

Unveiling the higher-order organization of multivariate time series

Andrea Santoro,¹ Federico Battiston,² Giovanni Petri,³ and Enrico Amico^{1,4,*}

¹*Neuro-X Institute, École Polytechnique Fédérale de Lausanne, 1202 Geneva, Switzerland*

²*Department of Network and Data Science, Central European University, 1100 Vienna, Austria*

³*CENTAI, Corso Re Inghilterra 3, Turin, Italy*

⁴*Department of Radiology and Medical Informatics,
University of Geneva, 1211 Geneva, Switzerland*

(Dated: September 13, 2022)

Time series analysis has proven to be a powerful method to characterize several phenomena in biology, neuroscience and economics, and to understand some of their underlying dynamical features. Despite a plethora of methods have been proposed for the analysis of multivariate time series, most of them neglect the effect of non-pairwise interactions on the emerging dynamics. Here, we propose a novel framework to characterize the temporal evolution of higher-order dependencies within multivariate time series. Using network analysis and topology, we show that, unlike traditional tools based on pairwise statistics, our framework robustly differentiates various spatiotemporal regimes of coupled chaotic maps, including chaotic dynamical phases and various types of synchronization. Hence, using the higher-order co-fluctuation patterns in simulated dynamical processes as a guide, we highlight and quantify signatures of higher-order patterns in data from brain functional activity, financial markets, and epidemics. Overall, our approach sheds new light on the higher-order organization of multivariate time series, allowing a better characterization of dynamical group dependencies inherent to real-world data.

Keywords: multivariate time series, higher-order dependencies, topological data analysis, complex systems, network theory

The growing availability of rich, often temporally resolved, data coming from many different complex systems, has led to the possibility of studying in detail their behaviour and –often– their internal mechanisms. Examples of such systems include epidemics, social contagion, as well as financial, brain, and biological signals. All of these systems are composed by large numbers of elementary units interacting in heterogeneous fashion with each other, and –in virtually all cases– displaying emergent properties at the macroscopic level. Due to the crucial importance of the patterns of interactions composing such systems, it is no surprise that complex networks emerged as a powerful framework to investigate the structure and dynamics of such systems [1, 2], and have helped to characterize several real-world phenomena, including disease spreading [3], synchronisation [4], diffusion [5], and opinion formation [6].

Despite being widely considered as the reference model for many real-world complex systems [7–9], networks are limited to describing interactions between two units (or nodes) at a time. This however clashes with the growing empirical evidence for group interactions in social systems [10], neuroscience [11–13], ecology [14] and biology [15]. In all the aforementioned cases, connections and relationships do not take place only between pairs of nodes, but also as collective actions of groups of nodes. By taking into account the higher-order (group) interactions in more refined models, such as hypergraphs and simplicial complexes [16, 17], several recent studies have shown that the presence of higher-order interactions can

have a substantial impact on the dynamics of interacting systems [18, 19], ranging from alterations of the synchronization [20] and diffusion [21, 22] properties, to new collective dynamics in social [23–25] and evolutionary processes [26].

Yet, direct measurements of pairwise or group interactions to inform and constrain such higher-order models are rarely available. Hence, one must typically rely on indirect data, commonly extracted from time series of node activities, under the assumption that the system’s repertoire of spatiotemporal activity patterns encodes information about the underlying interactions. Indeed, examples of these complex patterns are observed in the neuronal activity of the brain, supporting a wide variety of motor and cognitive functions [27, 28], in financial markets, where partially synchronized patterns often reflect periods of financial stress [29, 30], but also in the co-evolution of biological species [31–33].

While the inference of pairwise interactions has a long history [34], researchers have only recently taken the first steps towards reconstructing or filtering higher-order interactions [35–38]. In particular, methods relying only on pairwise statistics might be in principle insufficient as significant information can be present only in the joint probability distribution and not in the pairwise marginals, therefore failing at identifying higher-order behaviours [39]. To date, it remains unclear to what degree the information encoded in multivariate time series stems from independent individual entities or, rather, from their group interactions. A clear example of this issue is provided by the conventional “functional connectivity” between two brain regions [40, 41]: a pairwise connection is drawn irrespective of whether the activities

* Corresponding author: enrico.amico@epfl.ch

of the two regions peaked as a pair, or as part of a larger group of functionally coherent regions.

Existing proposals to address this issue are mostly limited in their capacity to describe either the temporality or the complexity of such higher-order interactions, with only few exceptions on the topic [42]. For example, a recent set of information-theoretic methods characterized higher-order dependencies in multivariate time series by quantifying the intrinsic statistical synergy and redundancy in groups of three or more interacting variables [43–46]. Moreover, a recent approach [47] at the interface of network science and random matrix theory has also proven suitable to unveil the mesoscopic organization of correlation matrices. Yet, while very powerful, these methods hardly capture the information about the dynamics of the system, because they require integration over time. A recent exception comes from a work introducing a spectral decomposition that resolve the statistical synergy and redundancy of groups of variables into different frequency bands, which allows to analyze locally in time higher-order dependencies [42]. By contrast, tools coming from network neuroscience and signal processing easily deal with the dynamics of multivariate time series [48–55], but only focus on pairwise statistics and neglect the effects of higher-order interactions. As a result, a principled approach to quantify the instantaneous dynamics of groups of nodes and possibly infer its higher-order representation is still missing.

Here, we propose a novel framework to characterize the instantaneous co-fluctuation patterns of signals at all orders of interactions (pairs, triangles, etc.), and to investigate the global topology of such co-fluctuations. We do this by bridging time series analysis, complex network theory, and topological data analysis [56]. We first validate the framework by exploring the rich high-dimensional dynamics displayed by canonical models of spatiotemporal chaos. In particular, we demonstrate that, unlike traditional tools of time-series analysis based on pairwise statistics [57–59], higher-order measures are able to reveal the subtleties of different spatiotemporal regimes at the level of individual frames. We then use the insights obtained from these synthetic models as a Rosetta stone to interpret the higher-order structures reconstructed from time series concerning three diverse real-world case studies: resting-state brain activity (as measured by fMRI data), stock option prices, and epidemiological incidence of various diseases in the United States. In all cases, we unveil additional rich higher-order information that is not captured at the node and dyadic network level, and highlight distinct topological dynamical regimes, which in turn yield to instantaneous classification of the system’s states.

Finally, we show how the inferred dynamical higher-order structure provides instantaneous topological snapshots of the spatial configuration of the system, which can be used as input for further tasks on real datasets, including detecting local integration of brain regions, exploring periods of financial crisis, or classifying disease type from

spatial spreading patterns.

RESULTS

Topological markers of higher-order structure in multivariate time series. Simplicial complexes are well suited as modelling framework to describe the co-existence of pairwise and higher-order interactions [16]. In its most basic definition, a k -simplex σ is a set of $k + 1$ vertices $\sigma = [p_0, \dots, p_k]$. A collection of simplices is a simplicial complex \mathcal{K} if for each simplex σ all its possible subfaces (defined as subsets of σ) are themselves contained in σ (see Methods and Ref. [60] for details). Via this representation it is then easy to distinguish between a group interaction among three elements, which can be represented as a 2-simplex (or “filled” triangle) $[p_0, p_1, p_2]$, and the three pairwise interactions between the nodes, that is, the collection of 1-simplices (edges) $[p_0, p_1]$, $[p_0, p_2]$, $[p_1, p_2]$. The relative importance of pairwise versus higher-order interactions can be encoded in weights over the simplices, resulting in the so-called *weighted simplicial complexes*.

We rely on this representation to describe the higher-order dependencies among multiple time series. Our approach can be summarised in five main steps: First, *i*) we z -score the N original time series (Fig. 1a), and then *ii*) we calculate the element-wise product of the z -scored time series for all the $\binom{N}{k}$ k -order patterns (i.e. edges, triangles, etc.). Here, the generic elements represent the instantaneous co-fluctuation magnitude between a $(k + 1)$ group interaction. Then, *iii*) the resulting new set of time series encoding the k -order co-fluctuations (which corresponds to the so-called *edge time series* [53] in the case of 1-order co-fluctuation) are then further z -scored across time, to make products comparable across k -orders (Fig. 1b). At this point a choice on how to assign signs to the resulting weights is required in order to distinguish fully concordant group interactions (all positive or negative fluctuations) from discordant ones (a mixture of positive and negative fluctuations) in a k -order product. Indeed, these two scenarios might end up having similar co-fluctuation z -scored values after a k -order product (with $k \geq 2$), even if they clearly represent different regimes of group synchronization. Hence, we opted to assign positive signs to the fully concordant group interactions, and negative signs to the discordant ones. The rationale behind the concordant mapping is such that any simultaneous increased (or decreased) activity relative to baseline – no matter the order of the co-fluctuation – is always marked as positive, therefore reflecting a synchronous co-activation pattern. This adjustment is particularly important for the simplicial filtration step, as detailed below. Next, *iv*) for each time frame t , we condense all the instantaneous k -order co-fluctuations in a single mathematical object, i.e. a weighted simplicial complex \mathcal{K}^t (Fig. 1c). Lastly, for each time t , *v*) we construct a filtration $\mathbb{F}(\mathcal{K}^t)$ [61], i.e. a sequence of simplicial

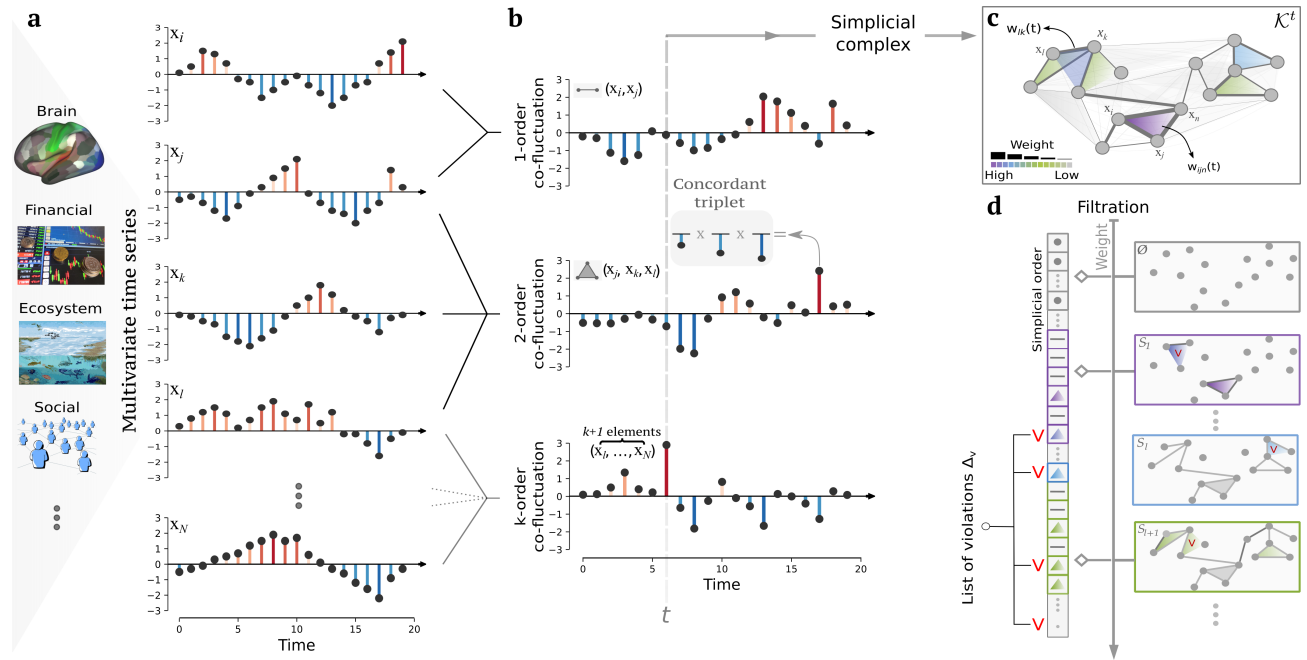


Figure 1. **Unveiling the higher-order structure of multivariate time series: Schematic representation.** (a) We start by extracting different co-fluctuation patterns according to their order from the “raw” nodal time series. (b) The generic element of a k -order co-fluctuation pattern is calculated by z -scoring each time series, and performing an element-wise product of $k + 1$ time series. We further z -score each of these group product time series so that the magnitude of the time-resolved co-fluctuation is comparable across group size (i.e. pairs, triangles, etc.). To distinguish concordant group interactions from discordant ones, we impose that concordant signs are always positively mapped, while discordant signs are negatively mapped. (c) For each time frame t , a weighted simplicial complex is constructed by merging together all the k -order co-fluctuations. (d) Finally, a weight filtration made of all the k -order co-fluctuations allows to identify weighted holes when k -order patterns are gradually included. For simplicity, we set $k = 2$ for panels (c, d).

complexes $\emptyset = S_0 \subset S_1 \subset \dots \subset S_l \subset \dots \subset S_n \subset \mathcal{K}^t$ by sorting all the k -order co-fluctuations by their weights (see Methods for details). The filtration proceeds in a top down fashion from larger weights to smaller weights – in the spirit of persistent homology [61–63] – so that when k -order simplices are gradually included, topological holes start to appear in the simplicial complexes of \mathbb{F} and then potentially close (i.e., descending from more coherent patterns to less coherent). Yet, to maintain valid simplicial complexes as each step of the filtration, only k -order simplices respecting the simplicial closure condition can be included. That is, simplices whose sub-faces are already contained in the simplicial complex at the previous step. To preserve this property, whenever we would add a simplex that does not satisfy this requirement (e.g. a triangle entering the complex before its edges), we consider it as a *simplicial violation*, and exclude it from the filtration. Note that such violating simplices can be considered as *hyper coherent* structures, as their co-fluctuations are stronger than those of its sub-components (see Fig. 1d and Methods for details). Note further that in this paper we present results when considering $k = 2$, so we take into account simplices only up to triangles. Nevertheless, our framework generalizes naturally to higher orders (i.e. $k \geq 3$).

In summary, for each time t , our framework produces

two different outputs:

1. a list of violating triangles, $\Delta_v = \{(i, j, k), w_{ijk}\}$, induced by the simplicial closure condition; these are 2-simplices (triangles) whose weights co-fluctuate more than at least one of their corresponding 1-simplices (edges). Intuitively, these triangles reflect higher-order states that cannot be merely captured by pairwise co-fluctuations. We then define the *hyper coherence indicator*, as the fraction of violating coherent triangles (i.e. violating triangles with a weight greater than zero) over all the possible coherent triangles (i.e. triangles with a weight greater than zero).
2. the simplicial filtration \mathbb{F} , a sequence of embedded simplicial complexes – sorted according to coherent patterns – starting with the empty complex and ending with the entire simplicial complex (see right panel of Fig. 1d). We then compute persistent homology of \mathbb{F} to characterize the persistency of certain topological features (connected components, 1-dimensional cycles, 3D-cavities, etc.) [63, 64]. Here, we focus on examining the 1D cycles in the filtration, i.e. the persistent generators of the first homology group H_1 , which provide insights

about where and when higher synchronised regions emerge. The classical output of persistent homology is a barcode (or equivalently, a persistence diagram), which is a compressed summary describing how long 1D cycles live along \mathbb{F} (see SI Fig. S1). We rely on this object and define the *hyper complexity indicator* as the Wasserstein distance [65] between the persistence diagram of H_1 and the empty persistence diagram, corresponding to a space with trivial H_1 homology (see Methods for details). We obtain in this way a measure of the topological complexity of the landscape of coherent and incoherent co-fluctuations.

Global and local topological markers classify different dynamical regimes. To gain insights into the performance of our topological indicators, we show here that hyper coherence and hyper complexity easily distinguish different dynamical regimes generated by canonical models of spatiotemporal chaos. As a case study, we consider diffusively coupled map lattices (CMLs) [67], which are high-dimensional dynamical systems defined on discrete time and space, with continuous state variables. CMLs are broadly used to model complex spatiotemporal dynamics in several different fields including biology [68], and finance [69, 70]. In particular, we consider a ring lattice with N sites, and we assume that the dynamical evolution of the system of the state x_i of each site i is the result of two different competing dynamics: an internal chaotic dynamic, and an external diffusive coupling dynamic among the first nearest-neighbour sites. Their dynamics can be expressed as

$$x_i(t+1) = (1 - \varepsilon)f[x_i(t)] + \frac{\varepsilon}{2} (f[x_{i-1}(t)] + f[x_{i+1}(t)])$$

where $i = 1, 2, \dots, N$, $\varepsilon \in [0, 1]$ is the coupling parameter (or coupling strength), and $f(x)$ is generally a chaotic map. In all our simulations, we have considered the logistic map, i.e. $f(x) = 1 - 1.75x^2$. It is now well established [66, 71] that by changing the values of the coupling strength ε of the chaotic map, CMLs exhibit a great variety of spatiotemporal patterns, including different degrees of synchronization and dynamical phases such as Fully Developed Turbulence (FDT, a phase with incoherent spatiotemporal chaos and high dimensional attractors), Pattern Selection (PS, a phase with suppression of chaos in favour of randomly selected periodic attractor, reflecting quasiperiodic behaviours), and different forms of spatiotemporal Intermittency (STI, chaotic pseudo-phases with low dimensional attractors that interpolate between FDT and PS). Together with the latter STI class, we also highlight two other different phases such as Brownian Motion with Defects (BMWD, a phase where defects exist in the system and fluctuate chaotically akin to Brownian motion), and Defect Turbulence (DT, a phase where many defects are generated and turbulently collide together) [72]. It is worth remarking that the origin of this very rich phase diagram comes from

the interplay between the local tendency towards inhomogeneity, which is induced by the chaotic dynamic of each single state, and the global tendency to homogenise the system in space, which is induced by the diffusion dynamic [72].

In Figure 2 we summarize the results of our higher-order approach when applied to these synthetic multivariate series with $N = 119$ nodes and $T = 1200$, obtained by concatenating five different dynamical phases of CMLs with fixed time length $L = 240$. Namely, from order to disorder, PS at $\varepsilon = 0.12$, BMWD at $\varepsilon = 0.08$, STI at $\varepsilon = 0.3$, DT at $\varepsilon = 0.068$, and FDT at $\varepsilon = 0.05$ for which a transient of 10^5 time points has been removed. A sample of such multivariate time-series is reported in SI Fig. S2, while we study the effect of the z -score in SI Section S2.2 and SI Fig. S3. Remarkably, the global hyper coherence indicator reported in Fig. 2a clearly distinguishes the different dynamical phases of the CMLs, while also preserving the ranking between ordered and disordered states. More precisely, it assigns high values to fully and partially synchronized regimes while, on the contrary, chaotic or turbulent regimes exhibit lower values of hyper coherence. While this indicator provides only global information, refined information can be obtained by projecting the magnitudes of the list of violating triangles Δ_v as a weighted graph (see Methods for the definition of downward projections). Also in this case, in fact, the edge weight distribution $P(w_{ij})$ reflects the nature and the “rank” of the different dynamical regimes (Fig. 2b). Periodic series, such as PS, convert into well-peaked distributions, akin to Poisson distributions. By contrast, as disorder enters in the pseudo-phases of the multivariate time series, the edge weight distribution gradually changes its shape, with the limit case of the FDT chaotic series converging towards a fat-tailed distribution.

Similar conclusions can be reached when investigating the temporal evolution of the hyper complexity (Fig. 2c). We found, however, some notable differences. For complexity, the lowest value is assigned to periodic patterns (e.g. PS), as these regimes require a low amount of information to be described. Contrarily, chaotic states such as FDT display the highest hyper complexity values. While also this higher-order indicator is able to differentiate the different dynamical regimes of CMLs, one might assume that the hyper coherence and hyper complexity indicators provide equivalent information as indicated by the strong negative correlation (i.e. Spearman’s rank correlation $\rho \approx -0.95$). We will show in the next section that this is not true in general for real-world multivariate time series.

Finally, in Fig. 2d we report the edge weight distributions of the persistence homological scaffolds, graphs constructed from the persistent homology generators of H_1 (see Methods and Ref. [11] for details). These distributions quantify the topological importance of edges in the co-fluctuation landscape in terms of how persistent are the homological generators to which they belong. No-

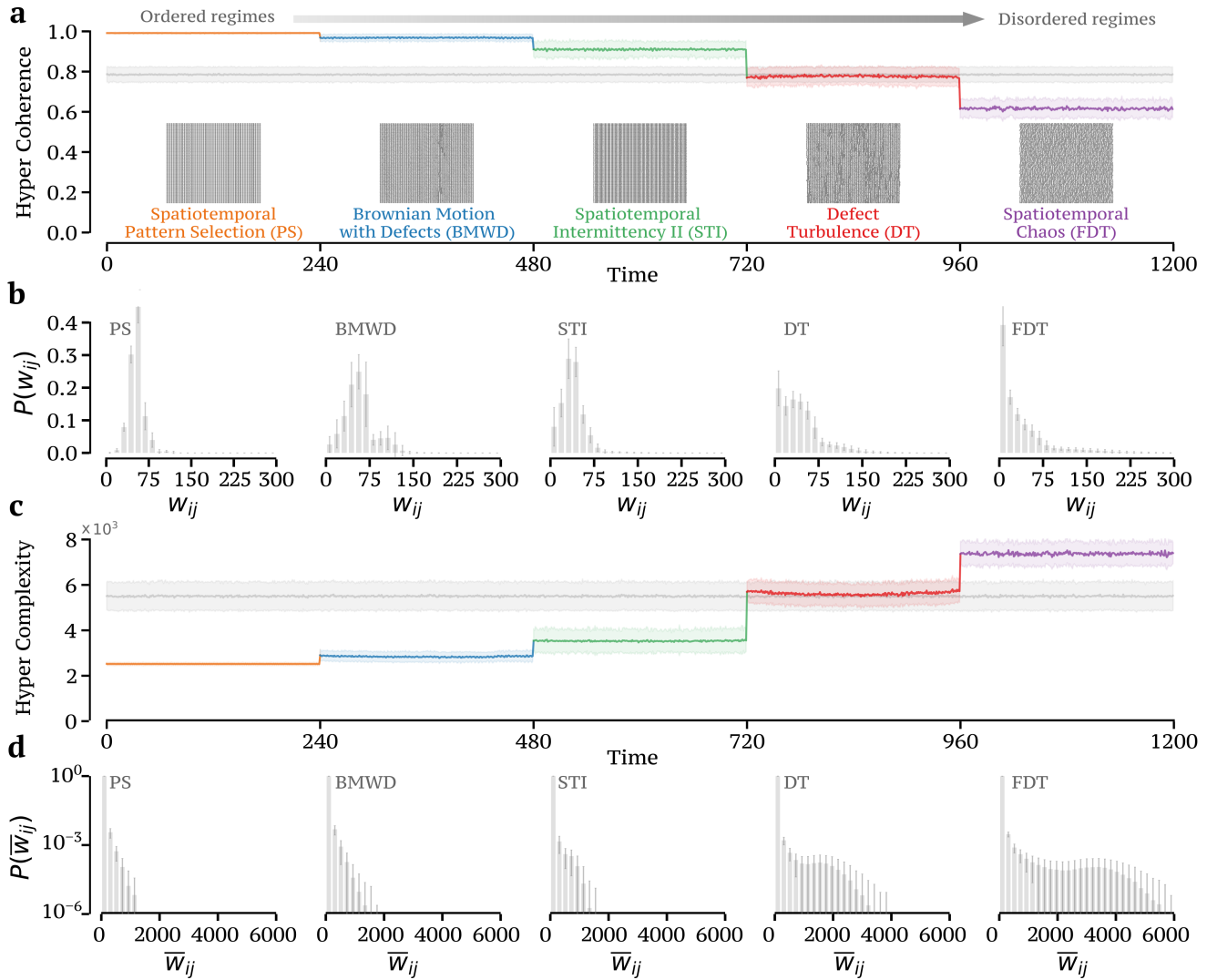


Figure 2. **Global and local higher-order indicators distinguish the dynamical regimes of coupled chaotic maps.** (a) We report the temporal evolution of the hyper coherence indicator for a multivariate time series with $N = 119$ nodes and $T = 1200$, obtained by concatenating five different CML regimes with fixed time length $L = 240$. Namely, from order to disorder, Pattern Selection (PS) at $\varepsilon = 0.12$, Brownian motion with Defects (BMWD) at $\varepsilon = 0.08$, Spatiotemporal Intermittency II (STI) at $\varepsilon = 0.3$, Defect Turbulence (DT) at $\varepsilon = 0.068$, and Fully Developed Turbulence (FDT) at $\varepsilon = 0.05$ [66]. (b) Notably, when projecting the list of violating triangles Δ_v as a weighted graph (see Methods for the definition of downward projections), the edge weight distribution $P(w_{ij})$ reflects the nature and the “ranking” of the different dynamical regimes from ordered to disordered. (c) We plot the temporal evolution of the hyper complexity indicator and the (d) distribution of weights $P(\bar{w}_{ij})$ of the homological scaffold constructed from the persistent homology generators of H_1 [11]. For comparison, we also report in panels (a,c) the same indicators for a null model obtained when reshuffling without any constraint the multivariate time series (grey curve). See SI Section S4 for the behaviour of the higher-order indicators in more conservative null models. Shaded regions and error bars represent standard deviations across 100 independent realizations.

tably, also these distributions change their overall shape as we move from periodic to chaotic multivariate time series preserving, also in this case, the rank between order and disorder.

These results qualitatively confirm that both the global and local topological information extracted with our approach well discriminate among the different dynamical regimes. We quantitatively assess the capacity of our higher-order indicators to differentiate between dy-

namical regimes with the intraclass correlation coefficient (ICC) [73, 74], a statistical measure commonly used to determine the agreement between units (or ratings/scores) of different groups. In other words, the ICC describes how strongly units in the same group resemble each other, so that the stronger the agreement, the higher its ICC value. In Figure 3, we report the comparison of several approaches when trying to differentiate the five dynamical regimes of CML. From a theoretical

standpoint, it is interesting to notice from Figure 3a that the five CMLs regimes map into distinct hyper coherence distributions mirroring some peculiar features intrinsic to each individual spatiotemporal regime. For instance, apart from a well-defined bulk, the STI regime exhibits a long tail towards lower hyper coherence values, which captures the rare appearance of short chaotic bursts arising from the mismatching of the dynamical phases [66, 75].

We find that both our higher-order measures (in Figure 3 only hyper coherence is shown) have high ICC values (i.e. approximately 0.95 and 0.96 for hyper coherence and hyper complexity, respectively). For a commonly used temporal low-order measure, the RSS statistic [54], which accounts for the magnitude of peak amplitude in all the 1-order co-fluctuations, i.e. the edge time series (see Methods for a formal definition), we find instead considerably smaller ICC values ($ICC \approx 0.57$) as shown in Figure 3b, implying that the higher-order effects are dominant. Indeed, our dynamic approach is comparable to “static” higher-order information-theoretic approaches [43, 45], i.e. for s -information (shown in Figure 3c) we find $ICC \approx 0.98$. However, these quantities are typically computed on temporal windows, while in the topological framework presented here it is possible to have instantaneous information. Finally, as expected, higher-order measures (both static and dynamic) consistently outperform static lower-order methods based on Pearson’s correlation, i.e. $ICC \approx 0$ and shown in Figure 3d. As a matter of fact, it appears clear that only higher-order approaches are able to effectively distinguish the various spatiotemporal regimes, while lower-order statistics are not able to capture the subtle differences between dynamical states. For detailed comparisons with other “static” higher-order and pairwise approaches [41, 47, 54], see also SI Section S3 and SI Figures S6-S7.

Real-world complex systems exhibit non-trivial hyper coherence structure. As examples of applications to the analysis of real-world multivariate time series, we report the results of the higher-order framework on fMRI signals from the Human Connectome Project (HCP) [76], on prices of financial assets in the New York Stock Exchange (NYSE), and on historical data of several infectious diseases in the US [77, 78].

For human brain data, we consider resting-state fMRI signals of the HCP 100 unrelated subjects, employing a cortical parcellation of 100 brain regions [79] and 19 sub-cortical ones as provided by the HCP release [80], for a total of $N = 119$ Regions of Interest (ROIs). For financial time series, we analyse the daily time evolution of $N = 119$ stock prices of some of US companies from the NYSE over the period 2000–2021. Finally, for the epidemic dataset, we investigate the weekly number of cases at the US state-level ($N = 50$) for chlamydia, gonorrhoea, influenza, measles, mumps, polio, and pertussis (see Methods for details on the datasets).

In Fig. 4a we report the distributions of hyper

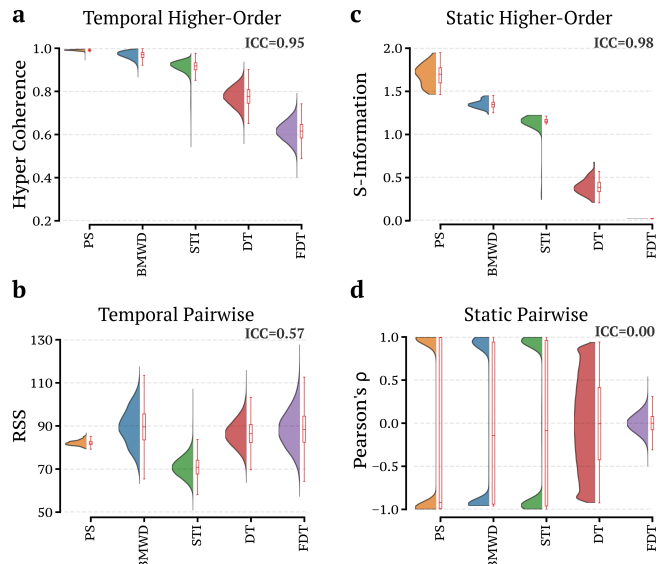


Figure 3. Higher-order approaches perform better in distinguishing the CML regimes. For the five dynamical regimes, we report the violin plots of the (a) Hyper Coherence, (b) RSS [54], (c) S -information [43, 45], and (d) Person’s correlation distributions. Notably, only higher-order approaches are able to distinguish the five dynamical regimes (i.e. $ICC > 0.9$). Also notice that static approaches can only be used when having prior knowledge on the position of each block.

coherence for the five CMLs dynamical regimes and the three datasets. For comparison, we also plot the null models obtained by independently reshuffling synthetic and real-world multivariate time series (see SI Sections S4 and SI Figures S8-S10 for the behaviour of the higher-order indicators in more conservative null models). Several things can be observed when examining the hyper coherence distributions for real-world systems. First, these distributions are always statistically distinct from the corresponding null models (all p -values $< 10^{-10}$, with the Kolmogorov-Smirnov test), yet they also exhibit specific profiles which strongly differ from each other. If we focus on the epidemic data, for example, it is already possible to differentiate the diseases by coarsely comparing the corresponding hyper coherence distributions. These distributions, in fact, reflect the unique higher-order spatiotemporal patterns inherent to the evolution of the disease. For the financial system, by contrast, we obtain a bimodal distribution mirroring the dichotomy between financial periods of crisis and stability. That is, economic crises are typically characterized by increased (hyper) synchronization, whereas periods of financial stability seem to unfold in a more chaotic fashion. Moreover, armed with the CMLs interpretational benchmarks, we find that, during rest, the human brain is mostly associated with chaotic states and few partially synchronized states, in agreement with studies on resting state brain dynamics [82–85].

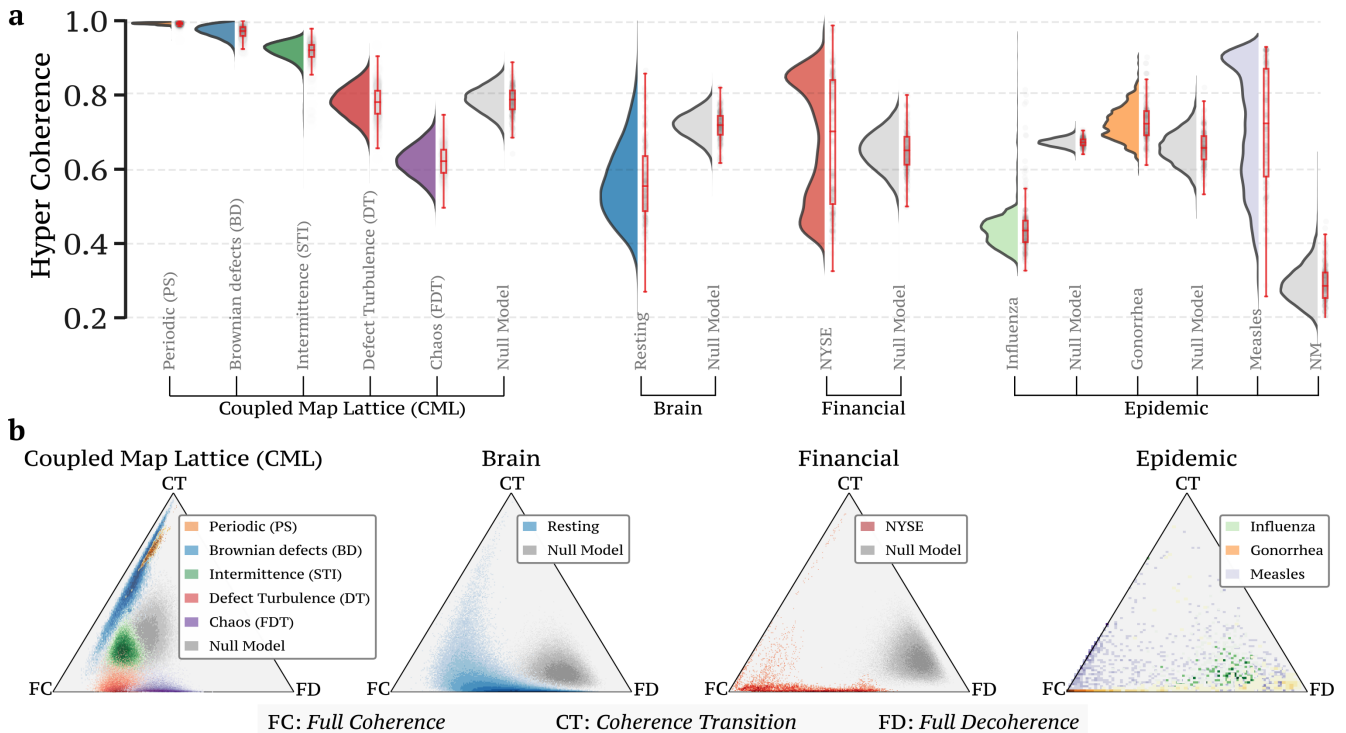


Figure 4. **Higher-order indicators for real-world multivariate time series.** (a) Violin plots showing the distribution of hyper coherence for three real-world datasets, namely, resting-state fMRI data ($N=119$ brain regions), financial prices of 119 assets in NYSE, and the US historical data of several infectious diseases at the US state-level ($N=50$). The real distributions are compared against the five CML dynamical regimes, as well as the corresponding null models obtained when independently reshuffling without any constraint synthetic and real-world multivariate time series (see SI Section S2.3, Fig. S5; SI section S4, Fig. S8-S10 for more details). Note how the distributions of the three real-world datasets employed exhibit noticeable differences in their profile, yet always statistically distinct from the corresponding null models. See SI Section S5 and SI Fig. S11 for a similar analysis with edge-based indicators. (b) Two-dimensional histogram of the different contributions associated with 1D cycles in the landscape of coherent and decoherent co-fluctuations. Here, the position of each point in the triangle is determined by the three different contributions associated with the 1D cycles. For example, a point would be at the centre of the triangle if the hyper complexity indicator splits into three equal contributions of Full Coherence (FC), Coherence Transition (CT), and Full Decoherence (FD), while a corner position is reserved for points whose mainly contributions come either from FC, CT, or FD.

Hyper complexity decomposition provides detailed information about dynamical regimes. To better characterize the evolution of 1D homological generators in the space of coherent and decoherent co-fluctuations, we decompose the hyper complexity indicator into three different contributions. That is, as we track the evolution of 1D cycles along the filtration, we focus on 1D cycles that are created and closed only by fully coherent structures, i.e. edges and triangles having a weight larger than zero, which we denote as a Full Coherence (FC) contribution; 1D cycles formed by coherent structures and closed by the decoherent ones (i.e. edges and triangles with a weight smaller than zero), which we denote as a Coherence Transition (CT) contribution; finally, 1D cycles created only by the fully decoherent structures, we denote as a Full Decoherence (FD) contribution. Clearly, by construction, the sum of these three contributions sums up to the total hyper complexity. We show an illustrative example in SI Fig. S1.

In Fig. 4b we plot the three fractional contributions to the hyper complexity indicator in a triangular representation. In this space, a point is placed on the bottom-left corner if all the 1D cycles are formed and closed only by fully coherent structures. Likewise, the bottom-right corner corresponds to an exclusive contribution from fully decoherent structures, and the top corner corresponds to a contribution uniquely determined by the coherence transition. Whenever the hyper coherence indicator splits into similar FC and FD contributions, the point is placed between the corresponding corners, so that its position reflects the relative importance of the contributions. For example, a point would be at the centre of the triangle if the hyper complexity indicator is split into three equal contributions of FC-CT-FD. Note that such decomposition carries completely different information with respect to the hyper coherence indicator, and yet we draw similar analogies to the results just presented.

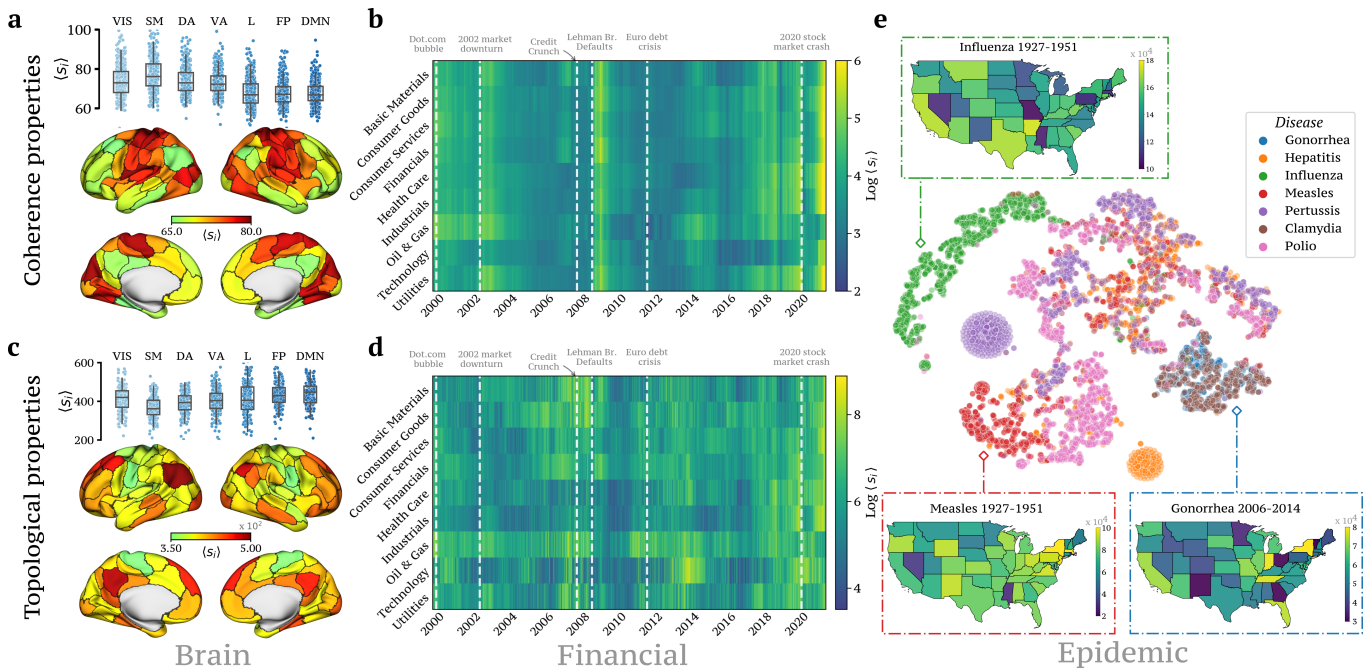


Figure 5. **Projection of higher-order measures provides local spatiotemporal information.** The nodal strength extracted from the violating triangles Δ_v can be used to track the importance of higher-order structures in time. (a) Brain map of the nodes involved in higher-order co-fluctuations obtained when isolating the top 15% coherent frames, which are the ones associated with a more synchronized dynamics during rest. We also show the boxplots reporting the mean coherence within the Yeo seven canonical functional networks [81], namely, the Visual (VIS), SomatoMotor (SM), Dorsal-Attention (DA), Ventral-Attention (VA), Limbic (L), FrontoParietal (FP), and Default Mode Network (DMN). (b) The temporal evolution of the nodal strength of violating triangles Δ_v at the level of industrial sectors discriminates crises from periods of financial stability. (c) The brain map obtained when selecting the 15% low-hyper complexity frames mainly encompasses the Default Mode Network, as confirmed by its high mean nodal score. (d) The temporal evolution of the nodal strength of the homological scaffold provides finer details on the downturns of certain economic sectors. Note that, from a topological perspective, the nodal strength extracted from the homological scaffold provides information about 1D loops in the space of co-fluctuations. (e) Finally, a planar embedding of the historical data of epidemic outbreaks can be obtained through t-SNE nonlinear dimensionality reduction when considering as features the higher-order indicators, such as hyper coherence, and the average edge violation (see Method for definition). As inset plots, we report the nodal strength of the violating triangles Δ_v at the US-state level when selecting the 15% high-coherent frames. Remarkably, the spatiotemporal evolution of the outbreaks is different across states and diseases.

Indeed, when examining the different contributions of the hyper complexity indicator in synthetic signals, we find that the five CML regimes appear to be separated in different clusters. Partially synchronized signals are characterized by a mixture of Fully Coherence and Coherence Transition contributions, while chaotic signals are mainly determined by Fully Coherence and Fully De-coherence (see Fig. 4b left panel). In comparison, for the human brain at rest, we find that most of the states are positioned between chaotic and partially synchronized regimes. This is in agreement with the results obtained when considering the hyper coherence indicator, which provides information of different nature, i.e. it is only based on the number of simplicial violations.

Real-world applications of higher-order topological markers

So far, we have mainly focused on the temporal evolution of our global higher-order indicators in synthetic and real-world multivariate time series. In what fol-

lows, we report some representative applications when considering higher-order measures on a more local level. Our goal is to characterize the higher-order states with the largest level of synchronization in both resting-state brain data and financial systems. To this end, in the context of the human brain, we isolated the top 15% coherent frames, which are those associated with a more synchronized dynamical phase. In Fig. 5a we report a brain map of the most discriminative nodes by projecting the magnitudes of the violating triangles Δ_v on a nodal level (see Methods for details and SI Fig. S15 for comparisons at other peaks percentages). This is equivalent to considering the nodal strength extracted from the list of violating triangles Δ_v . In other words, regions with the highest absolute value are the ones belonging to the most coherent higher-order structures. In particular, we find activity patterns with emphasized synchronized co-fluctuations mainly reflect sensorimotor areas, which belong to one of the well-known substrates present in

the resting-state network [86]. This is confirmed when considering the histogram reporting the mean coherence within the seven canonical functional networks [81] (see also SI Fig. S16 for the effects of higher-order indicators between the functional networks).

In Fig. 5b we report the temporal evolution of the nodal strength extracted from the list of violating triangles Δ_v , yet aggregated at the level of industrial sectors, for the financial time series. The highest values capture the onset of the major periods of financial instability (2002, corresponding to the market downturn, and 2007–2008, corresponding to the great recession that took place as a consequence of the subprime mortgage crisis), which are characterized by an increased synchronization of stock prices, which clearly distinguishes them from the unsynchronized intervals 2002–2007 and 2013–2018, which in turn corresponds to a more stable period of the economy.

Similar analyses can be produced by focusing on the hyper complexity indicator and the nodal strength of the homological scaffold constructed from the persistent homology generators of H_1 (see Methods for details). In particular, Fig. 5c depicts the brain map obtained when isolating the 15% low-hyper complexity frames, which, as previously shown, are the ones associated with a more synchronized dynamical phase. Here, the highest absolute values are the ones associated with the Default Mode Network (DMN), which is known for being the most active network during wakeful rest [87, 88].

By contrast, for the financial time series in Fig. 5d, the temporal evolution of the nodal strength of the homological scaffold provides fine details on the downturns of certain economic sectors. For instance, consumer goods, basic materials, as well as oil and gas, are the main sectors affected by the great recession of 2007.

Finally, by analysing the historical data of epidemic outbreaks in the US, we show that the temporal evolution of the higher-order measures (i.e. hyper coherence, the three contributions of hyper complexity, and the average edge violations; see Methods for definition) can be used to classify different infectious diseases. In particular, a support vector machine (SVM) classifier reports a high accuracy level, i.e. around 85 %, using a 10-fold cross-validation setting repeated 50 times (for a comparison between classifiers see SI Table S2). To provide a more intuitive representation of this result, we report in Fig. 5e a planar embedding of the historical data of epidemic outbreaks obtained using the t-distributed Stochastic Neighbor Embedding (t-SNE) nonlinear dimensionality reduction method. Note that nonlinear methods, such as t-SNE, allow to preserve the “local” structure in the original high-dimensional space after projection into the low-dimensional space, which is typically not possible with linear methods like Principal Component Analysis (PCA) or Multidimensional Scaling (MDS) [89]. In this space, we observe that diseases of different kinds cluster together to a great extent, somehow reflecting the unique spatiotemporal evolution of the outbreaks, which are in-

deed captured by the SVM classifier. At the same time, similarities between diseases can be observed. This is the case for sexually transmitted diseases, such as gonorrhoea and chlamydia, which are mostly overlapping in the planar embedding. As inset plots, we also report the map at the US-state level obtained when selecting the 15% high-coherent frames and considering the nodal strength of the violating triangles Δ_v . We find that the spatiotemporal evolution of the outbreaks is different across states and diseases, somehow reflecting the unique “higher-order” characteristics of the disease.

DISCUSSION

Inferring the dynamics of higher-order structures in multivariate time series is of utmost importance in many complex systems, from epidemiological, to financial, to biological systems. However, direct higher-order network measurements are often inaccessible [18]. As a matter of fact, the vast majority of complex spatiotemporal activity patterns commonly found in many biological, social, and financial systems are typically recorded on a nodal level, rather than directly measured at the level of edges or groups. The higher-order approach introduced in this work provides the first powerful and alternative method to dynamically reconstruct higher-order interactions from multivariate time series.

As a starting benchmark, we have first validated our method against signals whose underlying dynamics is well known. In particular, differently from various lower-order statistics [47, 54], the global higher-order indicators presented in this work are able to robustly separate several dynamical phases in high-dimensional coupled chaotic maps, which appear to be distinguishable only through methods based on higher-order statistics [43] (see Fig. 3 and SI Figures S6, S7). This provides further empirical evidence on the need of higher-order approaches for identifying higher-order behaviours [39]. On a more local level, i.e. when projecting the list of hyper coherent triangles as a weighted graph, the graph weight distribution reflects the global dynamic of the multivariate time series: synchronized periodic series convert into well-peaked distributions, while chaotic series convert into fat-tailed distributions. Armed with these theoretical foundations, we then applied our framework to real-world multivariate time series, specifically resting-state fMRI signals from 100 unrelated human subjects, prices of financial assets in the New York Stock Exchange, and historical data from 7 different epidemic outbreaks.

We found that, during rest, the human brain higher-order dynamics mainly oscillates between fully developed turbulence and partial synchronization. This is in agreement with recent studies supporting that the human brain operates in a turbulent regime [82, 90], at the edge of criticality [91], which seems to confer significant information processing advantages [92]. Moreover, when analysing brain states on a finer scale, we found two no-

table aspects. On the one hand, the maximally coherent higher-order structures reflect sensorimotor areas, which belong to one of the well-known substrates present in the resting-state brain network architecture [86]. These regions are known to play a major role in deciphering — over very fast time scales [55] — inputs that constantly change due to the external environment [83]. On the other hand, when examining the hyper complexity marker at its lowest points, we found that the nodal projection coarsely captures the Default-Mode Network (DMN), which is known to integrate high and low-order information in human brain networks [87, 88]. Hence, our two proposed markers provide complementary insights, which are not trivially deducible from an edge-wise approach (see SI Section S5 and SI Figures S12, S13), on how the brain network segregates and integrates higher-order information over time: hyper synchronous, less integrated, as measured by the number of simplicial violations, for sensorimotor regions; more integrated within the system, as measured in terms of “proper” higher-order simplices, for the DMN, whose brain dynamics acts mainly inward, in a constant state of internal exploration through the integration of low and higher-order dynamics [55, 93], at the “edge of instability” [84].

In the context of financial time series, instead, we provided evidence that the magnitude of higher-order structures efficiently discriminate crises from periods of financial stability, which cannot be obtained from different null models (see SI Section S4.2). In particular, maximally coherent higher-order structures strongly emerge in correspondence of major financial crises, mirroring the increase of synchronous co-activation patterns (i.e. stock prices tend to move to the same direction, therefore increasing their level of synchronization). While this is not new in the literature [30, 72, 94], we stress that, unlike our method, most of these approaches rely on correlation matrices estimated over sliding time-windows [29], therefore neglecting the information that one might want to capture at the level of individual frames (e.g. in high-frequency trading [95]). When examining the hyper complexity indicator at its lowest points, topological markers capture refined information regarding the different role of industrial sectors during crises, revealing strong variations in time and high heterogeneity across different industries [96], suggesting their potential to identify the building up of systemic risk [97, 98].

Finally, when analysing historical data of epidemic outbreaks in the US, we have shown that the temporal evolution of our higher-order measures can be used to classify, to a great extent, diseases of different kind. In particular, a planar embedding of diseases revealed the presence of interesting clusters based on their unique spatiotemporal pattern. While this result is interesting per se across disciplines [99–102], our higher-order markers may provide new tools for the quest of epidemic outbreak predictability [77, 103, 104], despite its limitation when increasing forecast length [77, 105].

Taken together, here we have developed a new flex-

ible tool to provide framewise estimates (therefore circumventing the limitations of sliding window approaches [106, 107]) of higher-order structures in multivariate time-series. We believe that our framework can be effectively used in all situations where the dynamics of signals is poorly understood or unknown, paving the way to further applications in the fields of biology, fluid dynamics, social sciences, or clinical neuroscience. In particular, the higher-order indicators and the corresponding lower-order projections can also provide topological *Polaroids*, that is, instantaneous topological snapshots of the spatial configuration of the system under study. Overall, our approach suggests that investigating the higher-order structure of multivariate time series might provide new insights compared to standard methods, allowing to better characterize group dependencies inherent to real-world data.

METHODS

Higher-order topology of multivariate time series — Let us consider a N -dimensional real valued time series $\{\mathbf{x}(t)\}_{i=1}^N$ with T time points, where the generic time series $\mathbf{x}_i = [x_i(1), x_i(2), \dots, x_i(T)]$ is usually measured empirically or extracted from a N -dimensional deterministic/stochastic dynamical system. It is well established that it is possible to construct correlation matrices by estimating the statistical dependency between every pair of time series [41, 57]. Here, the magnitude of that dependency is usually interpreted as a measure of how strongly (or weakly) those two time series are related to each other. Following the edge-centric approach proposed in Ref. [53], however, it is possible to estimate the instantaneous co-fluctuation magnitude between a pair of time series \mathbf{x}_i and \mathbf{x}_j —once they have been z -scored— by estimating their element-wise product. That is, for every pair of time series, a new time series encodes the magnitude of co-fluctuation between those signals resolved at every moment in time. We generalise such a concept to the case of higher-order interactions, i.e. triangles, tetrahedron, etc. We first z -score each original time series \mathbf{x}_i , such that $\mathbf{z}_i = \frac{\mathbf{x}_i - \mu[\mathbf{x}_i]}{\sigma[\mathbf{x}_i]}$, where $\mu[\bullet]$ and $\sigma[\bullet]$ are the time-averaged mean and standard deviation. We can then calculate the generic element at time t of the z -scored k -order co-fluctuations between $(k + 1)$ time series as

$$\xi_{0\dots k}(t) = \frac{\prod_{p=0}^k z_p(t) - \mu\left[\prod_{p=0}^k \mathbf{z}_p\right]}{\sigma\left[\prod_{p=0}^k \mathbf{z}_p\right]},$$

where also in this case $\mu[\bullet]$ and $\sigma[\bullet]$ are the time-averaged mean and SD functions. In order to differentiate concordant group interactions from discordant ones in a k -order product, concordant signs are always positively mapped,

while discordant signs are negatively mapped. Formally,

$$\text{sign}[\xi_{0\dots k}(t)] := (-1)^{\text{sgn}[(k+1) - |\sum_0^k \text{sgn}[z_i(t)]|]},$$

where $\text{sgn}[\bullet]$ is the signum function of a real number. In other words, the weight $w_{0\dots k}(t)$ at time t of the k -order co-fluctuations is defined as:

$$w_{0\dots k}(t) = \text{sign}[\xi_{0\dots k}(t)]|\xi_{0\dots k}(t)|$$

If we compute all the possible products up to order k , this will result in $\binom{N}{k}$ different co-fluctuation time series for each order k .

For each time t , we condense all the different k -order co-fluctuations into a weighted simplicial complex \mathcal{K}^t . Formally, a $(d-1)$ -dimensional simplex σ is defined as the set of d vertices, i.e. $\sigma = [p_0, p_1, \dots, p_{d-1}]$. A collection of simplices is a simplicial complex \mathcal{K} if for each simplex σ all its possible subfaces (defined as subsets of σ) are themselves contained in \mathcal{K} [60]. Weighted simplicial complexes are simplicial complexes with assigned values (called weights) on the simplices.

For simplicity, in this work we only consider co-fluctuations of dimension up to $k=2$, so that triangles represent the only higher-order structures in the weighted simplicial complex \mathcal{K} , and weights on the simplices, i.e. w_{ij} and w_{ijk} , represent the magnitude of edges and triangles co-fluctuations.

Note finally that, in order to compare our approach with the edge-based approach, we employed the Root Sum Square (RSS) of the edge-time series Ref. [54], which can be used as a direct proxy of the amplitude of the collective co-fluctuations of the edge time series. In other words, we compute the amplitude of the edge time series as the root sum of squared co-fluctuations, i.e. $RSS(t) = \sqrt{\sum_{i,j>i} e_{ij}(t)^2}$, where the vector $\mathbf{e}_{ij} = \mathbf{z}_i \mathbf{z}_j$ is the 1-order co-fluctuation (i.e. the edge time series) obtained as a product of the z-scores of the original time series.

Hyper coherence and hyper complexity — To analyse the structure of the weighted simplicial complex \mathcal{K}^t across multiple scales, we consider a topological data analysis approach [56], which has been shown to unveil new dynamical properties of different complex systems [108–111]. In particular, we rely on persistent homology, which is a recent technique in computational topology that has been largely used for the analysis of high dimensional datasets [64, 112] and in disparate applications [113–115]. The central idea is the construction of a sequence of successive simplicial complexes, which approximates with increasing precision the original weighted simplicial complex. This sequence of simplicial complexes, i.e. $\emptyset = \mathcal{S}_0 \subset \mathcal{S}_1 \subset \dots \subset \mathcal{S}_l \subset \dots \subset \mathcal{S}_n$, is such that $\mathcal{S}_i \subset \mathcal{S}_j$ whenever $i < j$ and is called a filtration. In our case, we construct a filtration building upon these steps:

- Sort the weights of the links and triangles in a decreasing order: the parameter $\epsilon_l \in \mathbb{R}$ scans the se-

quence. Equivalently, ϵ_l is the parameter that keeps track of the actual weight as we gradually scroll the list of weights.

- At each step l , remove all the triangles that do not satisfy the simplicial closure condition, i.e. $\exists!(i, j) : w_{ij} < w_{ijk}$. Such triangles are considered as a violation and inserted, along with the corresponding weights, in the *list of violations* $\Delta_v = \{(i, j, k), w_{ijk}\}$. The remaining links and triangles with a weight larger than ϵ_l belong to the simplicial complex \mathcal{S}_l

We then define the *hyper coherence indicator*, as the fraction of violating coherent triangles (i.e. violating triangles with a weight greater than zero) over all the possible coherent triangles (i.e. triangles with a weight greater than zero). Notice also that when identifying each violating triangle (i.e. by checking whether the triangle is entering the complex before its edges), we can keep track of the number of its edges $e_v \in [0, 2]$ that are already in the complex. We can then define the *average edge violation* indicator as the total number of those edges e_v averaged over all the violating triangles.

Persistent homology studies the changes of the topological structure along the filtration $\{\mathcal{S}_l\}$ and provides a natural measure of robustness for the topological features emerging across different scales. In particular, it is possible to keep track of these topological changes by looking at each k -dimensional cycle in the homology group H_k . In our case, we focus on the 1-dimensional holes (i.e. loops), therefore analysing the homology group H_1 . More precisely, at each step of the filtration process, a generator g uniquely identifies a 1-dimensional cycle by its constituting elements. The importance of the 1-dimensional hole g is encoded in the form of “time-stamps” recording its birth b_g and death d_g along the filtration $\{\mathcal{S}_l\}$ [11]. These two time-stamps can be combined to define the persistence $\pi_g = d_g - b_g$ of the one-dimensional cycle, which gives a notion of its importance in terms of its lifespan.

A typical way to visualize the results of persistent homology group H_1 is through multiset points in the two-dimensional persistence diagram. In this diagram, each point (b_g, d_g) represents a one-dimensional hole g that appears across the filtration. As a consequence, this diagram is a compressed summary describing how long 1D cycles live along the filtration and can be used as a proxy of the “complexity” of the underlying space. In fact, the sum of the persistences of the homological generators of H_k can be seen as the distance of the topological space from the trivial space (i.e. the space without $k+1$ -dimensional holes). In our case, we define the *hyper complexity indicator* as the Wasserstein distance [65] between the persistence diagram of H_1 and the empty persistence diagram, corresponding to a space with trivial H_1 homology. Finally, note that in SI Section S2.2 and SI Fig. S4, we briefly investigate the presence of 1D cycles and 3D-cavities in the context of CML when ex-

tending our framework to 4-body interactions.

Homological scaffold and lower-order projections —

To obtain a finer description of the topological features present in the persistent diagram, we consider the persistence homological scaffold as proposed in Ref. [11]. In a nutshell, this object is a weighted network composed of all the cycle paths corresponding to generators g_i weighted by their persistence π_{g_i} . In other words, if an edge e belongs to multiple 1-dimensional cycles g_0, g_1, \dots, g_s , its weight \bar{w}_e^π is defined as the sum of the generators' persistence, i.e.:

$$\bar{w}_e^\pi = \sum_{g_i | e \in g_i} \pi_{g_i}$$

The information provided by the homological scaffold allow us to decipher the role that different links have regarding the homological properties of the system. A large total persistence \bar{w}_e^π for a link e implies that such link acts as a locally strong bridge in the space of coherent and decoherent co-fluctuations [11].

Lastly, to analyse the information provided by the list of violations Δ_v on an edge/node level, we rely on downward projections. That is, for each edge (i, j) we assign a weight w_{ij} equal to the average sum of the weights of triangles defined by that edge, i.e. triangles of the form (i, j, \bullet) with a weight $w_{ij\bullet}$, and the average is computed over the number of triangles $n_{ij\bullet}$ defined by that edge. Similarly, we define the nodal strength w_i of node i as the average sum of weights of the triangles connected to node i . In the case of the homological scaffold, since it is a weighted network, the node strength \bar{w}_i of node i is defined, in the classical way [8, 116], as the sum of the weights of edges connected to the node i .

Real-world dataset — We analysed three datasets belonging to different domains. Specifically, we considered fMRI resting-state data from The Human Connectome Project (HCP, <http://www.humanconnectome.org/>), the stock prices of the NYSE financial market obtained from the Yahoo! finance API [117], and the historical data of several infectious diseases in the US [77, 78].

The fMRI dataset used in this work consists of resting-state data from 100 unrelated subjects (54 females, 46 males, mean age = 29.1 ± 3.7 years) as provided at the HCP 900 subjects data release Van Essen *et al.* [76, 118]. We added some extra steps to the HCP minimal preprocessing pipeline [80, 119]: First, we applied a standard general linear model regression that included detrending and removal of quadratic trends; removal of motion regressors and their first derivatives; removal of white matter, cerebrospinal fluid signals, and their first derivatives;

and global signal regression (and its derivative). Second, we bandpass-filtered the time series in the range of 0.01 to 0.15 Hz. Last, the voxel-wise fMRI time series were averaged into the $N = 100$ corresponding brain nodes of the Schaefer cortical atlas [79] and then z-scored. For completeness, 19 sub-cortical regions were added, as provided by the HCP release [80]. The interested reader can refer to Ref. [55] for details on these steps.

The financial dataset used in this study was obtained from the Yahoo! finance historical data API (via the Python library *yfinance* [117]). We have collected the daily prices of 119 US companies in the NYSE from Yahoo! finance in the period from January 1, 2000 to November 30, 2021.

We considered the weekly historical data at the US state-level of several infectious diseases including chlamydia, gonorrhea, influenza, measles, polio, and pertussis. This dataset was previously used in Ref. [77] and is freely available.

Limitations — One of the main limitations of our approach concerns the time complexity. Indeed, if we consider co-fluctuation patterns up to the order k , the resulting time complexity scales as $\mathcal{O}(N^k)$. Moreover, at the current stage, our framework does not allow to investigate the causality effect between two subsequent time frames (i.e. how much previous time points affect the next ones in terms of the proposed topological markers). Notice also that our dynamical higher-order approach, as many of other existing pairwise dynamical methods [48, 49, 54], can be heavily affected by noisy fluctuations in the time series. However, this issue can be smoothed out by analyzing statistics averaged over multiple time frames, as we have done in this work. Furthermore, the higher-order brain maps reported in Figure 5 appear to be robust after accounting for the presence of head motion volumes in the fMRI data (see also SI Fig. 14). Finally, we stress that our framework, with the exception of the hyper complexity indicator, mainly detect coherent synchronous patterns while it mostly ignores the effect of decoherent patterns, which are known to be important in the overall dynamics of a system. Future work should explore alternative approaches that deal in a more explicit way with decoherent patterns present in the data.

Data availability — The data used for this analysis will be available on Zenodo upon acceptance of the manuscript.

Code availability — The python code used in this work will be made available upon acceptance of the manuscript on AS EPFL webpage, as well as a maintained version on E.A.'s GitHub page (<https://github.com/eamico>).

[1] R. Albert and A.-L. Barabási, “Statistical mechanics of complex networks,” *Reviews of Modern Physics* **74**, 47–97 (2002).

[2] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, “Complex networks: Structure and dynamics,” *Physics Reports* **424**, 175–308 (2006).

- [3] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, “Epidemic processes in complex networks,” *Reviews of Modern Physics* **87**, 925–979 (2015).
- [4] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, “Synchronization in complex networks,” *Physics Reports* **469**, 93–153 (2008).
- [5] A. Barrat, M. Barthélemy, and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, 2008).
- [6] D. J. Watts and P. S. Dodds, “Influentials, Networks, and Public Opinion Formation,” *Journal of Consumer Research* **34**, 441–458 (2007).
- [7] A.-L. Barabási, *Network Science* (Cambridge University Press, 2016).
- [8] V. Latora, V. Nicosia, and G. Russo, *Complex Networks: Principles, Methods and Applications* (Cambridge University Press, 2017).
- [9] M. Newman, *Networks* (Oxford University Press, 2018).
- [10] A. R. Benson, D. F. Gleich, and J. Leskovec, “Higher-order organization of complex networks,” *Science* **353**, 163–166 (2016).
- [11] G. Petri, P. Expert, F. E. Turkheimer, R. L. Carhart-Harris, D. Nutt, P. Hellyer, and F. Vaccarino, “Homological scaffolds of brain functional networks,” *Journal of the Royal Society Interface* **10**, 186–198 (2014).
- [12] C. Giusti, E. Pastalkova, C. Curto, and V. Itskov, “Clique topology reveals intrinsic geometric structure in neural correlations,” *Proceedings of the National Academy of Sciences* **112**, 13455–13460 (2015).
- [13] A. E. Sizemore, C. Giusti, A. Kahn, J. M. Vettel, R. F. Betzel, and D. S. Bassett, “Cliques and cavities in the human connectome,” *Journal of Computational Neuroscience* **44**, 115–145 (2018).
- [14] J. Grilli, G. Barabás, M. J. Michalska-Smith, and S. Allesina, “Higher-order interactions stabilize dynamics in competitive network models,” *Nature* **548**, 210–213 (2017).
- [15] A. Sanchez-Gorostiaga, D. Bajić, M. L. Osborne, J. F. Poyatos, and A. Sanchez, “High-order interactions distort the functional landscape of microbial consortia,” *PLOS Biology* **17**, e3000550 (2019).
- [16] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J. G. Young, and G. Petri, “Networks beyond pairwise interactions: Structure and dynamics,” *Physics Reports* **874**, 1–92 (2020).
- [17] L. Torres, A. S. Blevins, D. Bassett, and T. Eliassirad, “The Why, How, and When of Representations for Complex Systems,” *SIAM Review* **63**, 435–485 (2021).
- [18] F. Battiston, E. Amico, A. Barrat, G. Bianconi, G. Ferraz de Arruda, B. Franceschiello, I. Iacopini, S. Kéfi, V. Latora, Y. Moreno, M. M. Murray, T. P. Peixoto, F. Vaccarino, and G. Petri, “The physics of higher-order interactions in complex systems,” *Nature Physics* **17**, 1093–1098 (2021).
- [19] F. Battiston and G. Petri, *Higher-Order Systems* (Springer, 2022).
- [20] A. P. Millán, J. J. Torres, and G. Bianconi, “Synchronization in network geometries with finite spectral dimension,” *Physical Review E* **99**, 022307 (2019).
- [21] M. T. Schaub, A. R. Benson, P. Horn, G. Lippner, and A. Jadbabaie, “Random Walks on Simplicial Complexes and the Normalized Hodge 1-Laplacian,” *SIAM Review* **62**, 353–391 (2020).
- [22] T. Carletti, F. Battiston, G. Cencetti, and D. Fanelli, “Random walks on hypergraphs,” *Physical Review E* **101**, 022308 (2020).
- [23] I. Iacopini, G. Petri, A. Barrat, and V. Latora, “Simplicial models of social contagion,” *Nature Communications* **10**, 1–9 (2019).
- [24] G. F. de Arruda, G. Petri, and Y. Moreno, “Social contagion models on hypergraphs,” *Physical Review Research* **2**, 023032 (2020).
- [25] R. Sahasrabudde, L. Neuhäuser, and R. Lambiotte, “Modelling non-linear consensus dynamics on hypergraphs,” *Journal of Physics: Complexity* **2**, 025006 (2021).
- [26] U. Alvarez-Rodriguez, F. Battiston, G. F. de Arruda, Y. Moreno, M. Perc, and V. Latora, “Evolutionary dynamics of higher-order interactions in social networks,” *Nature Human Behaviour* **5**, 586–595 (2021).
- [27] G. Deco, V. K. Jirsa, and A. R. McIntosh, “Emerging concepts for the dynamical organization of resting-state activity in the brain,” *Nature Reviews Neuroscience* **12**, 43–56 (2011).
- [28] A. Avena-Koenigsberger, B. Misic, and O. Sporns, “Communication dynamics in complex brain networks,” *Nature Reviews Neuroscience* **19**, 17–33 (2018).
- [29] R. N. Mantegna and H. E. Stanley, *Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, 1999).
- [30] T. K. D. Peron and F. A. Rodrigues, “Collective behavior in financial markets,” *EPL (Europhysics Letters)* **96**, 48004 (2011).
- [31] J. M. Olesen, J. Bascompte, H. Elberling, and P. Jordano, “Temporal dynamics in a pollination network,” *Ecology* **89**, 1573–1582 (2008).
- [32] J. Friedman, L. M. Higgins, and J. Gore, “Community structure follows simple assembly rules in microbial microcosms,” *Nature Ecology & Evolution* **1**, 1–7 (2017).
- [33] D. Ebert and P. D. Fields, “Host–parasite co-evolution and its genomic signature,” *Nature Reviews Genetics* **21**, 754–768 (2020).
- [34] I. Brugere, B. Gallagher, and T. Y. Berger-Wolf, “Network Structure Inference, A Survey: Motivations, Methods, and Applications,” *ACM Computing Surveys* **51**, 24:1–24:39 (2018).
- [35] J. G. Young, G. Petri, and T. P. Peixoto, “Hypergraph reconstruction from network data,” *Communications Physics* **4**, 1–11 (2021).
- [36] F. Musciotto, F. Battiston, and R. N. Mantegna, “Detecting informative higher-order interactions in statistically validated hypergraphs,” *Communications Physics* **4**, 1–9 (2021).
- [37] H. Wang, C. Ma, H.-S. Chen, Y.-C. Lai, and H.-F. Zhang, “Full reconstruction of simplicial complexes from binary time-series data,” (2022).
- [38] S. Lizotte, J.-G. Young, and A. Allard, “Hypergraph reconstruction from noisy pairwise observations,” arXiv preprint [arXiv:2208.06503](https://arxiv.org/abs/2208.06503) (2022).
- [39] F. E. Rosas, P. A. M. Mediano, A. I. Luppi, T. F. Varley, J. T. Lizier, S. Stramaglia, H. J. Jensen, and D. Marinazzo, “Disentangling high-order mechanisms and high-order behaviours in complex systems,” *Nature Physics* **18**, 476–477 (2022).
- [40] E. Bullmore and O. Sporns, “Complex brain networks: Graph theoretical analysis of structural and functional systems,” *Nature Reviews Neuroscience* **10**, 186–198

- (2009).
- [41] O. Sporns, *Networks of the Brain* (MIT press, 2010).
- [42] L. Faes, G. Mijatovic, Y. Antonacci, R. Pernice, C. Barà, L. Sparacino, M. Sammartino, A. Porta, D. Marinazzo, and S. Stramaglia, “A Framework for the Time- and Frequency-Domain Assessment of High-Order Interactions in Brain and Physiological Networks,” arXiv:2202.04179 (2022).
- [43] F. E. Rosas, P. A. Mediano, M. Gastpar, and H. J. Jensen, “Quantifying high-order interdependencies via multivariate extensions of the mutual information,” *Physical Review E* **100**, 32305 (2019).
- [44] F. E. Rosas, P. A. Mediano, H. J. Jensen, A. K. Seth, A. B. Barrett, R. L. Carhart-Harris, and D. Bor, “Reconciling emergences: An information-theoretic approach to identify causal emergence in multivariate data,” *PLoS Computational Biology* **16**, 1–22 (2020), arXiv:2004.08220.
- [45] M. Gatica, R. Cofré, P. A. Mediano, F. E. Rosas, P. Orío, I. Diez, S. P. Swinnen, and J. M. Cortes, “High-Order Interdependencies in the Aging Brain,” *Brain Connectivity* **00**, 1–11 (2021).
- [46] S. Stramaglia, T. Scagliarini, B. C. Daniels, and D. Marinazzo, “Quantifying Dynamical High-Order Interdependencies From the O-Information: An Application to Neural Spiking Dynamics,” *Frontiers in Physiology* **11**, 1–11 (2021), arXiv:2007.16018.
- [47] M. MacMahon and D. Garlaschelli, “Community Detection for Correlation Matrices,” *Physical Review X* **5**, 021006 (2015).
- [48] E. Tagliazucchi, P. Balenzuela, D. Fraiman, and D. Chialvo, “Criticality in Large-Scale Brain fMRI Dynamics Unveiled by a Novel Point Process Analysis,” *Frontiers in Physiology* **3** (2012), 10/gf2vtk.
- [49] X. Liu and J. H. Duyn, “Time-varying functional network information extracted from brief instances of spontaneous brain activity,” *Proceedings of the National Academy of Sciences* **110**, 4392–4397 (2013).
- [50] F. I. Karahanoglu and D. Van De Ville, “Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks,” *Nature Communications* **6**, 7751 (2015).
- [51] M. G. Preti, T. A. Bolton, and D. Van De Ville, “The dynamic functional connectome: State-of-the-art and perspectives,” *NeuroImage Functional Architecture of the Brain*, **160**, 41–54 (2017).
- [52] X. Liu, N. Zhang, C. Chang, and J. H. Duyn, “Co-activation patterns in resting-state fMRI signals,” *NeuroImage Brain Connectivity Dynamics*, **180**, 485–494 (2018).
- [53] J. Faskowitz, F. Z. Esfahlani, Y. Jo, O. Sporns, and R. F. Betzel, “Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture,” *Nature Neuroscience* **23**, 1644–1654 (2020).
- [54] F. Z. Esfahlani, Y. Jo, J. Faskowitz, L. Byrge, D. P. Kennedy, O. Sporns, and R. F. Betzel, “High-amplitude co-fluctuations in cortical activity drive functional connectivity,” *Proceedings of the National Academy of Sciences* **117**, 28393–28401 (2020).
- [55] D. Van De Ville, Y. Farouj, M. G. Preti, R. Liégeois, and E. Amico, “When makes you unique: Temporality of the human brain fingerprint,” *Science Advances* **7**, eabj0751 (2021).
- [56] L. Wasserman, “Topological Data Analysis,” *Annual Review of Statistics and Its Application* **5**, 501–532 (2018).
- [57] W. W. Wei, *Time Series Analysis* (Addison Wesley, 2005).
- [58] W. W. Wei, *Multivariate Time Series Analysis and Applications* (John Wiley & Sons, 2019).
- [59] Y. Zou, R. V. Donner, N. Marwan, J. F. Donges, and J. Kurths, “Complex network approaches to nonlinear time series analysis,” *Physics Reports* **787**, 1–97 (2019).
- [60] A. Hatcher, *Algebraic Topology* (Cambridge University Press, 2005).
- [61] G. Petri, M. Scolamiero, I. Donato, and F. Vaccarino, “Topological Strata of Weighted Complex Networks,” *PLoS ONE* **8** (2013), 10/gm8bs6.
- [62] H. Edelsbrunner, D. Letscher, and A. Zomorodian, “Topological persistence and simplification,” in *Proceedings 41st Annual Symposium on Foundations of Computer Science* (2000) pp. 454–463.
- [63] A. Zomorodian and G. Carlsson, “Computing Persistent Homology,” *Discrete & Computational Geometry* **33**, 249–274 (2005).
- [64] R. Ghrist, “Barcodes: The persistent topology of data,” *Bulletin of the American Mathematical Society* **45**, 61–75 (2008).
- [65] M. Carrière, M. Cuturi, and S. Oudot, “Sliced Wasserstein Kernel for Persistence Diagrams,” in *Proceedings of the 34th International Conference on Machine Learning* (PMLR, 2017) pp. 664–673.
- [66] K. Kaneko, “Pattern dynamics in spatiotemporal chaos,” *Physica D: Nonlinear Phenomena* **34**, 1–41 (1989).
- [67] K. Kaneko, “Overview of coupled map lattices,” *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2**, 279–282 (1992).
- [68] M. Bevers and C. H. Flather, “Numerically Exploring Habitat Fragmentation Effects on Populations Using Cell-Based Coupled Map Lattices,” *Theoretical Population Biology* **55**, 61–76 (1999).
- [69] A. Hilgers and C. Beck, “Turbulent Behavior of Stock Exchange Indices and Foreign Currency Exchange Rates,” *International Journal of Bifurcation and Chaos* **07**, 1855–1859 (1997).
- [70] W. Huang and Z. Chen, “Heterogeneous agents in multi-markets: A coupled map lattices approach,” *Mathematics and Computers in Simulation Complex Dynamics in Economics and Finance - Papers of the MDEF 2012 Workshop*, **108**, 3–15 (2015).
- [71] K. Kaneko and I. Tsuda, *Complex Systems: Chaos and beyond: A Constructive Approach with Applications in Life Sciences* (Springer Science & Business Media, 2011).
- [72] L. Lacasa, V. Nicosia, and V. Latora, “Network structure of multivariate time series,” *Scientific Reports* **5**, 1–9 (2015).
- [73] K. O. McGraw and S. P. Wong, “Forming inferences about some intraclass correlation coefficients,” *Psychological Methods* **1**, 30–46 (1996).
- [74] P. E. Shrout and J. L. Fleiss, “Intraclass correlations: Uses in assessing rater reliability,” *Psychological Bulletin* **86**, 420–428 (1979).
- [75] Á. M. Núñez, B. Luque, L. Lacasa, J. P. Gómez, and A. Robledo, “Horizontal visibility graphs generated by type-I intermittency,” *Physical Review E* **87**, 052801 (2013).

- [76] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. J. Behrens, E. Yacoub, and K. Ugurbil, “The WU-Minn Human Connectome Project: An overview,” *NeuroImage Mapping the Connectome*, **80**, 62–79 (2013).
- [77] S. V. Scarpino and G. Petri, “On the predictability of infectious disease outbreaks,” *Nature Communications* **10** (2019), 10/ggrfdr.
- [78] W. G. van Panhuis, J. Grefenstette, S. Y. Jung, N. S. Chok, A. Cross, H. Eng, B. Y. Lee, V. Zadorozhny, S. Brown, D. Cummings, and D. S. Burke, “Contagious Diseases in the United States from 1888 to the Present,” *New England Journal of Medicine* **369**, 2152–2158 (2013).
- [79] A. Schaefer, R. Kong, E. M. Gordon, T. O. Laumann, X.-N. Zuo, A. J. Holmes, S. B. Eickhoff, and B. T. T. Yeo, “Local-Global Parcellation of the Human Cerebral Cortex from Intrinsic Functional Connectivity MRI,” *Cerebral Cortex* **28**, 3095–3114 (2018).
- [80] M. F. Glasser, S. N. Sotiropoulos, J. A. Wilson, T. S. Coalson, B. Fischl, J. L. Andersson, J. Xu, S. Jbabdi, M. Webster, J. R. Polimeni, D. C. Van Essen, and M. Jenkinson, “The minimal preprocessing pipelines for the Human Connectome Project,” *NeuroImage Mapping the Connectome*, **80**, 105–124 (2013).
- [81] B. T. Yeo, F. M. Krienen, J. Sepulcre, M. R. Sabuncu, D. Lashkari, M. Hollinshead, J. L. Roffman, J. W. Smoller, L. Zöllei, J. R. Polimeni, *et al.*, “The organization of the human cerebral cortex estimated by intrinsic functional connectivity,” *Journal of Neurophysiology* **106**, 1125–1165 (2011).
- [82] G. Deco and M. L. Kringelbach, “Turbulent-like Dynamics in the Human Brain,” *CellReports* **33**, 108471 (2020).
- [83] G. Deco, M. L. Kringelbach, V. K. Jirsa, and P. Ritter, “The dynamics of resting fluctuations in the brain: Metastability and its dynamical cortical core,” *Scientific Reports* **7**, 3095 (2017).
- [84] G. Deco, V. K. Jirsa, and A. R. McIntosh, “Resting brains never rest: Computational insights into potential cognitive architectures,” *Trends in Neurosciences* **36**, 268–274 (2013).
- [85] D. R. Chialvo, “Emergent complex neural dynamics,” *Nature Physics* **6**, 744–750 (2010).
- [86] S. M. Smith, P. T. Fox, K. L. Miller, D. C. Glahn, P. M. Fox, C. E. Mackay, N. Filippini, K. E. Watkins, R. Toro, A. R. Laird, and C. F. Beckmann, “Correspondence of the brain’s functional architecture during activation and rest,” *Proceedings of the National Academy of Sciences* **106**, 13040–13045 (2009).
- [87] M. E. Raichle, A. M. MacLeod, A. Z. Snyder, W. J. Powers, D. A. Gusnard, and G. L. Shulman, “A default mode of brain function,” *Proceedings of the National Academy of Sciences* **98**, 676–682 (2001).
- [88] M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, and M. E. Raichle, “The human brain is intrinsically organized into dynamic, anticorrelated functional networks,” *Proceedings of the National Academy of Sciences* **102**, 9673–9678 (2005).
- [89] L. Van der Maaten and G. Hinton, “Visualizing data using t-SNE.” *Journal of machine learning research* **9** (2008).
- [90] G. Deco, M. Kemp, and M. L. Kringelbach, “Leonardo da Vinci and the search for order in neuroscience,” *Current Biology* **31**, R704–R709 (2021).
- [91] Y. S. Perl, A. Escrichs, E. Tagliazucchi, M. L. Kringelbach, and G. Deco, “On the edge of criticality: Strength-dependent perturbation unveils delicate balance between fluctuation and oscillation in brain dynamics,” *BioRxiv*, 2021.09.23.461520 (2022).
- [92] G. Deco, Y. S. Perl, P. Vuust, E. Tagliazucchi, H. Kennedy, M. L. Kringelbach, G. Deco, Y. S. Perl, P. Vuust, E. Tagliazucchi, and H. Kennedy, “Rare long-range cortical connections enhance human information processing,” *Current Biology* **31**, 4436–4448.e5 (2021).
- [93] E. Amico, A. Arenas, and J. Goñi, “Centralized and distributed cognitive task processing in the human connectome,” *Network Neuroscience* **3**, 455–474 (2019).
- [94] R. Kutner, M. Ausloos, D. Grech, T. Di Matteo, C. Schinckus, and H. Eugene Stanley, “Econophysics and sociophysics: Their milestones & challenges,” *Physica A: Statistical Mechanics and its Applications* **516**, 240–253 (2019).
- [95] F. Musciotto, J. Pillo, and R. N. Mantegna, “High-frequency trading and networked markets,” *Proceedings of the National Academy of Sciences* **118**, e2015573118 (2021).
- [96] N. Musmeci, V. Nicosia, T. Aste, T. Di Matteo, and V. Latora, “The Multiplex Dependency Structure of Financial Markets,” *Complexity* **2017**, e9586064 (2017).
- [97] T. Squartini, I. van Lelyveld, and D. Garlaschelli, “Early-warning signals of topological collapse in interbank networks,” *Scientific Reports* **3**, 3357 (2013).
- [98] G. Iori and R. N. Mantegna, “Chapter 11 - Empirical Analyses of Networks in Finance,” in *Handbook of Computational Economics*, Vol. 4, edited by C. Hommes and B. LeBaron (Elsevier, 2018) pp. 637–685.
- [99] C. M. Bishop, *Pattern Recognition and Machine Learning*, Vol. 4 (Springer New York, 2006).
- [100] Z. Xing, J. Pei, and E. Keogh, “A brief survey on sequence classification,” *ACM SIGKDD Explorations Newsletter* **12**, 40–48 (2010).
- [101] D. Kobak and P. Berens, “The art of using t-SNE for single-cell transcriptomics,” *Nature Communications* **10**, 5416 (2019).
- [102] H. Peng, Q. Ke, C. Budak, D. M. Romero, and Y.-Y. Ahn, “Neural embeddings of scholarly periodicals reveal complex disciplinary organizations,” *Science Advances* **7**, eabb9004 (2021).
- [103] G. E. P. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung, *Time Series Analysis: Forecasting and Control* (John Wiley & Sons, 2015).
- [104] S. Pei, S. Kandula, W. Yang, J. Shaman, “Forecasting the spatial transmission of influenza in the United States,” *Proceedings of the National Academy of Sciences* **115**, 2752–2757 (2018).
- [105] J. D. Farmer and J. J. Sidorowich, “Predicting chaotic time series,” *Physical Review Letters* **59**, 845–848 (1987).
- [106] R. M. Hutchison, T. Womelsdorf, E. A. Allen, P. A. Bandettini, V. D. Calhoun, M. Corbetta, S. Della Penna, J. H. Duyn, G. H. Glover, J. Gonzalez-Castillo, D. A. Handwerker, S. Keilholz, V. Kiviniemi, D. A. Leopold, F. de Pasquale, O. Sporns, M. Walter, and C. Chang, “Dynamic functional connectivity: Promise, issues, and interpretations,” *NeuroImage Mapping the Connectome*, **80**, 360–378 (2013).
- [107] R. Hindriks, M. H. Adhikari, Y. Murayama, M. Ganzetti, D. Mantini, N. K. Logothetis, and G. Deco, “Can sliding-window correlations reveal dy-

- dynamic functional connectivity in resting-state fMRI?" *NeuroImage* **127**, 242–256 (2016).
- [108] P. Y. Lum, G. Singh, A. Lehman, T. Ishkanov, M. Vejdemo-Johansson, M. Alagappan, J. Carlsson, and G. Carlsson, "Extracting insights from the shape of complex data using topology," *Sci Rep* **3**, 1236 (2013).
- [109] M. Nicolau, A. J. Levine, and G. Carlsson, "Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival," *Proceedings of the National Academy of Sciences* **108**, 7265–7270 (2011).
- [110] M. Sagar, O. Sporns, J. Gonzalez-Castillo, P. A. Bandettini, G. Carlsson, G. Glover, and A. L. Reiss, "Towards a new approach to reveal dynamical organization of the brain using topological data analysis," *Nat Commun* **9**, 1399 (2018).
- [111] M. Sagar, J. M. Shine, R. Liégeois, N. U. F. Dosenbach, and D. Fair, "Precision dynamical mapping using topological data analysis reveals a hub-like transition state at rest," *Nat Commun* **13**, 4791 (2022).
- [112] G. Carlsson, T. Ishkhanov, V. de Silva, and A. Zomorodian, "On the Local Behavior of Spaces of Natural Images," *International Journal of Computer Vision* **76**, 1–12 (2008).
- [113] H. Lee, M. K. Chung, H. Kang, B.-N. Kim, and D. S. Lee, "Discriminative persistent homology of brain networks," in *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro* (2011) pp. 841–844.
- [114] C. J. Carstens and K. J. Horadam, "Persistent Homology of Collaboration Networks," *Mathematical Problems in Engineering* **2013**, e815035 (2013).
- [115] D. Horak, S. Maletić, and M. Rajković, "Persistent homology of complex networks," *Journal of Statistical Mechanics: Theory and Experiment* **2009**, P03034 (2009).
- [116] A. Barrat, M. Barthelemy, R. Pastor-Satorras, and A. Vespignani, "The architecture of complex weighted networks," *PNAS* **101**, 3747–3752 (2004).
- [117] R. Aroussi and al., "Yahoo! finance market data downloader," (2022).
- [118] D. C. Van Essen, K. Ugurbil, E. Auerbach, D. Barch, T. E. J. Behrens, R. Bucholz, A. Chang, L. Chen, M. Corbetta, S. W. Curtiss, S. Della Penna, D. Feinberg, M. F. Glasser, N. Harel, A. C. Heath, L. Larson-Prior, D. Marcus, G. Michalareas, S. Moeller, R. Oostenveld, S. E. Petersen, F. Prior, B. L. Schlaggar, S. M. Smith, A. Z. Snyder, J. Xu, and E. Yacoub, "The Human Connectome Project: A data acquisition perspective," *NeuroImage Connectivity*, **62**, 2222–2231 (2012).
- [119] S. M. Smith, C. F. Beckmann, J. Andersson, E. J. Auerbach, J. Bijsterbosch, G. Douaud, E. Duff, D. A. Feinberg, L. Griffanti, M. P. Harms, M. Kelly, T. Laumann, K. L. Miller, S. Moeller, S. Petersen, J. Power, G. Salimi-Khorshidi, A. Z. Snyder, A. T. Vu, M. W. Woolrich, J. Xu, E. Yacoub, K. Ugurbil, D. C. Van Essen, and M. F. Glasser, "Resting-state fMRI in the Human Connectome Project," *NeuroImage Mapping the Connectome*, **80**, 144–168 (2013).

Supplementary Material:

Unveiling the higher-order organization of multivariate time series

CONTENTS

S1.	Filtration and persistent homology	18
S2.	Global behaviour of Coupled Map Lattice (CML)	20
	S2.1. Effect of z -scores and signs assignment in CML time series	20
	S2.2. Impact of 3D-cavities	21
	S2.3. Higher-order approach on Gaussian multivariate time series	22
S3.	Comparison of time series approaches to distinguish the CML regimes	24
S4.	Impact of different null models	27
	S4.1. CML synthetic multivariate time series	27
	S4.2. Block null model in financial time series	30
S5.	Edge-based indicators in real-world systems	31
S6.	Effect of noisy fluctuations in brain data	35
S7.	Local higher-order indicators in real-world systems	36
	S7.1. Role of thresholds in brain maps	36
	S7.2. Higher-order indicators in the brain's functional networks	36
	S7.3. Comparison of classifiers for the US historical data	37

S1. FILTRATION AND PERSISTENT HOMOLOGY

In the main text, we introduced the simplicial filtration $\mathbb{F}(\mathcal{K}^t)$ as one of the main steps of our higher-order framework. More precisely, for each time t , this object represents a sequence of embedded simplicial complexes, sorted according to coherent patterns, starting with the empty complex and ending with the entire simplicial complex \mathcal{K}^t , i.e. $\emptyset = \mathcal{S}_0 \subset \mathcal{S}_1 \subset \dots \subset \mathcal{S}_l \subset \dots \subset \mathcal{S}_n \subset \mathcal{K}^t$. We then considered the persistent homology of \mathbb{F} to characterize the persistency of 1D cycles in the filtration, i.e. the persistent generators of the first homology group H_1 , which provide insights about where and when higher synchronised regions emerge.

In Fig. S1 we report a schematic representation of the persistent homology computation and the corresponding definition of the hyper complexity contributes considered in this work. In particular, Fig. S1a reports an illustrative example of a simplicial filtration in two-dimensions. The outputs of persistent homology are barcodes, which represent a compressed description of the homological features of a space (see Fig. S1b). Each bar corresponds to a specific topological feature, which can be identified in terms of a “time-stamp” recording the birth w_b and death w_d of that feature along the filtration. In the main text, we characterized the evolution of 1D cycles along such a filtration — the blue bars in Fig. S1b — by relying on persistence diagrams, which provide an equivalent description of bar codes. In this 2-dimensional plot, each 1D cycle is represented by a point with coordinates (w_b, w_d) (see Fig. S1c).

Furthermore, we defined the *hyper complexity indicator* as the Wasserstein distance [1] between the persistence diagram of H_1 and the empty persistence diagram, corresponding to a space with trivial H_1 homology. However, to better characterize the evolution of 1D cycles in the space of coherent and decoherent co-fluctuations, we decompose the hyper complexity indicator into three different contributions. That is, as we track the evolution of 1D cycles along the filtration, we focus on 1D cycles that are created and closed only by fully coherent structures, i.e. edges and triangles having a weight larger than zero, which we renamed as a Full Coherence (FC) contribution; 1D cycles formed by coherent structures and closed by the decoherent ones (i.e. edges and triangles with a weight smaller than zero), which we renamed as a Coherence Transition (CT) contribution; 1D cycles created only by the fully decoherent structures, renamed as a Full Decoherence (FD) contribution.

In Fig. S1d, the three fractional contributions of the hyper complexity indicator are reported in a triangular representation.

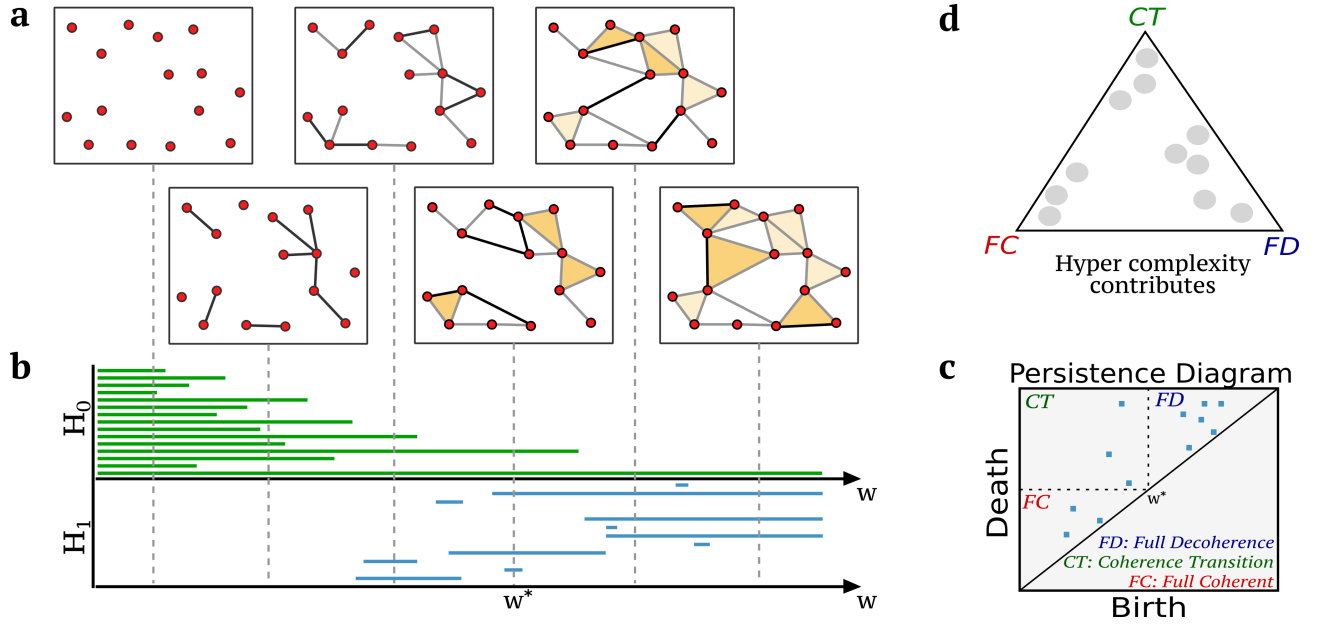


Figure S1. **Pictorial representation of persistent homology computation and definition of hyper complexity contributes.** (a) Example of a simplicial filtration in two-dimensions. (b) Barcodes describe the lifetime of different topological features across multiple scales. Here, green bars identify the persistence of various connected components (describing H_0), progressively merging into each other until only one survives, whereas blue bars describe the lifetime of 1-dimensional cycles (describing H_1). Hence, each bar corresponds to a specific topological feature, which can be identified in terms of a “time-stamp” recording the birth w_b and death w_d of that feature along the filtration. (c) Persistence diagrams provide an equivalent description of barcodes. For example, if we focus only on H_1 , each 1D cycle is represented in the 2-dimensional plot by a point with coordinates (w_b, w_d) . In this work, we further distinguish the nature of 1D cycles depending on the corresponding time-stamps, i.e. whether 1D cycles are created and closed before/after $w^* = 0$, therefore reflecting pure coherent (resp. incoherent) structures. (d) The hyper complexity indicator, defined as the Wasserstein distance [1] between the persistence diagram of H_1 and the empty persistence diagram, can be then decomposed into three different contributes according to the nature of 1D cycles and plotted in a triangular representation.

S2. GLOBAL BEHAVIOUR OF COUPLED MAP LATTICE (CML)

In the main text, we considered $N = 119$ diffusively coupled fully chaotic maps to generate synthetic multivariate time series with different dynamical regimes. In Fig. S2a we report a sample of such time series in three different lattice sites, showing the general behaviour of the five different dynamical states. Moreover, the space-amplitude plot [2, 3] reported in Fig. S2b, which depicts the temporal evolution of $x_i(t)$ at each lattice site i , clearly shows the overall behaviour of the five dynamical regimes.

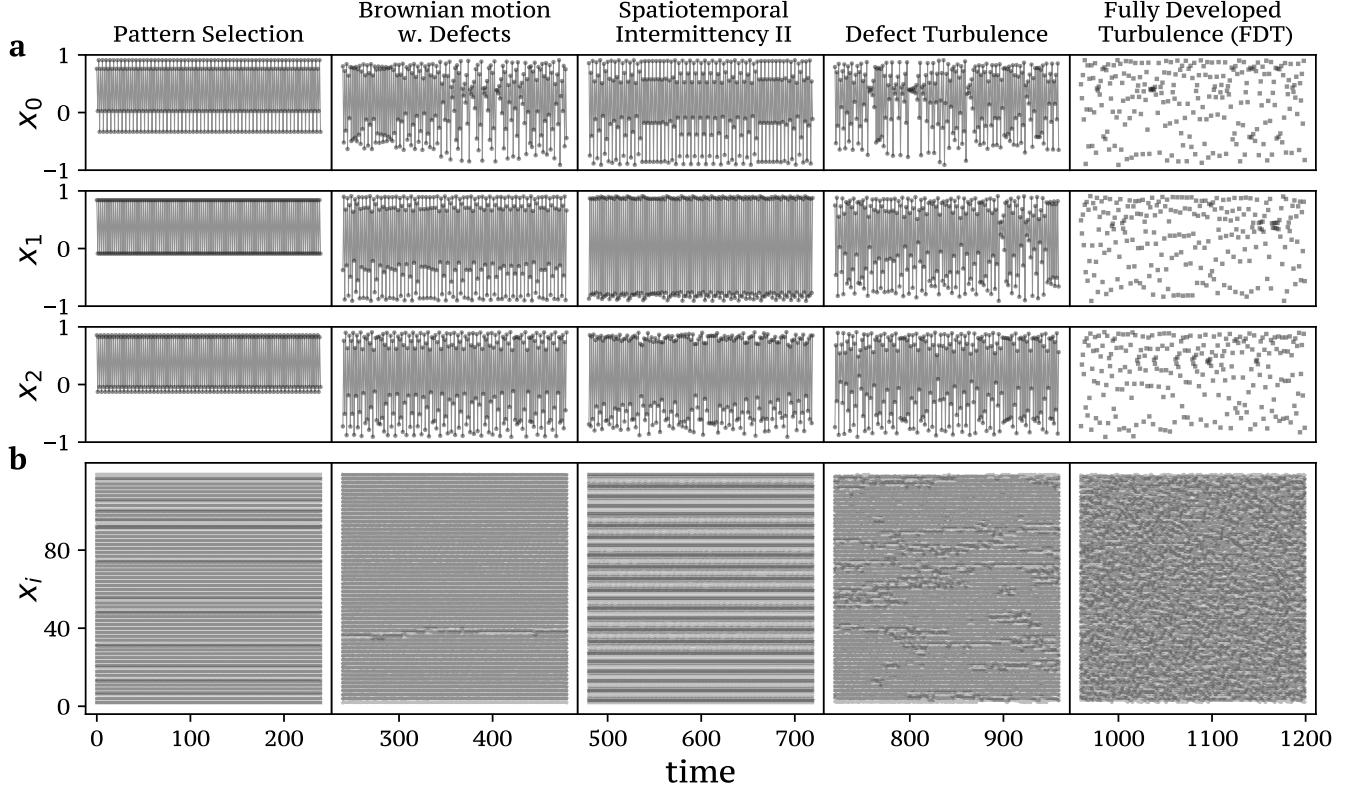


Figure S2. **Sample of the coupled map lattice** (a) We report a sample of the time series considered in the main text and generated by $N = 119$ diffusively coupled fully chaotic maps at different values of the coupling strength ε . Namely, from ordered to disordered systems: Pattern Selection (PS) at $\varepsilon = 0.12$, Brownian motion with Defects (BMWD) at $\varepsilon = 0.08$, Spatiotemporal Intermittency II (STI) at $\varepsilon = 0.3$, Defect Turbulence (DT) at $\varepsilon = 0.068$, and Fully Developed Turbulence (FDT) at $\varepsilon = 0.05$. (b) The space-amplitude plot, reporting the temporal evolution of $x_i(t)$ at each lattice site i , clearly reflects the different global behaviour of the five dynamical regimes.

S2.1. Effect of z -scores and signs assignment in CML time series

In one of the main steps of the simplicial framework described in the main text, we rely on the z -score for each group product time series to allow comparison across k -orders and to create proper simplicial filtrations. After that, a choice on how to assign signs to the resulting weights is required in order to distinguish fully concordant group interactions (all positive or negative fluctuations) from discordant ones (a mixture of positive and negative fluctuations) in a k -order product. In order to distinguish fully concordant group interactions (all positive or negative fluctuations) from discordant ones (a mixture of positive and negative fluctuations) in a k -order product, we assign positive signs to the fully concordant group interactions, and negative signs to the discordant ones. However, this way of assigning signs means that the baseline score might be slightly altered and, as a result, the concordance might be harder to achieve for a triangle than for an edge.

In Figure S3, we analyse the impact of z -score and signs assignment in edges and triangles distributions for the coupled

map lattice dynamical states. In particular, the z -scores and signs assignment for triangles produce a negative offset with respect to that of edges. This phenomenon does not have a deep impact in the hyper coherence indicator, since it is defined as the fraction of violating coherent triangles (i.e. violating triangles with a weight greater than zero) over all the possible coherent triangles (i.e. triangles with a weight greater than zero). By contrast, the negative offset of the triangle weight distribution over the edges one might have an impact in the hyper complexity indicator. This is due to the fact that this indicator is constructed by considering the ensemble of both edges and triangles weight distribution. However, the effect on hyper complexity is a shift in the contributions from the fully coherent to the other two contributions, rather than a restructuring of the whole simplicial filtration. In this sense, the effect is a biasing one.

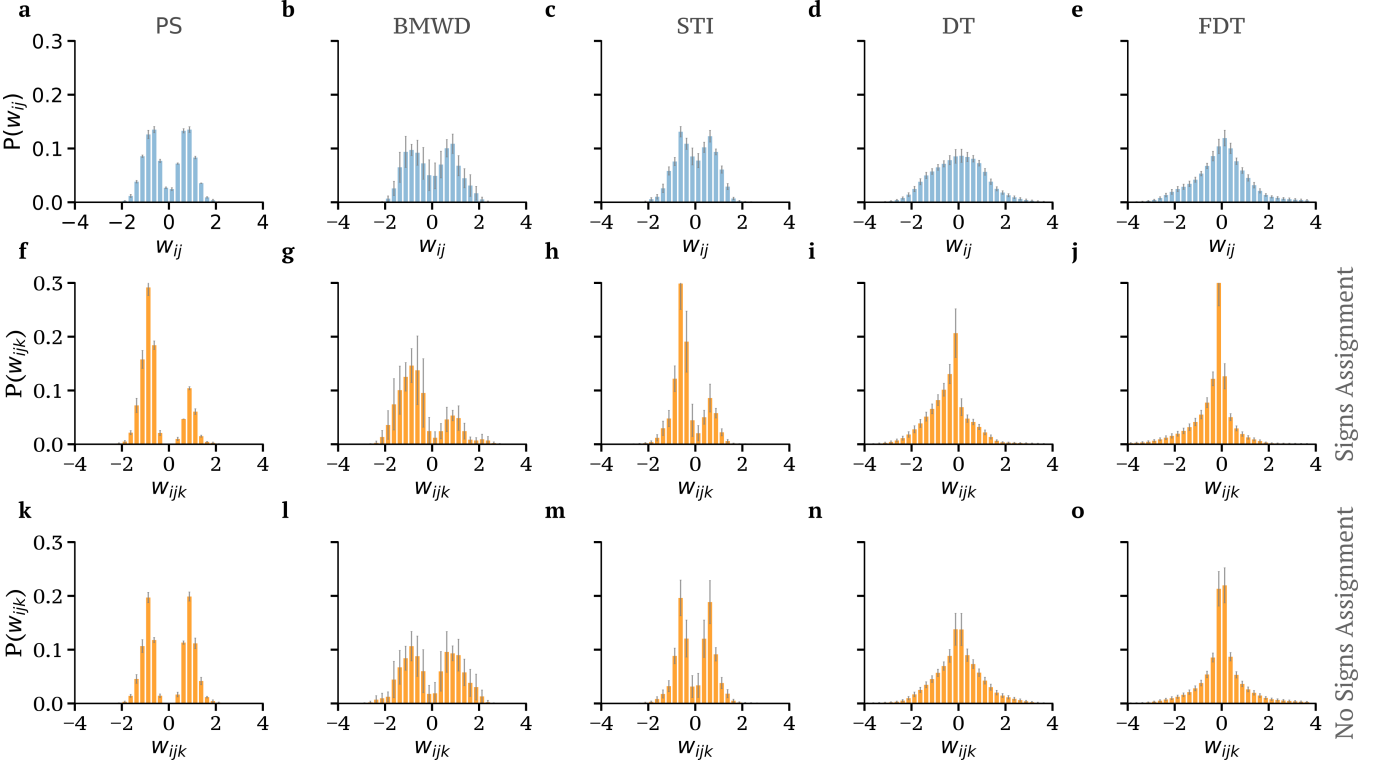


Figure S3. **Impact of z -score and signs assignment in edge and triangles distributions for the coupled map lattice.** For each of the five CML dynamical states, we report the distribution of weights of edges (a-e) and triangles with (f-j) and without the sign assignment (k-o), respectively. Notice that the z -scores and signs assignment produce a negative offset for triangles with respect to that of edges. While this phenomenon does not have a deep impact in the hyper coherence indicator for the way such a measure is constructed, the hyper complexity indicator might be shifted by such offset. It is also worth remarking that when sign remapping is not performed, it is not possible to identify a triangle with a synchronous co-activation pattern from its weight, given that decoherent triangles might end up having the same weights as fully coherent ones. Error bars represent standard deviations over 240 time points.

S2.2. Impact of 3D-cavities

In this section we briefly investigate the presence of 3D-cavities (i.e. the generators of the homology group H_2) when including in the analysis all the interactions up to 4-body interactions. Indeed, in the main text, we mainly focused on analyzing 1D cycles (i.e. the generators of the homology group H_1) of the simplicial filtration $\mathbb{F}(\mathcal{K}^t)$ that we generate at each time t , while we explicitly did not consider 3D-cavities. This is because we limited our analysis up to 3-body interactions (i.e. up to triangles), and 3D-cavities do not add meaningful information. As a matter of fact, if only 3-body interactions are present, then all the 3D-cavities have an infinite persistence due to the absence of 4-body interactions that close cavities. In other words, without 4-body interactions, 3D-cavities will have a birth in the filtration but not a death. Nevertheless, the presence of cavities in the simplicial filtration $\mathbb{F}(\mathcal{K}^t)$ might add

important information when characterizing the higher-order structure of a multivariate time series, as shown when analysing the human connectome data [4].

Figure S4 reports the persistence diagrams for the homology groups H_1 , H_2 obtained in the context of Coupled Map Lattices with $N = 50$ nodes when extending our framework to 4-body interactions. Remarkably, we find that the persistence diagram for H_2 is diverse for the five dynamical regimes and, as a result, it might be used together with H_1 to further differentiate the CML dynamical regimes.

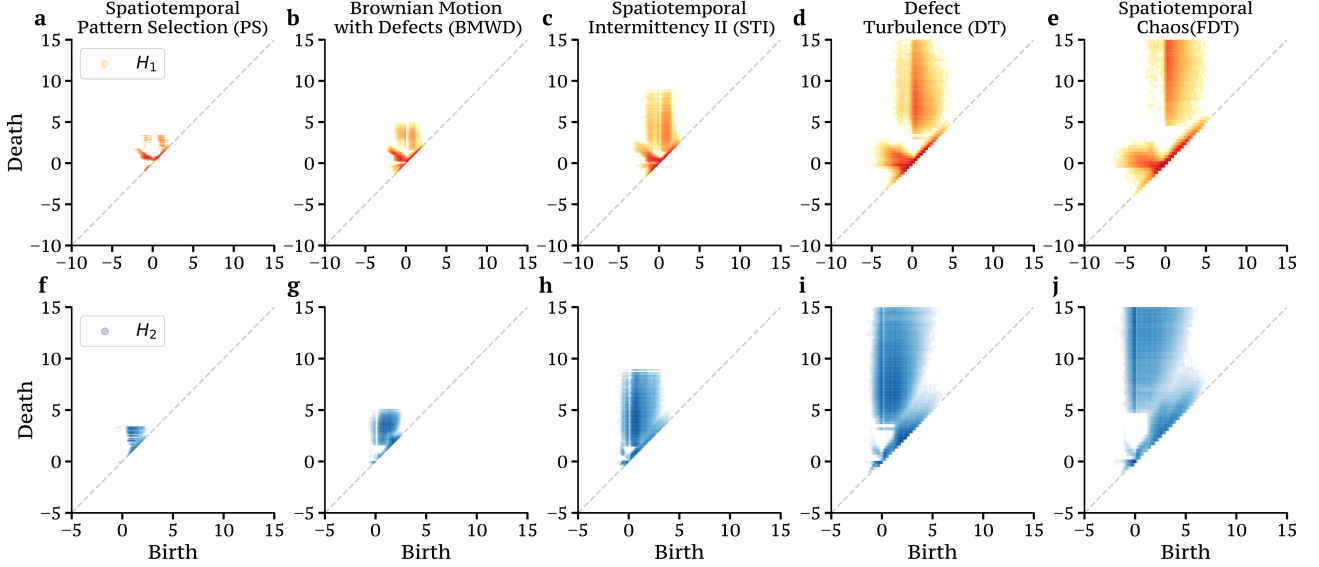


Figure S4. **2D histograms for the H_1 and H_2 generators in Coupled Map Lattices.** We report, for each dynamical state of the Coupled Map Lattice with $N = 50$, the persistence diagram for the homology groups H_1 and H_2 . Interestingly, the H_2 persistence diagram is different between the CML dynamical regimes and, as a result, it might help to further differentiate the five dynamical states when considered in synergy with H_1 . Results are obtained considering 10 independent realization and 240 time points for each block in analogy with the results reported in the main text. The dynamical states are: Pattern Selection (PS) at $\varepsilon = 0.12$, Brownian motion with Defects (BMWD) at $\varepsilon = 0.08$, Spatiotemporal Intermittency II (STI) at $\varepsilon = 0.3$, Defect Turbulence (DT) at $\varepsilon = 0.068$, and Fully Developed Turbulence (FDT) at $\varepsilon = 0.05$.

S2.3. Higher-order approach on Gaussian multivariate time series

In this section we show the Hyper Coherence distribution obtained when considering a N -dimensional Gaussian multivariate time series, where each of the time series is sampled from a normal distribution $\mathcal{N}(0, 1)$ with zero mean and unit variance.

We report in Figure S5 the violin plots reporting the hyper coherence distribution computed from resting-state fMRI data, the corresponding null model obtained when independently reshuffling the time series, and from Gaussian multivariate time series. It is interesting to notice how the null model of the resting-state fMRI leads to a hyper coherence distribution that is statistically equivalent to the one obtained when analysing Gaussian multivariate time series (p -value $< 10^{-10}$ with the Kolmogorov-Smirnov test). This is due to the fact that, even though each BOLD signal is not Gaussian per se, the entire ensemble can be considered as Gaussian. As a result, the simple shuffling procedure destroys all the temporal dependencies and the resulting co-fluctuations of the multivariate signals tend to a Gaussian distribution.

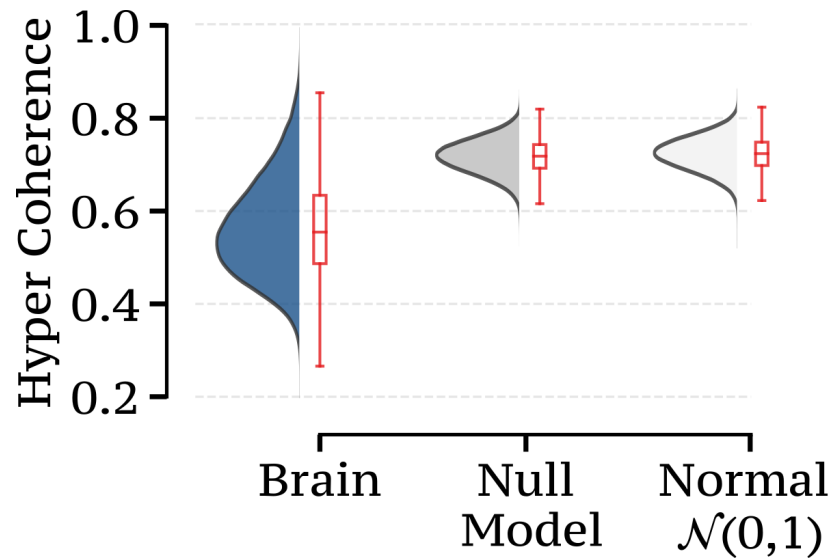


Figure S5. **Hyper coherence distribution computed from resting-state fMRI data and Gaussian multivariate time series.** Violin plots showing the distribution of Hyper Coherence for the resting-state fMRI data (N=119 brain regions), the corresponding null model obtained when independently reshuffling the time series, and for a multivariate time series, where each of the time series is sampled from a normal distribution $\mathcal{N}(0, 1)$. The null model of the resting-state fMRI data leads to a distribution of Hyper Coherence that is statistically equivalent to the one obtained when analysing Gaussian multivariate time series (p -value $< 10^{-10}$ with the Kolmogorov-Smirnov test).

S3. COMPARISON OF TIME SERIES APPROACHES TO DISTINGUISH THE CML REGIMES

In this section, we provide a detailed investigation of the ability of different indicators to differentiate the five regimes generated by $N = 119$ diffusively coupled fully chaotic maps, extensively considered in the main text. The multivariate time series consists of $N = 119$ nodes and $T = 1200$, obtained by concatenating five different CML regimes with fixed time length $L = 240$.

We first compare our higher-order measures with the lower-order dynamical indicator originally proposed in Ref. [5]. This is the Root Sum Square (RSS) of the edge-time series, which was recently used to identify important “events” in fMRI signals [5, 6], and it is a direct proxy of the amplitude of the collective co-fluctuations of the edge time series. In other words, we compute the amplitude of the edge time series as the root sum of squared co-fluctuations, i.e. $RSS(t) = \sqrt{\sum_{i,j>i} e_{ij}(t)^2}$. Here, the vector $\mathbf{e}_{ij} = \mathbf{z}_i \mathbf{z}_j$ is the edge time series obtained as a product of the z -scores of the original time series. We then assess the performance of several other “static” approaches, namely, (i) the information-theoretic approaches introduced in Refs. [7, 8] accounting for higher-order interactions, (ii) the dyadic method at the interface of network science and random matrix theory [9], and (iii) the classical approach based on Pearson’s correlation coefficient [10].

Top panels of Figure S6 depict the violin plots of the distribution of RSS and of the three higher-order measures for the five regimes of the CMLs. Surprisingly, the RSS distributions of certain dynamical states (BMWD, DT, and FDT) are highly similar to each other, somehow mirroring the inability to capture the subtleties of these regimes. By contrast, the higher-order measures seem to differentiate the dynamical states in a qualitatively better way. As a matter of fact, the ability to distinguish the different dynamical regimes is quantitatively confirmed by the ICC values associated with each metric and reported in the bottom panels. In particular, we have RSS with an $ICC \approx 0.57$, hyper coherence with $ICC \approx 0.95$, hyper complexity with $ICC \approx 0.96$, and average edge violation with $ICC \approx 0.91$. For the sake of completeness, for each dynamical measure we also report an ICC matrix encoding the ICC values

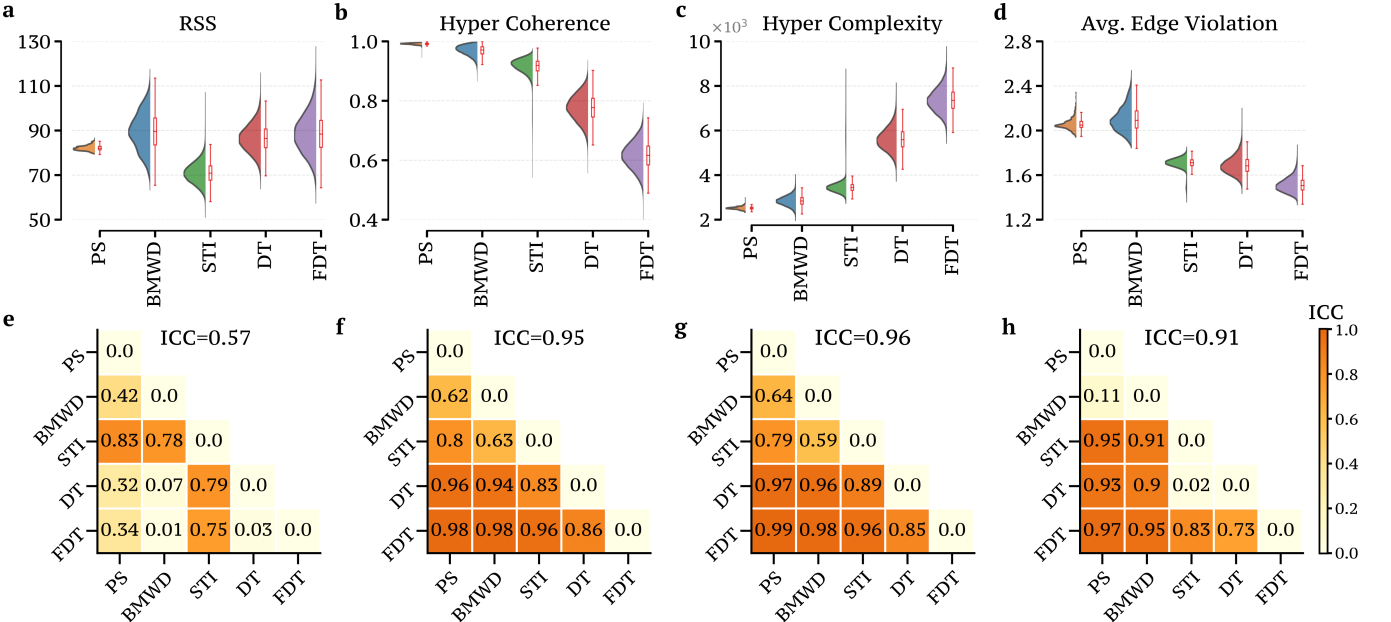


Figure S6. **Comparison of different dynamical metrics.** (a-d) Violin plots showing the distribution of RSS (Root Sum Square) and three higher-order measures for the five different dynamical regimes generated by the diffusively coupled chaotic maps. Remarkably, the RSS statistic, which only captures the effect of lower-order structures (i.e. the edges), is not able to well-separate the five dynamical regimes. By contrast, the higher-order metrics introduced in this work distinguish the different regimes. (e-h) The performance of each indicator is quantitatively measured by the ICC values computed either considering the five distributions or between all the possible pairs. When considering all the five distributions, we find RSS with an $ICC \approx 0.57$, hyper coherence with $ICC \approx 0.95$, hyper complexity with $ICC \approx 0.96$, and average edge violation with $ICC \approx 0.91$. Results are averaged over 100 independent realizations.

between all the possible pairwise comparison of dynamical regimes.

Lastly, for each of the five CML regimes, we independently compute diverse “static” higher-order and pairwise measures [7–10], aggregating the results across 100 independent realizations. In other words, after isolating each dynamical state by selecting the precise intervals of size $T = 240$, we compute the selected measure on that interval. This is necessary, since such measures are not dynamic (i.e. they are not defined on a single frame), but instead they require a time-window to be computed.

Figure S7 reports the results of our analyses for the different approaches. Remarkably, several of the information-theoretic measures introduced by Rosas and colleagues [7] [see Figs. S7(a-d)] are able to differentiate in a quantitative

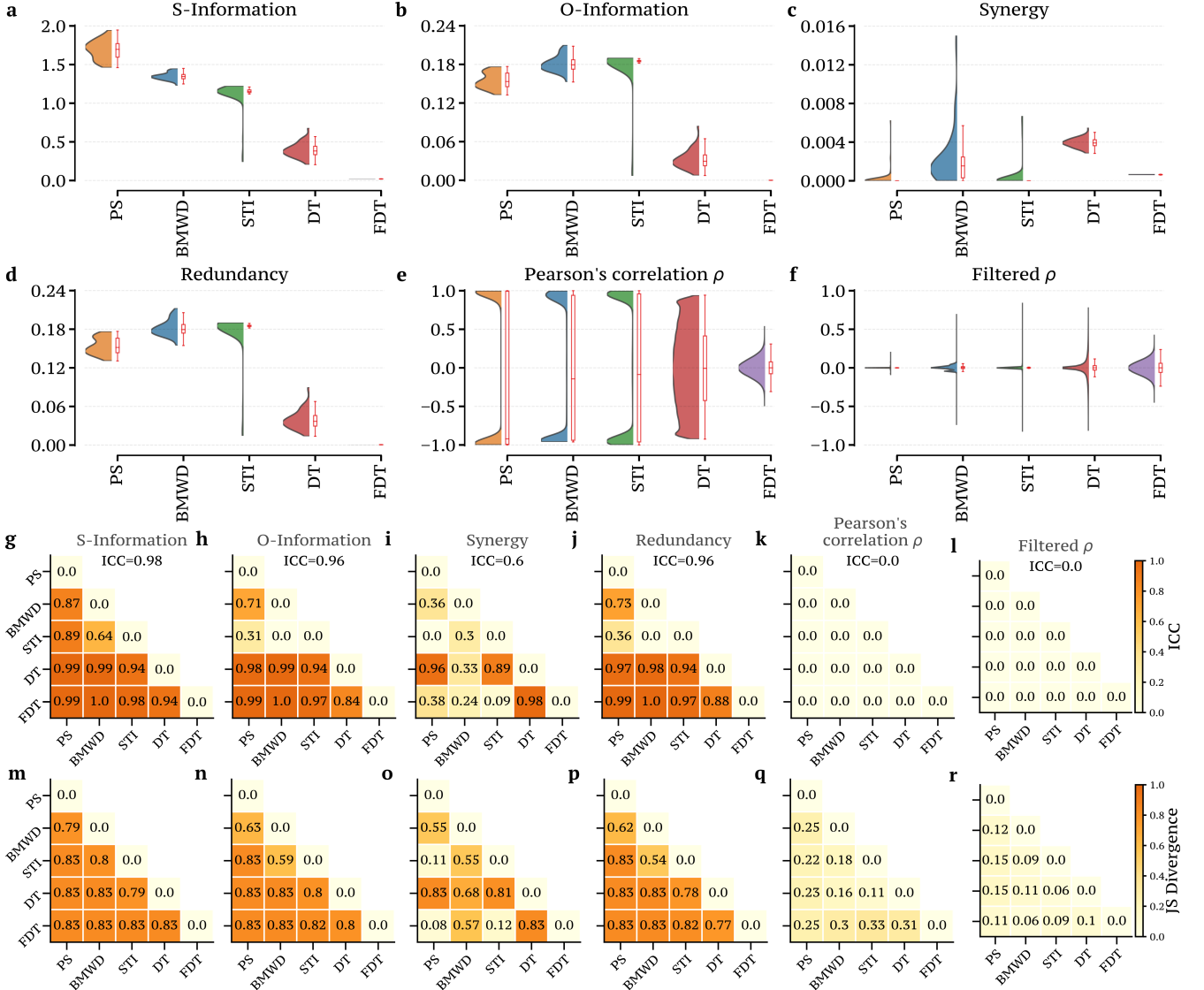


Figure S7. **Comparison between approaches for distinguishing the CML regimes.** (a-f) For the five dynamical regimes generated by the diffusively coupled chaotic maps, we report the violin plots showing the distribution of “static” higher-order measures [7], and “static” pairwise measures, namely, the Pearson’s correlation coefficient ρ and filtered ρ obtained by considering the approach based on random matrix theory [9]. Remarkably, only higher-order measures are able to separate the five dynamical regimes. (g-l) We report the *ICC* values associated with each metric, either computed considering the five distributions or between all the possible pairs, thus demonstrating that higher-order measures outperform statistics based on pairwise correlations. (m-r) We repeat the same analysis but considering the Jensen-Shannon (JS) divergence to measure distances between distributions, leading to the same conclusions of the ICC approach.

way the different CML dynamical states, as confirmed by the high ICC values reported in panels S7(g-j)¹. By contrast, approaches only based on pairwise statistical dependencies [see Figs. S7(e-f)], such as “connectome” analyses based on Pearson’s correlation ρ or the filtered ρ relying on random matrix theory [9], fail at differentiating the five dynamical regimes [see Figs. S7(k-l)]. To confirm the quantitative analyses provided by the ICC metric, we repeat an analogous analysis but considering the Jensen-Shannon (JS) divergence to measure distances between distributions [see Figs. S7(m-r)].

¹ The code for these analyses has been adapted from Ref. [8].

S4. IMPACT OF DIFFERENT NULL MODELS

S4.1. CML synthetic multivariate time series

After having demonstrated that standard metrics to assess two-body similarity of time series, such as RSS, Pearson correlation, or more refined approaches based on random matrix theory [9], fail to distinguish the CML dynamical regimes, we now shift the focus to investigating the impact of alternative null models [11–13]. Indeed, the null reported in Figs. 2-3 of the main text — obtained by independently shuffling the time series — might be in principle too simplistic to detect regime-specific features of the co-fluctuations in the CML case. This is because the empirical signal of the co-fluctuations in each dynamical regime is compared to products of node-level signals at different times and for different regimes. As a consequence, we test the global behaviour of our two higher-order indicators when considering two alternative null models, namely, the *(i)* block null model, and the *(ii)* phase randomization null model. The first null model, hereafter denoted as *block null model*, is constructed such that the shuffling of each dynamical regime of the CML is performed separately for each region (i.e. the blocks consisting of 240 time points). In this way, we are ensuring that the dynamical regimes are not mixed during the reshuffling, and we can therefore detect regime-specific features of the co-fluctuations.

The second null model, hereafter denoted as *phase randomization null model*, is a conventional approach [11] to generate surrogate time series preserving the empirical power spectra, while randomizing the temporal dependencies. More specifically, such time series are generated by transforming the CML time series to the frequency domain via Fourier transform, shuffling the phase coefficients, and then taking the inverse transform to the time domain [13].

We first aim at replicating the panels of Fig. 2(a,c) of the main text when considering the two alternative null models. In particular, Figure S8 summarizes the results of our higher-order approach when applied to the two surrogate multivariate series. We find that the distributions of hyper coherence in the block null model become less distinguishable from each other, while keeping certain regime-specific features of the co-fluctuations. Furthermore, the block null model does not preserve any more the ranking between order (i.e. the PS regime) and disorder (i.e. the Spatiotemporal Chaos), see Fig. S8(a). In addition, the distributions of hyper complexity in the block null model are mostly identical to each other for the five blocks, so that it is clear that there are some intrinsic topological properties that are only present in the CML multivariate time series [see Fig. S8(b)]. When considering the phase randomization null model, we find a flat behaviour for both the hyper coherence and hyper complexity distributions, see Fig. S8(c-d).

As done in previous sections, to quantitatively test the ability to distinguish the different CML dynamical states in the null models, we rely again on the ICC values, which are either computed considering the five distributions or between all the possible pairs of regimes. We report in Figure S9 the distributions of hyper coherence and hyper complexity for the actual CML time series and for all the different surrogates time series.

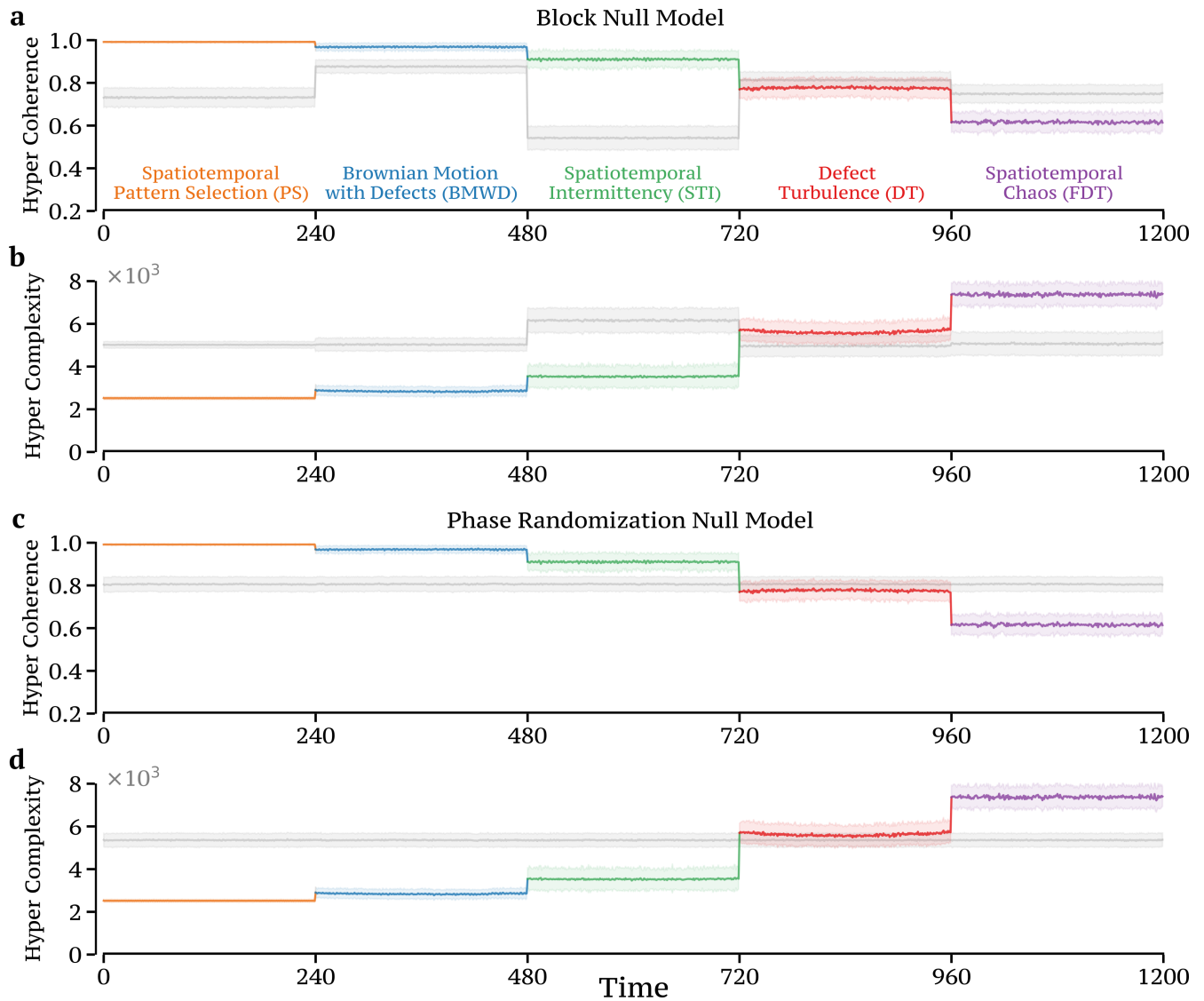


Figure S8. **Global behaviour of our higher-order indicators for two alternative null models.** (a-b) We report the behaviour of the two higher-order indicators, as in Fig. 2 of the main text, respectively computed for the original CML and for the surrogate time series generated by the block null model, which shuffles independently each block of 240 time points. Remarkably, the resulting blocks of hyper coherence in this null model become less distinguishable from each other, while keeping certain regime-specific features of the co-fluctuations. However, the ranking from ordered to disordered states is now lost. In addition, the blocks of hyper complexity cannot be separated in the block null model, somehow supporting the intuition that this topological-based measure is more fine-grained than the hyper coherence. (c-d) The same analyses are repeated when considering the phase randomization null model, which preserves the power spectra of the CML time series.

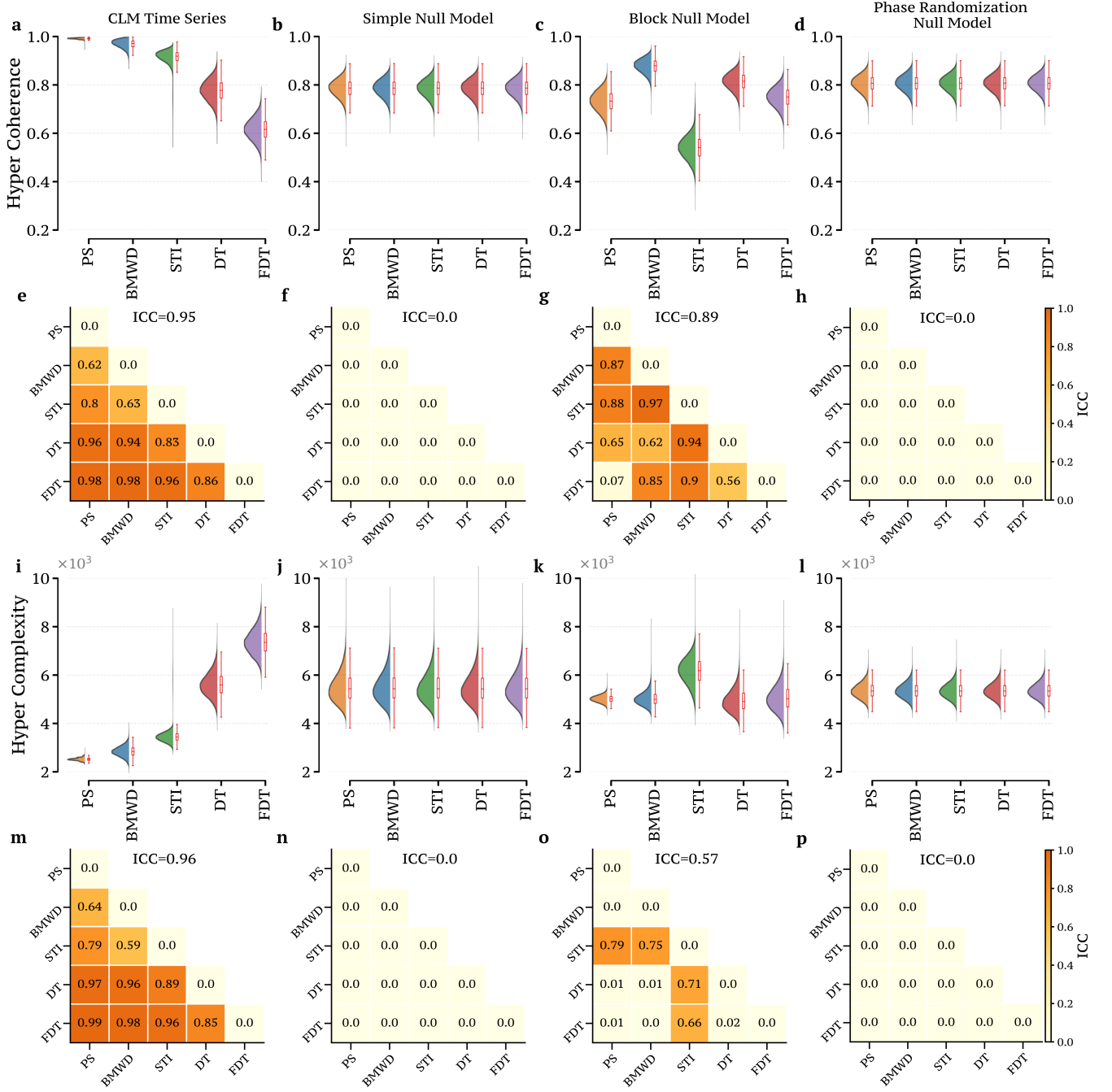


Figure S9. **Comparison of null models in Coupled Map Lattice.** (a-d) We report the distribution of hyper coherence for the classical CML time series and for the three null models considered in this work, namely, the simple null model considered in the main text, the block null model, and the phase randomization null model. (e-h) Both the global and pairwise analysis of the ICC metric reveal that only the block null model is able to partly separate the CML regimes, hence keeping certain regime-specific features of the co-fluctuations, while breaking the order-disorder gradient present in the real Hyper Coherence distributions. (i-l) When repeating the analyses of panels (a-d) for the distribution of hyper complexity, we find that the distributions of hyper complexity in the null models are mostly identical to each other for the five blocks, so that it is clear that there are some intrinsic topological properties that are only present in the CML multivariate time series. (m-p) Global and pairwise analysis of the ICC metric provide a quantitative way of assessing the similarities between each regimes in the different null models.

S4.2. Block null model in financial time series

Here, we assess whether the results of hyper coherence of the financial time series presented in the main text can be trivially obtained by analysing block null models which account for different global trends, including upwards and downwards ones. As a matter of fact, in the considered time span of 21 years (2000-2021) the market experienced different global trends. Therefore, when considering a simple null model, as the one described in the main text, which independently reshuffles each stock price in the whole window, we might neglect the importance of different financial trends. In order to assess this potential confound, we consider three different block null models with respectively 3, 6, and 9 temporal blocks, and test whether the findings on financial time series in the main text could not be simply deduced from more sophisticated null models. The blocks are selected in order to capture the major financial trends of the last 21 years and to consider crises as midpoints of equally spaced time windows.

In Figure S10 we report the distributions of hyper coherence for the financial dataset along with the four null models obtained by independently reshuffling the real-world multivariate time series over the whole time window of 21 years, or restricting the shuffling over 3, 6, or 9 temporal blocks, respectively. It is interesting to notice that null models with finer information about the global financial trends seem to increasingly approximate the distribution of hyper coherence of the empirical dataset, despite some notable differences are present even in the finer model consisting of 9 temporal blocks. For example, the distribution of the 9 blocks null model is almost 4-modal, while it is bi-modal in the empirical case, and the two distribution are statistically different, i.e. $p < 10^{-10}$ after the Kolmogorov-Smirnov test. Finally, for the sake of completeness, we report the list of periods considered for the block null models in Figure S10(b).

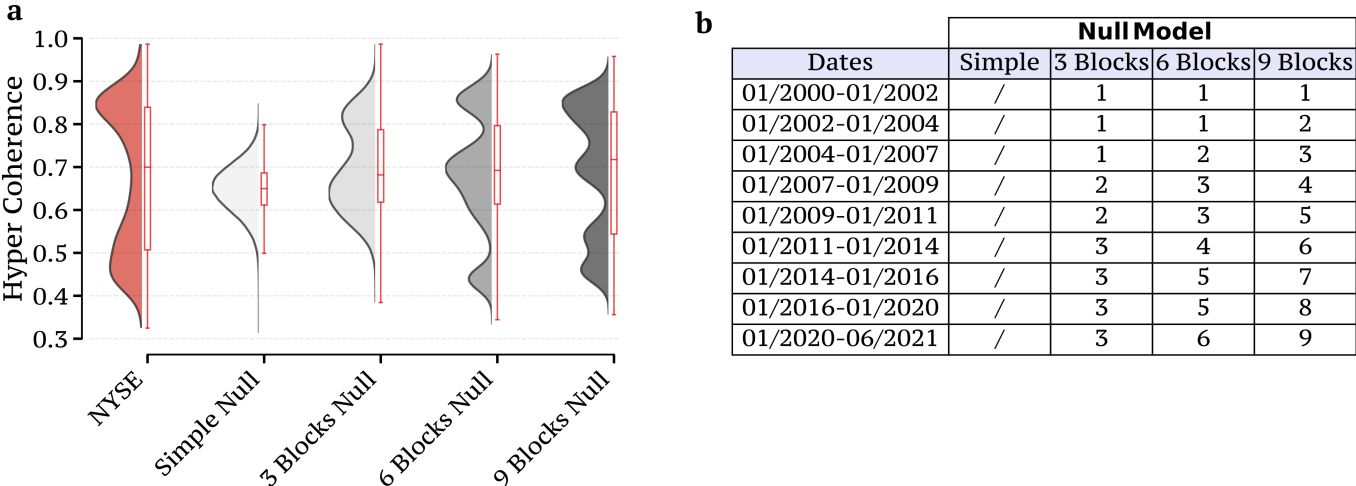


Figure S10. **Distribution of hyper coherence for different block null models.** (a) Violin plots showing the distribution of hyper coherence for the financial prices of 119 assets in NYSE spanning a period from 01/2000 to 06/2021. We also report the corresponding null models obtained when independently reshuffling the empirical time series over the whole time window of 21 years (i.e. Simple Null), or restricting the shuffling over 3, 6, or 9 blocks, respectively. Even when considering the null model obtained with 9 blocks (dark grey curve), the hyper coherence distribution is distinct from the one obtained for the empirical multivariate time series (red curve). (b) We report the periods considered for the various null models, which include finer information about the global financial trends. For instance, the 3 block null model spans three different periods, namely, 01/2000 – 01/2007, 01/2007 – 01/2011, and 01/2011 – 06/2021.

S5. EDGE-BASED INDICATORS IN REAL-WORLD SYSTEMS

In this section we investigate the impact of edge-based statistics in the context of real-world multivariate time series when compared to the higher-order approach presented in the main text. More specifically, we first analyse the RSS distributions for the three types of real-world systems considered in this work. Subsequently, we aim at reproducing some of the results of Fig. 5 of the main paper by relying on metrics derived from RSS or, more in general, from measures based on edge signals. Our final goal is to compare several different group-order representations in the context of the real-world systems considered in this study.

As global analysis, we report in Figure S11 the distributions of RSS for the three real-world datasets. For the sake of comparison, we also plot the null models obtained by independently reshuffling the real-world multivariate time series. We first notice that, in contrast with the results reported in the main text, these distributions are not always statistically distinct from the corresponding null models, as evident when looking at the distributions of the epidemic data (e.g. $p < 0.001$ for Gonorrhea with the Kolmogorov-Smirnov test). Moreover, when examining the RSS distribution of resting-state brain and of financial systems, it is not possible to extract insights on the nature of different temporal dynamics present in the system (e.g. we cannot discriminate crises from periods of financial stability).

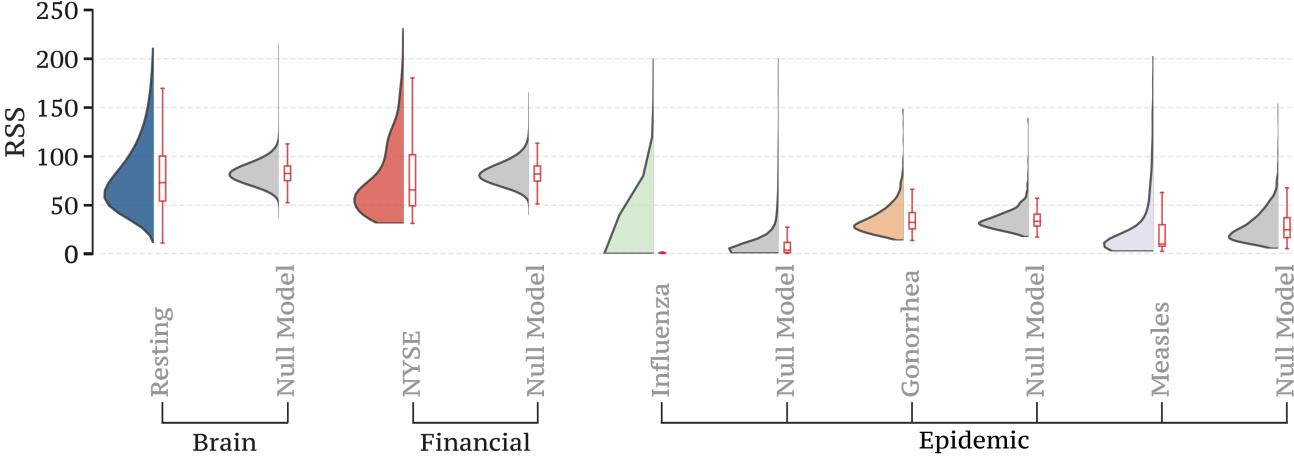


Figure S11. **RSS distribution for the empirical time series.** Violin plots showing the distribution of RSS for the three real-world datasets considered in the main text, namely, resting-state fMRI data (N=119 brain regions), financial prices of 119 assets in NYSE in the period 2000-2021, and the US historical data of several infectious diseases at the US state-level (N=50). While for the brain and financial systems the empirical RSS distributions are different from the corresponding null models (obtained by independently reshuffling each time series), the RSS distributions for the epidemic time series are very similar to the ones obtained when considering the corresponding null models ($p < 0.001$ for Gonorrhea, with the Kolmogorov-Smirnov test).

Once we established that global edge-wise measures such as RSS provide moderate information on the empirical systems, we now shift the focus on investigating the impact of different group-order representations on a more local level. In particular, we performed several analyses on the empirical time series by projecting the magnitudes of higher- and lower-order approaches on nodal level and compare the overall results. Yet, we remind that any lower-order projection might lead to a moderate/high reduction of the total amount of information and therefore such representations might be potentially misleading if not carefully analysed.

Figure S12 summarizes the results of our comparisons for the brain data. Here, we mapped the node strength extracted from the violating triangles Δv , the homological scaffold, and the edges time series onto the cortical surface and found various patterns of co-fluctuations. In particular, we first observe that brain patterns uncovered by the 15% top hyper coherent frames provide quite similar brain configurations as the ones constructed from both top- and bottom-RSS edge-wise representations. By contrast, all the other representations based on higher-order measures provide different brain activity patterns, with the nodal strength of the homological scaffold from the 15% bottom hyper complexity frames mostly encompassing the Default Mode Network (DMN), see Fig. S12(d).

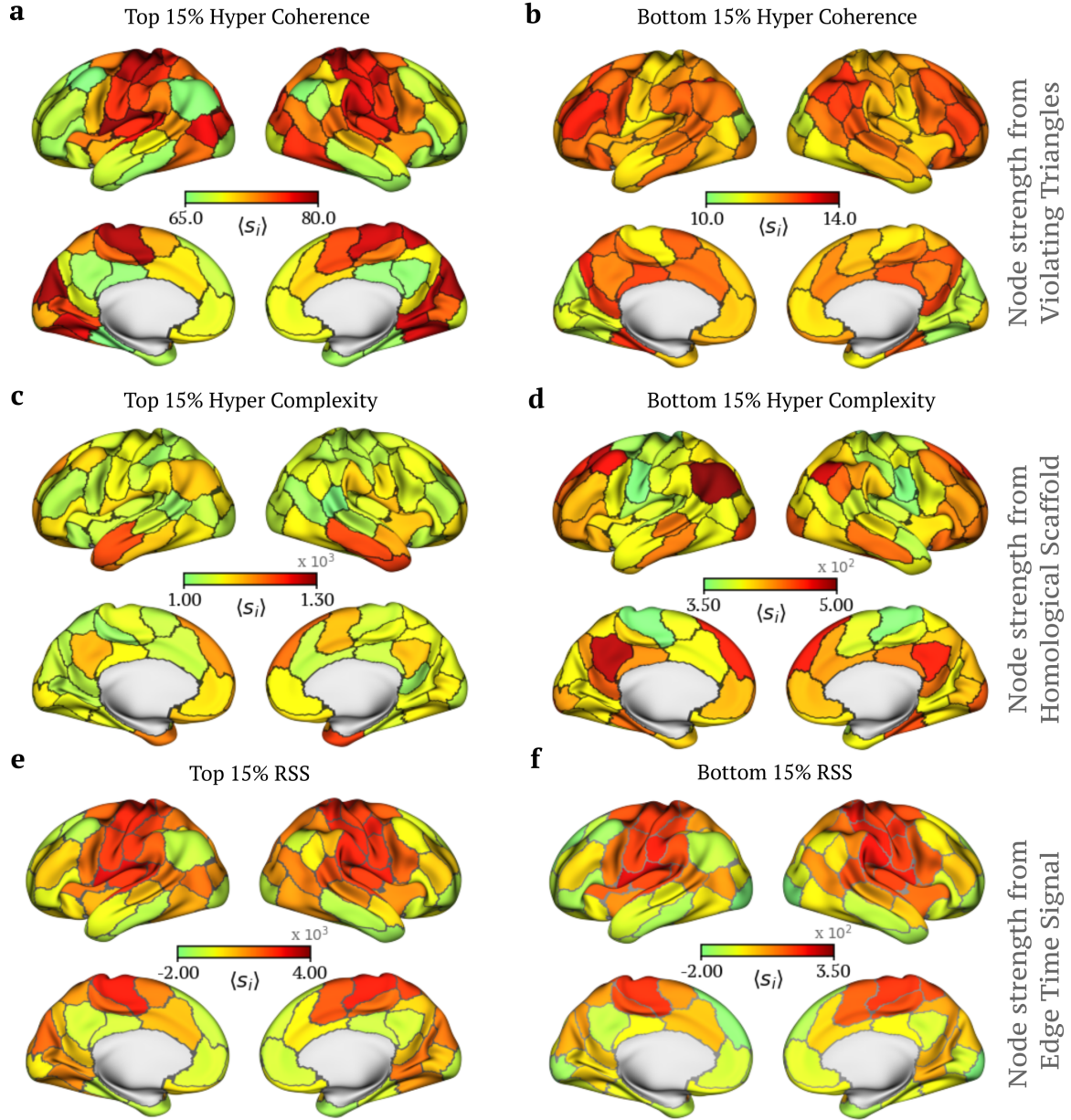


Figure S12. **Impact of different group-order representations in brain patterns.** Node strength extracted from the violating triangles Δv , homological scaffold, and edge time series can be used to track the importance of group-order structures. **(a-b)** We report the brain map of the nodes involved in higher-order co-fluctuations obtained when isolating the 15% high- and low- coherent frames, respectively. **(c-d)** We report the brain map of the nodes involved in homological scaffold obtained when isolating the 15% high- and low- complexity frames. To compare different group-order representations, we also performed a similar analysis as panels (a-d), yet considering the nodal strength extracted from the edge-wise representation. In particular, **(e-f)** we map onto the cortical surface the nodal strength extracted directly from the 1-order co-fluctuations (i.e. the edge time series), computed from the 15% high- and low-amplitude frames, which have been selected using the RSS statistics. Despite losing part of the information by projecting each group-order representation on a nodal level, we observe that the brain patterns uncovered by top- and low- peaks of the higher-order approach provide different brain configurations compared to edge-wise representation, with the only exception being the top hyper coherence frames (see Fig. S13 for all the correlations of activity patterns). Results are averaged over all 100 HCP subjects and scans.

Yet, the analyses presented so far provide only a qualitative comparison between the approaches. We therefore quantitatively investigate the spatial patterns identified on both resting-state brain data and financial systems by the various methods and compare them considering nodal projections. For clarity, we first provide in Table S1 a summary description of each method, while we report in Figure S13

Method	Description
$\uparrow \Delta_v$ ($\downarrow \Delta_v$)	The top (bottom) 15% of hyper-coherent frames are selected using the Hyper Coherence indicator. For only those frames, we compute the average spatial distributions by projecting the list of violating triangles Δ_v on a nodal level.
$\uparrow HC$ ($\downarrow HC$)	The top (bottom) 15% of the hyper-complexity frames are selected using the Hyper Complexity indicator. For only those frames, we compute the average nodal strength from the homological scaffold, which is obtained from the persistent homology generators of H_1 [14].
$\uparrow eFC$ ($\downarrow eFC$)	The top (bottom) 15% of the amplitude frames are selected using the RSS measure. From those frames, we then compute the edge functional connectivity matrix [15] and report the nodal strength obtained from such a matrix.
$\uparrow eTS$ ($\downarrow eTS$)	The top (bottom) 15% of the amplitude frames are selected using the RSS measure. For those frames, we then compute the average activity of the 1-order co-fluctuation (i.e. the edge time series) and report the nodal strength.
$\uparrow nodal$ ($\downarrow nodal$)	The top (bottom) 15% of the frames are selected using a RSS-like statistic but computed on the raw signal. For those frames, we then compute the average activity pattern to obtain nodal maps.

Table S1. **Summary of the different approaches.** We provide a short description of the methods considered in this study to construct spatial maps from multivariate time series. We remind also that Δ_v and HC are higher-order measures, eFC and eTS are edge-wise statistics, while $nodal$ corresponds to spatial measures based on the original nodal signal.

the matrices encoding all the possible Pearson’s correlations between the various spatial distributions for both the brain data and financial data. In particular, by analysing Figure S13(a), we confirm in a quantitative way that the spatial patterns uncovered by the 15% top-peaks of the higher-order approach using the hyper coherence (i.e. $\uparrow \Delta_v$) provide brain patterns which are very similar to the ones constructed from both top- and bottom-RSS edge-wise representations. Therefore, at least in brain resting data, such representation might lead to similar results as the one obtained using edges. By contrast, the other higher-order approaches provide patterns that are not entirely deducible from an edge-wise approach and, as such, they might provide new insights into brain dynamics. However, when analysing the same correlation matrix for financial systems [Figure S13(b)], we surprisingly find a rather different picture. In this case, the spatial pattern of the top 15% hyper coherent frames (i.e. $\uparrow \Delta_v$) is very dissimilar from all those obtained using edge approaches (i.e. $\rho \approx 0$), while the other higher-order approaches still provide patterns that are not captured by any edge-wise approach. This confirms our intuition that various real-world systems might be affected, in varied ways, by group-interactions at different instants in time.

Finally, by analysing the historical data of epidemic outbreaks in the US, we show that the temporal evolution of the RSS measure cannot be effectively used to classify different infectious diseases. Indeed, with a Random Forest classifier we obtain an accuracy level of around 67% using a 10-fold cross-validation setting repeated 50 times. See Section S7 S7.3 for a comparison with classification based on higher-order information.

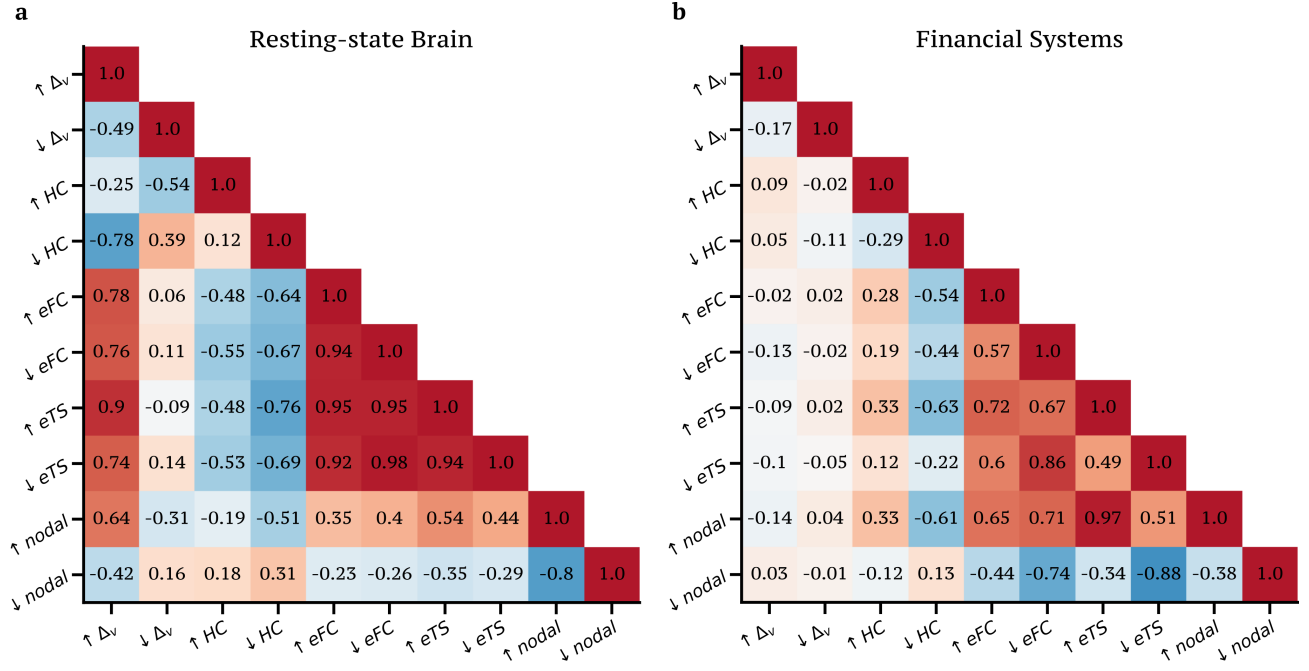


Figure S13. **Correlations between group-order representations of brain and financial patterns.** We report the Pearson's correlation coefficient ρ obtained when comparing the nodal projections for (a) brain and (b) financial systems. Here we have that Δ_v and HC are higher-order measures, eFC and eTS are edge-wise statistics, while $nodal$ corresponds to measures based on the nodal signal. \uparrow (\downarrow) correspond to the 15% top- (respectively bottom-) peaks of the selected measures. For the sake of clarity, in Table S1 we reported a short description of the methods. Interestingly, while the spatial patterns uncovered by the 15% top-peaks of the higher-order approach using the hyper coherence (i.e. $\uparrow \Delta_v$) provide brain patterns very similar to the ones constructed from both top- and bottom-RSS edge-wise representations, all the other three higher-order approaches provide patterns that are not entirely deducible from an edge-wise approach. However, in financial systems we have an opposite view, so that the patterns obtained from the top 15% hyper coherent frames (i.e. $\uparrow \Delta_v$) is now very dissimilar from all the ones obtained using edge approaches (i.e. $\rho \approx 0$).

S6. EFFECT OF NOISY FLUCTUATIONS IN BRAIN DATA

As mentioned in the main text, one of the limitations of our higher-order approach working on single-frame timescales is that it might be affected by noisy fluctuations in the time series, as many of other existing methods [5, 16, 17]. However, in practice, this problem can be smoothed out by analyzing the statistics of multiple time frames. In this section we confirm that, even in presence of noisy fluctuations in brain data (i.e., head movement), the resulting spatial distribution of the hyper coherence appear to be robust and stable. More precisely, similarly to Ref. [18], we considered frames (i.e. fMRI volumes) affected by head motion if the relative Root Mean Square (RMS) between frames was greater than 0.2, i.e. $RMS > 0.2$.

Figure S14 reports the nodal spatial distributions obtained from the violating triangles Δ_v using the top- and bottom 15% hyper coherent frames when analysing the fMRI resting-state data in presence and absence of fMRI volumes containing high head motion. Notably, we obtain a Pearson's correlation coefficient $\rho \approx 0.999$ when comparing the resulting brain maps in presence and absence of fMRI volumes affected by head motion.

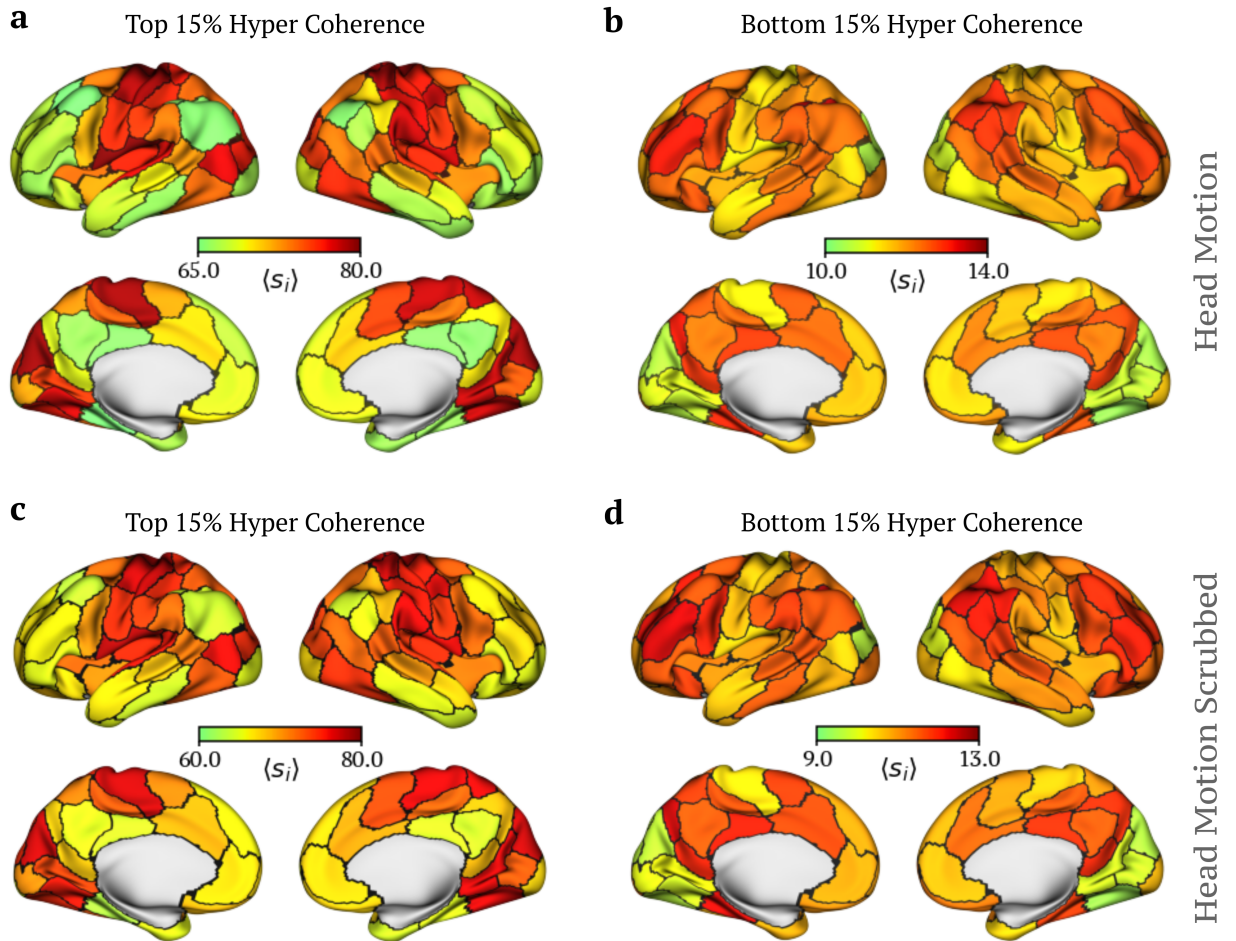


Figure S14. **Effect of head motion in brain maps.** We report the nodal spatial distributions obtained from the violating triangles Δ_v using the top- and bottom 15% hyper coherent frames when analysing the fMRI resting-state data in (a-b) presence and (c-d) absence of fMRI volumes containing noisy fluctuations (head motion). When comparing the resulting brain maps we obtain a Pearson's correlation coefficient of $\rho \approx 0.999$.

S7. LOCAL HIGHER-ORDER INDICATORS IN REAL-WORLD SYSTEMS

S7.1. Role of thresholds in brain maps

In the main text, we have shown that the nodal strength obtained when projecting either the magnitudes of the violating triangles Δ_v or the homological scaffold on a nodal level might provide interesting information about the group co-fluctuations of certain brain regions. In particular, since our goal was to characterize the higher-order states with the largest level of synchronization, we isolated only the 15% high-coherent (resp. low-complex) frames. Yet, in principle, a different percentage of selected peaks might translate into a distinct brain activation maps. As a consequence, in Fig. S15 we report the same brain maps as shown in the main text, yet we vary the percentage of peaks selected.

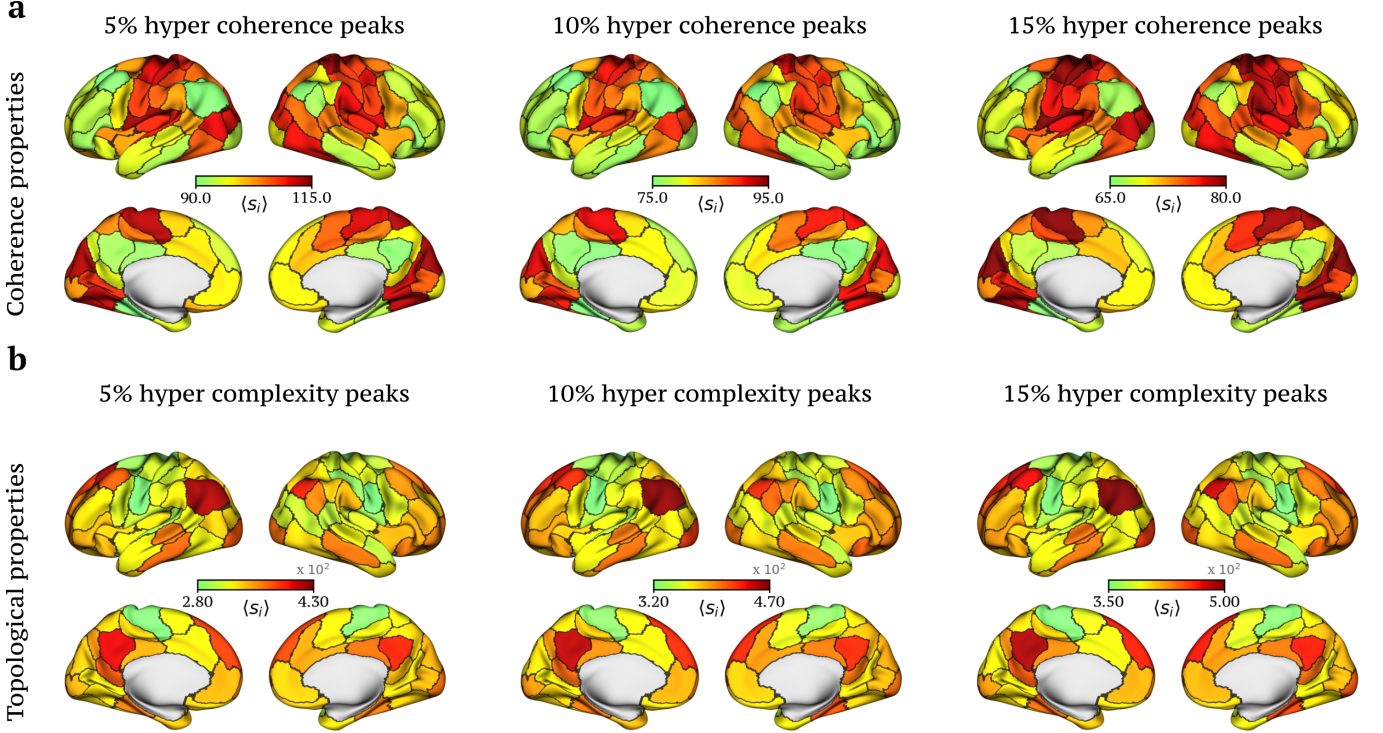


Figure S15. **Brain maps with different percentage of peaks.** The nodal strength extracted from either the violating triangles Δ_v or the homological scaffold can be used to track the importance of higher-order structures in time. (a) Brain maps of the nodes involved in higher-order co-fluctuations obtained when respectively isolating the 5, 10, 15% high-coherent frames, which are those associated with a more synchronized dynamical phase. Interestingly, brain maps are highly consistent across different values of thresholds. From a topological perspective, the nodal strength extracted from the homological scaffold provides information about 1D loops in the space of co-fluctuations. (b) Brain maps obtained when respectively selecting the 5, 10, 15% low-hyper complexity frames reveal the consistent activation of the Default Mode Network. Indeed, also in this case, the brain maps are mostly the same for different threshold values.

S7.2. Higher-order indicators in the brain's functional networks

In this section, we consider the local higher-order indicators introduced in the main text to analyze the brain at the level of the 7 functional Yeo networks [19] and subcortical regions. In particular, we project the normalized magnitude of the violating triangles Δ_v extracted from the 15% top hyper coherent frames and of the homological scaffold extracted from the 15% low hyper complexity frames at the nodal and edge level. This provides us with information about the most prominent functional networks when isolating a certain percentage of frames both at the level of nodes and edges. Figure S16(a) reports the mean hyper coherence within- and between- the 7 functional Networks plus subcortical regions, namely, the Visual (VIS), SomatoMotor (SM), Dorsal-Attention (DA), Ventral-

Attention (VA), Limbic (L), FrontoParietal (FP), Default Mode Network (DMN), and subcortical regions (SUBC). While the top boxplots (also shown in the main text) report the mean hyper coherence within the 7 functional Networks, the entries of the matrix encode the mean hyper coherence within- and between- the 7 functional networks and subcortical regions obtained when projecting Δ_v at the level of their interacting edges. Notably, hyper coherent interactions (i.e. the triangles) are mainly overrepresented in sensorimotor areas. By contrast, when analysing the mean persistence of 1D holes considering the homological scaffold, we find that such loops are highly concentrated around the DMN and FP areas, as shown in Fig. S16(b), mirroring the importance of such areas in integrating high and low-order systems.

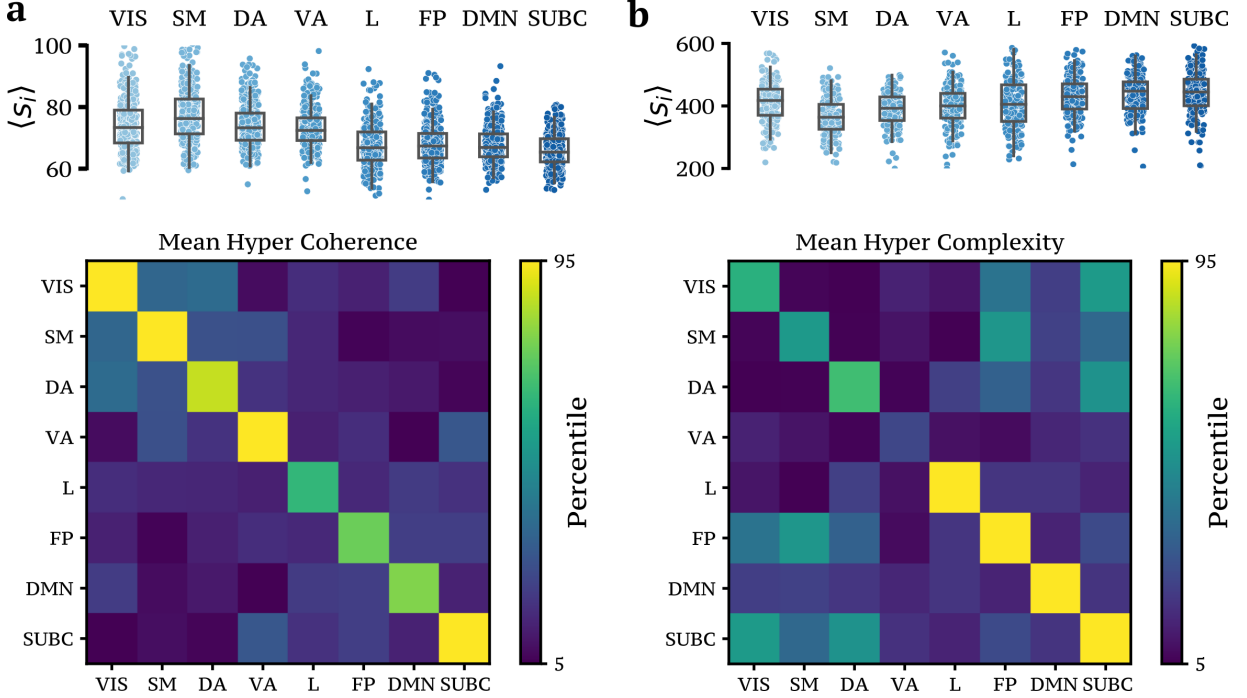


Figure S16. **Mean within- and between higher-order indicators in the seven functional networks.** For the two local higher-order indicators considered in the main text, namely, (a) the list of violating triangles Δ_v and (b) the homological scaffold, we report the mean nodal strength and average edge strength within- and between- the 7 functional Networks and subcortical regions, namely, the Visual (VIS), SomatoMotor (SM), Dorsal-Attention (DA), Ventral-Attention (VA), Limbic (L), FrontoParietal (FP), Default Mode Network (DMN), and subcortical regions (SUBC). When analysing the list of violating triangles Δ_v , we find that activity patterns with emphasized synchronized co-fluctuations mainly reflect sensorimotor areas, whereas the persistence of 1D holes obtained from homological scaffold are situated around the DMN and FP areas.

S7.3. Comparison of classifiers for the US historical data

In the main text, we reported the average accuracy score when considering the Support Vector Machine (SVM) method applied to the classification problem of the US historical data of several infectious diseases at the US state-level.

In particular, we have considered the following set of features for the classification task: hyper coherence, the three different contributes of hyper complexity, and the average edge violation. In Table S2, we report the scores of several classifiers obtained when considering a 10-fold cross-validation repeated 50 times with different training-test data partitions. The classifiers considered are: Gaussian naive Bayes (Gaussian NB), SVM using a Gaussian radial basis function as kernel (RBF SVM), Decision Tree, random decision forest (Random Forest), and k-nearest neighbors algorithm (k-NN, with $k=5$). Remarkably, the RBF SVM method, together with Random Forest, leads the pack with the highest accuracy of 0.85. All the analysis reported were obtained using the scikit-learn 1.0.2 python library.

Classifier	Avg. accuracy	F1 weighted score
Gaussian NB	0.47	0.43
RBF SVM	0.85	0.85
Decision Tree	0.81	0.81
Random Forest	0.85	0.85
k-NN	0.83	0.83

Table S2. **Comparison of classifier scores.** We report the average accuracy and F1 weighted scores for the classification of the US historical data of several infectious diseases at the US state-level. We consider hyper coherence, avg. edge violations, and the three different contributes of hyper complexity as features for our classification task. The classifiers considered are: Gaussian naive Bayes (Gaussian NB), SVM using a Gaussian radial basis function as kernel (RBF SVM), Decision Tree, random decision forest (Random Forest), and k-nearest neighbors algorithm (k-NN, with k=5).

-
- [1] Carriere, M., Cuturi, M. & Oudot, S. Sliced wasserstein kernel for persistence diagrams. In *International conference on machine learning*, 664–673 (PMLR, 2017).
- [2] Kaneko, K. Pattern dynamics in spatiotemporal chaos: Pattern selection, diffusion of defect and pattern competition intermittency. *Physica D: Nonlinear Phenomena* **34**, 1–41 (1989).
- [3] Kaneko, K. Overview of coupled map lattices. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **2**, 279–282 (1992).
- [4] Sizemore, A. E. *et al.* Cliques and cavities in the human connectome. *J Comput Neurosci* **44**, 115–145 (2018).
- [5] Esfahlani, F. Z. *et al.* High-amplitude cofluctuations in cortical activity drive functional connectivity. *Proceedings of the National Academy of Sciences* **117**, 28393–28401 (2020).
- [6] Pope, M., Fukushima, M., Betzel, R. F. & Sporns, O. Modular origins of high-amplitude cofluctuations in fine-scale functional connectivity dynamics. *Proceedings of the National Academy of Sciences* **118** (2021).
- [7] Rosas, F. E., Mediano, P. A., Gastpar, M. & Jensen, H. J. Quantifying high-order interdependencies via multivariate extensions of the mutual information. *Physical Review E* **100**, 32305 (2019).
- [8] Gatica, M. *et al.* High-Order Interdependencies in the Aging Brain. *Brain Connectivity* **00**, 1–11 (2021).
- [9] MacMahon, M. & Garlaschelli, D. Community Detection for Correlation Matrices. *Physical Review X* **5**, 021006 (2015).
- [10] Sporns, O. *Networks of the Brain* (MIT press, 2010).
- [11] Schreiber, T. & Schmitz, A. Surrogate time series. *Physica D: Nonlinear Phenomena* **142**, 346–382 (2000).
- [12] Donges, J. F. *et al.* Unified functional network and nonlinear time series analysis for complex systems science: The pyunicorn package. *Chaos: An Interdisciplinary Journal of Nonlinear Science* **25**, 113101 (2015).
- [13] Váša, F. & Mišić, B. Null models in network neuroscience. *Nature Reviews Neuroscience* 1–12 (2022).
- [14] Petri, G. *et al.* Homological scaffolds of brain functional networks. *Journal of the Royal Society Interface* **10**, 186–198 (2014).
- [15] Faskowitz, J., Esfahlani, F. Z., Jo, Y., Sporns, O. & Betzel, R. F. Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture. *Nature neuroscience* **23**, 1644–1654 (2020).
- [16] Tagliazucchi, E., Balenzuela, P., Fraiman, D. & Chialvo, D. Criticality in Large-Scale Brain fMRI Dynamics Unveiled by a Novel Point Process Analysis. *Frontiers in Physiology* **3** (2012).
- [17] Liu, X. & Duyn, J. H. Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proceedings of the National Academy of Sciences* **110**, 4392–4397 (2013).
- [18] Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage* **17**, 825–841 (2002).
- [19] Yeo, B. T. *et al.* The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *Journal of neurophysiology* (2011).