corresponding mean-field prediction on the vertical axis; the dashed line represents the identity.*

Appendix 5.B Approximations of the Poisson-binomial distribution

As discussed in Section 5.2.2, for each subject and each pair or triplet of regions, we count the number of co-activations V_{ij} and V_{ijk} from the bipartite matrix $\mathbf{B}$. Under the BiCM and BiPLOM null models (Section 5.3.2), the entries b_{it} are modeled as independent Bernoulli variables with success probabilities p_{it}. For a fixed n-tuple of regions ($n = 2, 3$), the contribution of each time point t is Bernoulli with probability

$$q_t = \begin{cases} p_{it}p_{jt}, & n = 2, \\ p_{it}p_{jt}p_{kt}, & n = 3, \end{cases} \tag{5.59}$$

so that the total number of co-activations can be written in the general form

$$S = \sum_{t=1}^{T} X_t, \tag{5.60}$$

where the X_t are independent Bernoulli variables with heterogeneous success probabilities q_t. In this setting, S follows a Poisson-binomial distribution:

$$f_{\text{BP}}(S = s) = \sum_{\substack{A \subseteq \{1,\dots,T\} \\ |A| = s}} \prod_{t \in A} q_t \prod_{t \notin A} (1 - q_t), \qquad s = 0, 1, \dots, T, \tag{5.61}$$

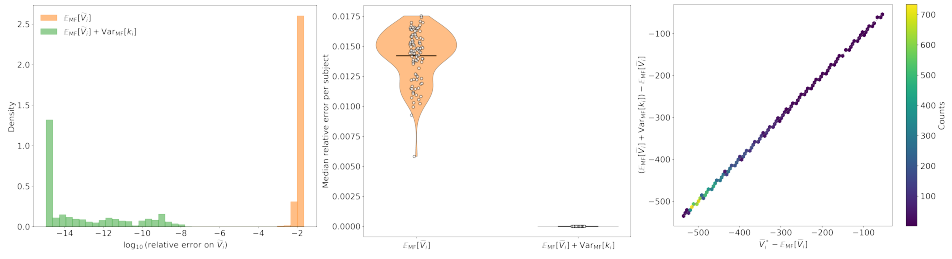


Figure 5.10: ***BIPLOM error diagnostics for the aggregate overlaps.*** *Left panel: distribution of* $\log_{10}$ *relative errors on* $\widehat{V}_i$ *across all nodes and subjects, comparing the prediction* $\mathbb{E}_{MF}[\widehat{V}_i]$ *(orange) with the variance corrected prediction* $\mathbb{E}_{MF}[\widehat{V}_i] + \mathrm{Var}_{MF}[k_i]$ *(green). Central panel: violin plot of the per-subject median relative errors for the same two predictions. Right panel: hexbin plot of* $\widehat{V}_i^* - \mathbb{E}_{MF}[\widehat{V}_i]$ *versus* $(\mathbb{E}_{MF}[\widehat{V}_i] + \mathrm{Var}_{MF}[k_i]) - \mathbb{E}_{MF}[\widehat{V}_i]$*, with color indicating the number of nodes falling in each hexagonal bin.*

whose mean, variance, and skewness are

$$\mu = \mathbb{E}[S] = \sum_{t=1}^{T} q_t, \tag{5.62}$$

$$\sigma^2 = \mathrm{Var}(S) = \sum_{t=1}^{T} q_t(1 - q_t), \tag{5.63}$$

$$\gamma = \sigma^{-3} \sum_{t=1}^{T} q_t(1 - q_t)(1 - 2q_t). \tag{5.64}$$

Given an observed co-activation count V_{obs}, the exact upper-tail p-value under the Poisson-binomial model is

$$p_{\mathrm{PB}}(V_{\mathrm{obs}}) = \mathbb{P}_{\mathrm{PB}}(S \geqslant V_{\mathrm{obs}}) = \sum_{s=V_{\mathrm{obs}}}^{T} f_{\mathrm{PB}}(s), \tag{5.65}$$

where f_{PB} is the Poisson-binomial probability mass function. In our implementation, this tail is evaluated using an FFT-based Hong-type algorithm and serves as the reference for assessing the accuracy of the approximations described below.

Exact Poisson-binomial tails provide a principled baseline but become computationally expensive when applied to all pairs and triplets of all 100 subjects. Following [55], we therefore consider three approximations that replace the Poisson-binomial distribution with simpler surrogates, all sharing the same mean μ and variance σ^2.

The Poisson approximation assumes that the success probabilities q_t are small and that $\sum_t q_t^2$ is not too large. In this regime, S is approximated by a Poisson variable with rate μ, and the upper-tail p-value is

$$p_{\mathrm{Poisson}}(V_{\mathrm{obs}}) \approx \mathbb{P}\big(\mathrm{Poisson}(\mu) \geqslant V_{\mathrm{obs}}\big) = \sum_{s=V_{\mathrm{obs}}}^{\infty} \frac{\mu^s \, \mathrm{e}^{-\mu}}{s!}. \tag{5.66}$$

This approximation is simple and efficient, but it can underestimate the tail probability when the q_t are heterogeneous or not uniformly small.

A second approximation relies on the central limit theorem. For sufficiently large T and moderate skewness, S can be approximated by a Gaussian variable with mean μ and variance σ^2. Introducing the continuity-corrected standardized statistic

$$Z = \frac{V_{\text{obs}} + 0.5 - \mu}{\sigma}, \tag{5.67}$$

the upper tail is approximated by

$$p_{\text{normal}}(V_{\text{obs}}) \approx \mathbb{P}\big(\mathcal{N}(0,1) \geqslant Z\big) = \Phi(-Z), \tag{5.68}$$

where Φ is the standard normal cumulative distribution function. The continuity correction improves accuracy in the discrete setting, particularly near the validation threshold.

The refined normal (rna) approximation incorporates the skewness γ of S into the Gaussian prediction, providing a one-parameter correction to the normal tail that remains tractable for large numbers of tests. Following [55], define

$$G(x) = \Phi(-x) - \gamma \frac{1 - x^2}{6}\, \varphi(x), \tag{5.69}$$

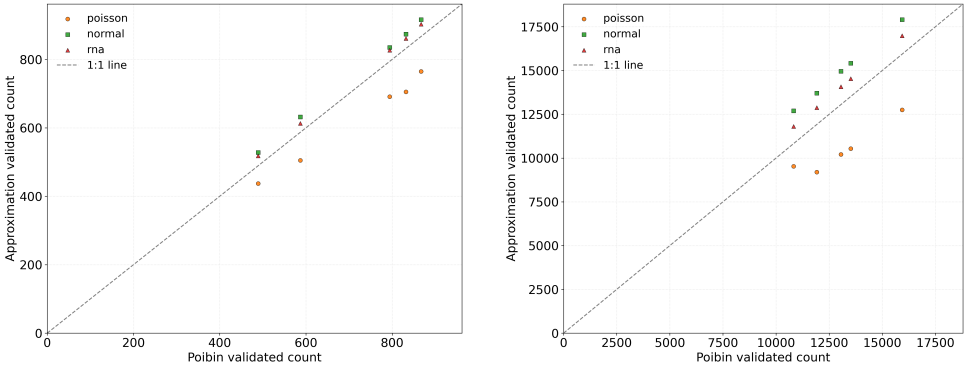
where φ is the standard normal density. The upper-tail p-value is then approximated as

$$p_{\text{rna}}(V_{\text{obs}}) \approx G\left(\frac{V_{\text{obs}} + 0.5 - \mu}{\sigma}\right). \tag{5.70}$$

For $\gamma = 0$, the refined approximation reduces to the ordinary normal approximation. For moderate skewness, the skewness correction improves the agreement with the Poisson-binomial tail without sacrificing computational efficiency. In our implementation, these three approximations are available as alternative options in the validation routines for pairs and triplets (BiPLOM).

To select a surrogate that can be safely used on the full cohort of 100 subjects, where exact Poisson-binomial calculations for all pairs and triplets would be computationally prohibitive, we benchmark Poisson, normal, and rna against the exact Poisson-binomial model on five randomly selected subjects, identified as 100307_treshold1, 125525_treshold1, 149539_treshold1, 198451_treshold1, and 899885_treshold1. For each subject and structure type (pairs for BiCM, triplets for BiPLOM), we compute, for each method, the total number of validated structures, the distribution of associated p-values, and the overlap between the corresponding sets of validated structures.

The scatter plots in Figure 5.11 compare, for each subject, the number of validated structures under Poisson, normal, and rna approximations with respect to the Poisson-binomial baseline. For BiCM pairs, the Poisson approximation validates on average about 0.87 times as many pairs as the Poisson-binomial model (mean ratio Poisson/PoiBin ≈ 0.871), whereas the normal and rna approximations are slightly more liberal, with mean ratios normal/PoiBin ≈ 1.063 and rna/PoiBin ≈ 1.044. The same pattern becomes more pronounced for BiPLOM triplets, where Poisson validates on average only about 0.80 times as many triplets as Poisson-binomial (mean ratio ≈ 0.803), while normal and rna validate about 1.15 and 1.08 times as many triplets, respectively (mean ratios ≈ 1.147 and ≈ 1.079). The points lie close to



(a) *(b)*

Figure 5.11: ***Validated counts under Poisson-binomial surrogates.*** *Scatter plots comparing the number of validated structures obtained with Poisson (orange), normal (green), and rna (red) approximations against the Poisson-binomial baseline (x-axis) across the five benchmark subjects. (a) BiCM pairs. (b) BiPLOM triplets. The dashed diagonal marks equality.*

the diagonal for all methods, indicating that the approximations track the baseline counts across subjects, but with a systematic conservative bias for Poisson and a mild anti-conservative bias for normal and rna. Across both structure types, rna remains close to the diagonal and avoids the strong underestimation of Poisson and the more pronounced over-validation of normal.

To characterize p-value distributions beyond simple counts, we analyze the transformed scores $z = -\log p$ (natural logarithm). For each approximation m and structure type, the violin plots in Figure 5.12 display the empirical distribution of z across structures and subjects. Each violin encodes a smoothed kernel density estimate

$$\hat{f}_m(z) = \frac{1}{Nh} \sum_{k=1}^{N} K\left(\frac{z - z_k}{h}\right), \tag{5.71}$$

where z_k are the observed $-\log p$ values for method m, K is a kernel function and h is the bandwidth. Higher and wider violins toward large z indicate heavier tails of small p-values and thus more aggressive validation. For both BiCM pairs and BiPLOM triplets, the rna approximation produces $-\log p$ distributions that closely follow the Poisson-binomial reference, with only a modest shift towards larger values. By contrast, the Poisson approximation concentrates more mass at smaller z, that is, larger p-values, which is consistent with its systematically smaller number of validated structures.

We also inspect the kernel density of the original p-values. For each approximation m, the curves in Figure 5.13 represent the kernel density estimate

$$\hat{g}_m(p) = \frac{1}{Nh} \sum_{k=1}^{N} K\left(\frac{p - p_k}{h}\right), \tag{5.72}$$

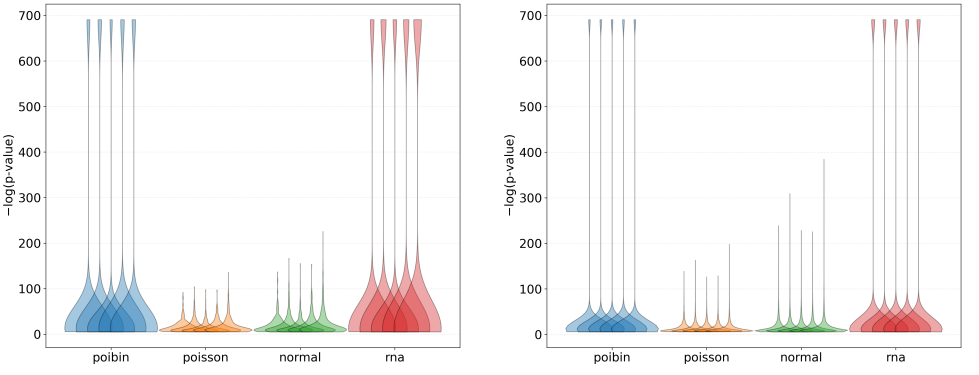
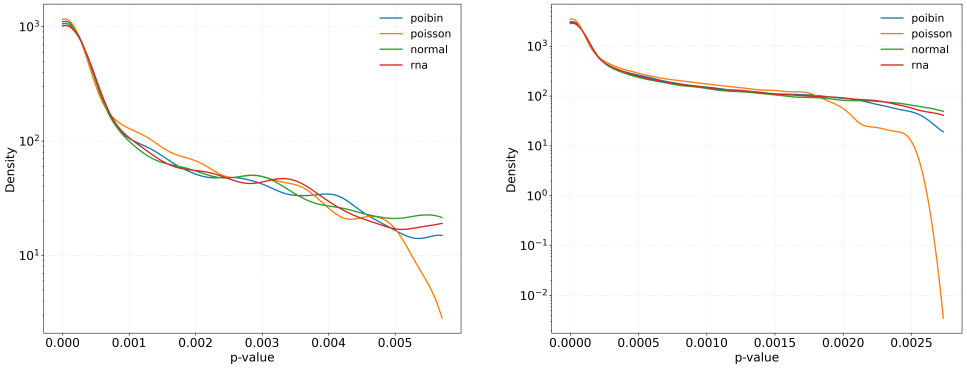


Figure 5.12: ***Distribution of significance scores across approximations.*** *Violin plots of $z = -\log p$ (natural logarithm) for Poisson, normal, and rna approximations, shown separately for each benchmark subject. (a) BiCM pairs. (b) BiPLOM triplets. Colors follow Fig. 5.11.*

constructed from all p-values p_k pooled across the five subjects. The horizontal axis reports the p-value on a linear scale, while the vertical axis reports the estimated density on a logarithmic scale, which emphasizes the behavior near $p = 0$. For both structure types, Poisson shows a marked depletion of small p-values compared to the Poisson-binomial baseline, while normal and rna closely track the Poisson-binomial density. The rna curves are slightly enhanced in the extreme tail, in line with the modestly higher number of validated structures reported in Figure 5.11, but remain very close to the Poisson-binomial reference.

The stacked overlap plots in Figure 5.14 decompose, for each benchmark subject, the union of the Poisson-binomial and approximate validated sets into structures validated only by Poisson-binomial, by both methods, or only by the approximation. For BiCM pairs, the Poisson approximation recovers on average about 87% of Poisson-binomial validated pairs, and 100% of Poisson-validated pairs are also validated by Poisson-binomial. The normal and rna approximations recover essentially all Poisson-binomial validated pairs (mean coverage close to 100%). In contrast, about 94% and 96% of normal- and rna-validated pairs, respectively, are also validated by the Poisson-binomial model. For BiPLOM triplets, the pattern is similar but more pronounced: Poisson recovers about 80% of Poisson-binomial triplets, and all Poisson-validated triplets are contained in the Poisson-binomial set, whereas normal and rna recover essentially 100% of Poisson-binomial triplets and about 87% and 93% of their own validated triplets overlap with Poisson-binomial. For the representative subject 899885, the rna approximation validates 14539 triplets at a subject-specific threshold $p \leqslant 0.00287$, corresponding to approximately $14539/253460 \approx 5.7\%$ of all possible triplets, and 92.98% of these rna-validated triplets are also validated under Poisson-binomial, while all Poisson-binomial validated triplets are recovered by rna. All proportions are computed on linear scales in both axes.



(a) *(b)*

Figure 5.13: ***p-value densities under Poisson-binomial surrogates.*** *Kernel density estimates of p-values for Poisson-binomial (blue), Poisson (orange), normal (green), and rna (red), pooled across benchmark subjects. (a) BiCM pairs. (b) BiPLOM triplets. The density axis is shown on a logarithmic scale to emphasize the small-p regime driving validation.*

To summarize the similarity between the Poisson-binomial and approximate validated sets, we use the Jaccard index. For each subject and approximation, we define

$$J\big(H_{\mathrm{PoiBin}}, H_{\mathrm{approx}}\big) = \frac{|H_{\mathrm{PoiBin}} \cap H_{\mathrm{approx}}|}{|H_{\mathrm{PoiBin}} \cup H_{\mathrm{approx}}|}, \tag{5.73}$$

where H_{PoiBin} is the set of structures validated by the Poisson-binomial model and H_{approx} is the set validated by the approximation. The Jaccard index equals 1 when the two sets coincide and 0 when they are disjoint. The bar plots in Figure 5.15 report the mean Jaccard index across the five subjects, together with one standard deviation. For BiCM pairs, the mean Jaccard index is about 0.87 for Poisson, 0.95 for normal, and 0.96 for rna. For BiPLOM triplets, the mean Jaccard index decreases slightly to about 0.78 for Poisson, 0.87 for normal, and 0.93 for rna. These values confirm that all approximations remain strongly nested within the Poisson-binomial benchmark, while highlighting systematic differences: Poisson is clearly conservative, with smaller validated sets and lower Jaccard values, whereas normal and especially rna achieve a higher recall of Poisson-binomial validated structures at the cost of a modest number of additional validations.

Taken together, the evidence from validated counts, p-value distributions, overlap decompositions, and Jaccard indices consistently points to rna as the most faithful approximation to the Poisson-binomial model. Compared with Poisson, rna avoids the strong underestimation of the number of validated structures and the depletion of small p-values. Compared with normal, rna achieves higher Jaccard indices and a closer match to the Poisson-binomial p-value density, particularly in the small p-value regime that drives validation. Based on this five-subject benchmark, we therefore adopt rna as the default surrogate for the Poisson-binomial model in the large-scale analysis of all 100 subjects.

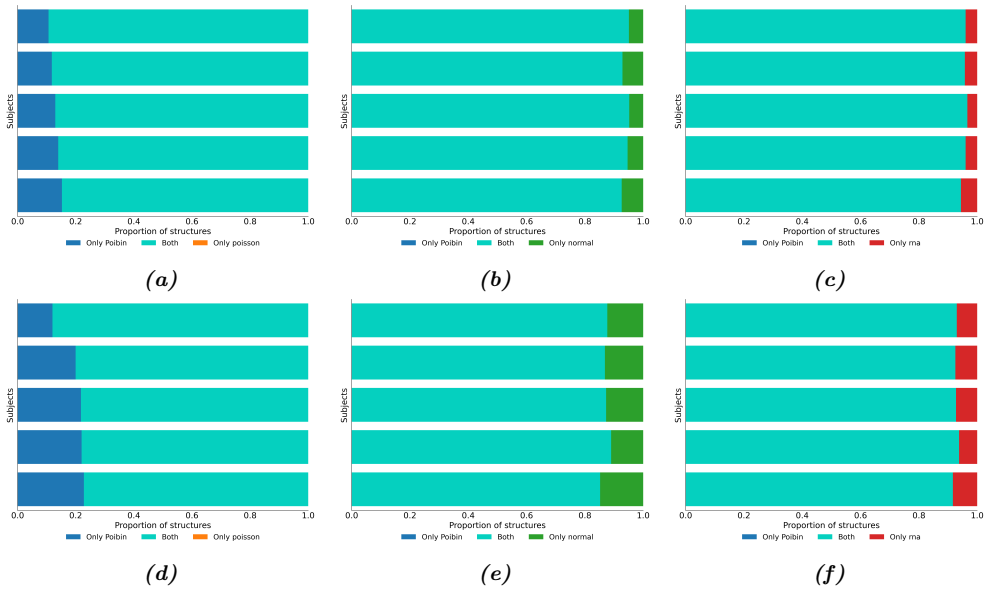


Figure 5.14: ***Overlap between Poisson-binomial and approximate validated sets.*** *Stacked decomposition of the union between Poisson-binomial validations and each approximation into structures validated only by Poisson-binomial, by both methods, or only by the approximation. Panels compare (a,d) Poisson, (b,e) normal, and (c,f) rna approximations against Poisson-binomial. Top row: BiCM pairs. Bottom row: BiPLOM triplets. For each subject, bar lengths are normalized by the union size so that each bar sums to one; subjects are ordered by decreasing fraction of Poisson-binomial-only structures.*

Appendix 5.C Atlas

The atlas considered, being the union between the Schaefer Atlas and the Tian Atlas, includes both subcortical and cortical regions. Subcortical regions are divided into left and right hippocampus, left and right amygdala, left and right posterior and anterior thalamus, left and right nucleus accumbens, left and right globus pallidus, left and right putamen, and left and right caudate nucleus. Cortical regions, instead, are grouped based on the seven canonical Resting-State Networks (RSNs) defined by Yeo et al. [135]. Specifically, these include 17 different areas corresponding to the visual system (VS), 14 accounting for the sensorimotor network (SMN), 15 for the Dorsal Attention Network (DAN), 12 for the Salience Ventral Attention Network (SVA), 5 for the limbic system, 13 for the Central Executive Network (CEN), also called Fronto-Parietal Network (FPN), and 24 for the Default Mode Network (DMN). The above partition is represented in Fig.5.16.

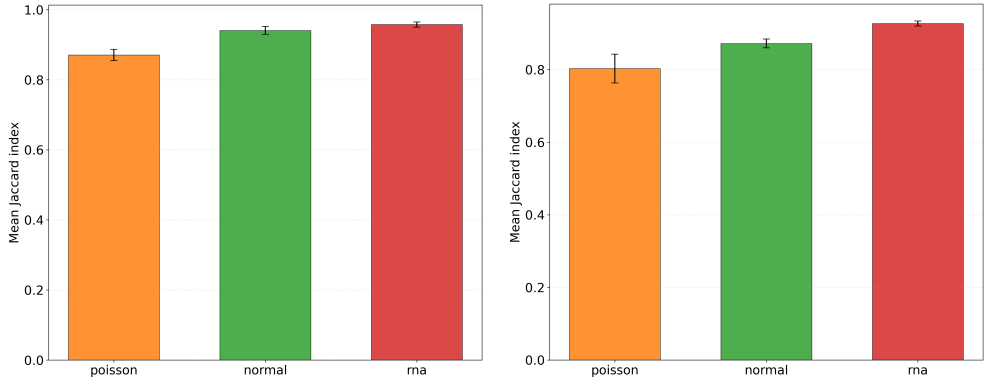


Figure 5.15: ***Agreement with Poisson-binomial validated sets.*** *Mean Jaccard index between the Poisson-binomial validated set and the sets obtained with Poisson, normal, and rna approximations, computed across the five benchmark subjects. (a) BiCM pairs. (b) BiPLOM triplets. Error bars indicate one standard deviation across subjects.*

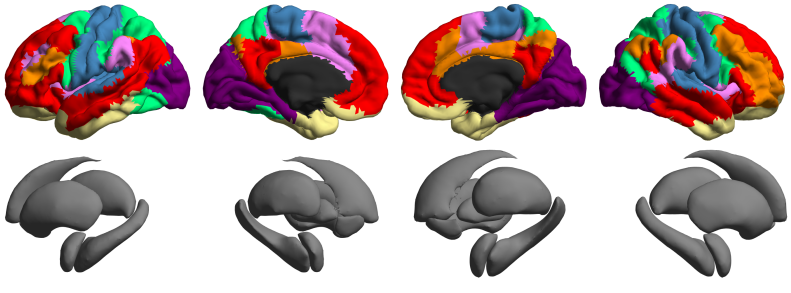


Figure 5.16: ***Yeo partition illustration on the brain.*** *Here, we show Yeo's partition through the Enigma Toolbox [153]. In gray, the subcortical structures that were not actually included in the Yeo analysis and here have been grouped together, in blue the SMN, in purple VS, in orange FPN, in green DAN, in pink VAN, in yellow the limbic system, in red the DMN.*