

# Higher-order organization of multivariate time series

Received: 18 March 2022

Accepted: 21 October 2022

Published online: 2 January 2023

 Check for updates

Andrea Santoro<sup>1</sup>, Federico Battiston<sup>2</sup>, Giovanni Petri<sup>3</sup> & Enrico Amico<sup>1,4</sup> 

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. Several methods have been proposed for the analysis of multivariate time series, yet most of them neglect the effect of non-pairwise interactions on the emerging dynamics. Here, we propose a framework to characterize the temporal evolution of higher-order dependencies within multivariate time series. Using network analysis and topology, we show that our framework robustly differentiates various spatiotemporal regimes of coupled chaotic maps. This includes 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 light on the higher-order organization of multivariate time series, allowing a better characterization of dynamical group dependencies inherent to real-world data.

Complex networks have emerged as a powerful framework to investigate the structure and dynamics of many real-world systems<sup>1</sup> and have helped to characterize several phenomena<sup>2</sup>, including synchronization<sup>3</sup>, diffusion<sup>4</sup> and opinion formation<sup>5</sup>.

Despite being widely considered the reference model for many real-world complex systems<sup>6</sup>, networks are limited to describing interactions between two nodes at a time. This however clashes with the growing empirical evidence for group interactions in social systems<sup>7</sup>, neuroscience<sup>8–10</sup>, ecology<sup>11</sup> and biology<sup>12</sup>, where 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<sup>13,14</sup>, several studies have shown that the presence of higher-order interactions substantially impacts the dynamics of interacting systems<sup>15,16</sup>, ranging from alterations of synchronization<sup>17</sup> and diffusion<sup>18,19</sup> properties, to new collective dynamics in social<sup>20–22</sup> and evolutionary processes<sup>23</sup>.

Yet, measurements of pairwise or group interactions are rarely available. Hence, one must 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 underlying interactions. Examples of these complex patterns are observed in brain's neuronal activity, supporting a variety of motor and cognitive functions<sup>24,25</sup>, or in financial markets, where partially synchronized patterns often reflect periods of financial stress<sup>26,27</sup>. While the inference of pairwise interactions has a long history<sup>28</sup>, researchers have only recently taken the first steps towards reconstructing or filtering higher-order interactions<sup>29–32</sup>. In particular, methods relying on pairwise statistics might be insufficient as significant information can only be present in the joint probability distribution and not in the pairwise marginals, therefore failing at identifying higher-order behaviours<sup>33</sup>. To date, it remains unclear to what degree the information encoded in multivariate time series (MTS) stems from independent individual entities or, rather, from their group interactions. Existing proposals to address this issue are limited in their capacity to describe either the temporality or the complexity of such higher-order interactions, with few exceptions<sup>34</sup>. For instance, recent information theoretic methods characterized higher-order

<sup>1</sup>Neuro-X Institute, École Polytechnique Fédérale de Lausanne (EPFL), Geneva, Switzerland. <sup>2</sup>Department of Network and Data Science, Central European University, Vienna, Austria. <sup>3</sup>CENTAI, Turin, Italy. <sup>4</sup>Department of Radiology and Medical Informatics, University of Geneva, Geneva, Switzerland.  
 e-mail: [enrico.amico@epfl.ch](mailto:enrico.amico@epfl.ch)

dependencies in MTS by quantifying the intrinsic statistical synergy and redundancy in groups of interacting variables<sup>35,36</sup>. Moreover, an approach<sup>37</sup> rooted in random matrix theory has proven suitable to unveil the mesoscopic organization of correlation matrices. Yet, these methods hardly capture the information about the system dynamics, because they require integration over time. By contrast, tools of network neuroscience and signal processing easily deal with the dynamics of MTS<sup>38–42</sup> but only focus on pairwise statistics, neglecting higher-order effects. As a result, a principled approach to quantify the instantaneous dynamics of groups of nodes and possibly infer its higher-order representation is missing.

Here, we propose a 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 first demonstrate that, unlike traditional tools of time series analysis, higher-order measures reveal the subtleties of different spatiotemporal regimes displayed by canonical models of spatiotemporal chaos at the level of individual frames. We then use the insights from these synthetic models as a Rosetta Stone to interpret the higher-order structures reconstructed from time series concerning three real-world case studies: functional magnetic resonance imaging (fMRI) resting-state brain activity, stock option prices and the epidemiological incidence of various diseases in the United States. In all cases, we unveil additional higher-order information that is not captured at the node and dyadic network level, and highlight distinct topological dynamical regimes, which 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 for further tasks on real data sets, 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 MTS

Simplicial complexes are well suited as a modelling framework to describe the co-existence of pairwise and higher-order interactions<sup>13</sup>. A  $k$ -simplex  $\sigma$  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  $\sigma$ , all its possible subsfaces (defined as subsets of  $\sigma$ ) are themselves contained in  $\sigma$  (see Methods for details). Via this representation, it is then easy to distinguish between a group interaction among three elements, 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.

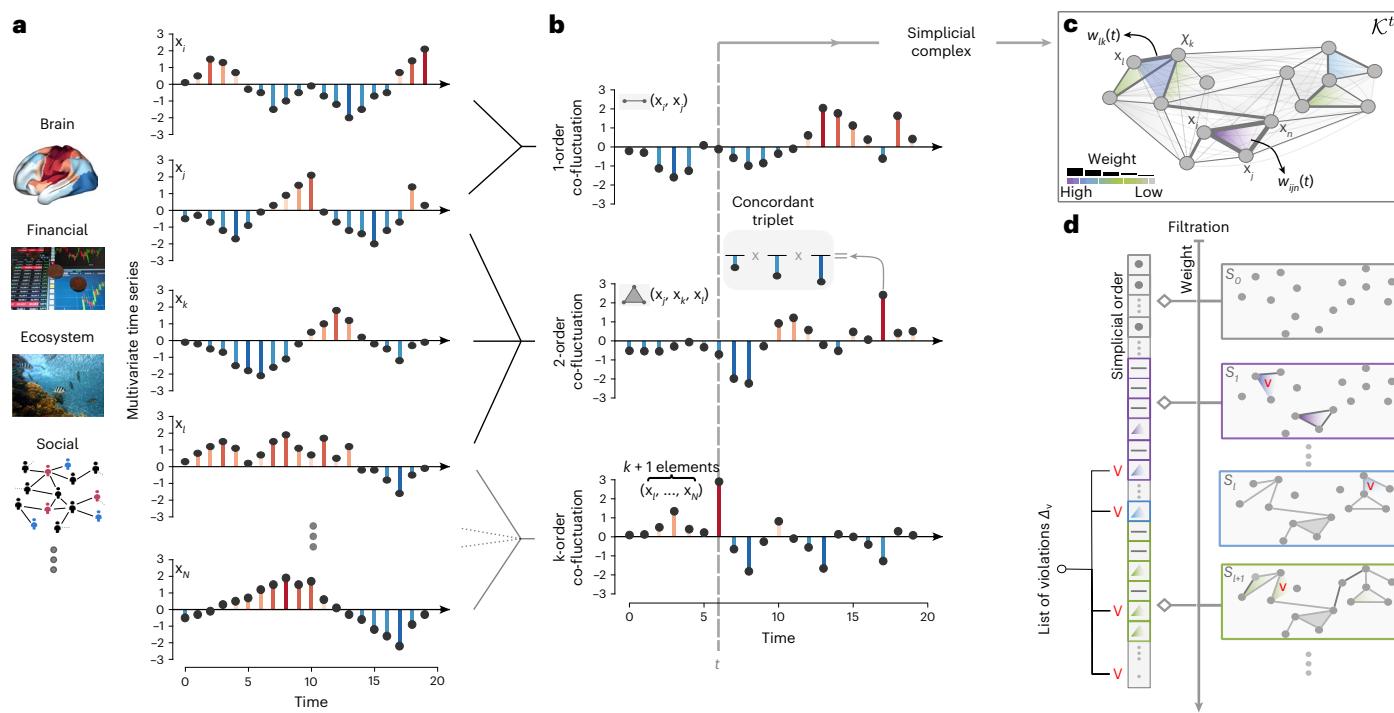
We rely on this representation to describe the higher-order dependencies between time series. Our approach has four main steps: (1) We z-score the  $N$  original time series (Fig. 1a) and calculate the element-wise product of the z-scored time series for all the  $\binom{N}{k}$   $k$ -order patterns (that is, edges, triangles, etc.). Here, the generic elements represent the instantaneous co-fluctuation magnitude between a  $(k+1)$ -group interaction. (2) The resulting set of time series encoding the  $k$ -order co-fluctuations (that is, the so-called edge time series<sup>40</sup> in the case of 1-order co-fluctuation) are then 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 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 represent different regimes of group synchronization. Hence, we assigned positive signs to fully concordant group interactions and negative signs to 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. (3) For each time  $t$ , we condense all instantaneous  $k$ -order co-fluctuations in a single mathematical object, that is, a weighted simplicial complex  $\mathcal{K}^t$  (Fig. 1c). (4) For each time  $t$ , we construct a filtration  $\mathbb{F}(\mathcal{K}^t)$  (ref. <sup>43</sup>), that is, a sequence of simplicial complexes  $\mathcal{S}_0 \subset \mathcal{S}_1 \subset \dots \subset \mathcal{S}_l \subset \dots \subset \mathcal{S}_n \subset \mathcal{K}^t$ , by sorting all  $k$ -order co-fluctuations by their weights (see Methods for details). The filtration proceeds in a top-down fashion from larger to smaller weights—in the spirit of persistent homology<sup>43,44</sup>—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 (that is, descending from more coherent to less coherent patterns). Yet, to maintain valid simplicial complexes at each step of the filtration, only  $k$ -order simplices respecting the simplicial closure condition can be included. That is, simplices whose subfaces 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 (for example, 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 their subcomponents (see Fig. 1d and Methods for details). Note further that here we present results considering  $k=2$ , so we take into account simplices only up to triangles. Nevertheless, our framework generalizes to  $k \geq 3$ . In summary, for each time  $t$ , our framework produces two outputs:

1. A list of violating triangles,  $\Delta_v = \{(i, j, k), w_{ijk}\}$ , induced by the simplicial closure condition. These are triangles whose weights co-fluctuate more than at least one of their corresponding edges. Intuitively, these triangles reflect higher-order states that cannot be captured by pairwise co-fluctuations. We then define the hyper-coherence indicator as the fraction of violating coherent triangles (that is, violating triangles with a weight greater than zero) over all the possible coherent triangles (that is, triangles with a weight greater than zero).
2. The simplicial filtration  $\mathbb{F}$ , a sequence of inclusions of simplicial complexes—sorted according to coherent patterns—starting with the empty complex and ending with the entire simplicial complex (Fig. 1d, right). We then compute the persistent homology of  $\mathbb{F}$  to characterize the persistence of certain topological features such as connected components, one-dimensional (1D) cycles, three-dimensional (3D) cavities, etc.<sup>44,45</sup>. Here, we focus on examining the 1D cycles in the filtration, that is, the persistent generators of the first homology group  $H_1$ , which provide insights about where and when higher synchronized 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  $\mathbb{F}$  (Supplementary Fig. S1). We rely on this object and define the hyper-complexity indicator as the Wasserstein distance<sup>46</sup> 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 then obtain 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 that the hyper-coherence and hyper-complexity distinguish different dynamical regimes generated by canonical models of spatiotemporal chaos. As a case study, we consider diffusively coupled



**Fig. 1 | Higher-order structure of MTS: schematic representation.** **a**, We start by extracting different co-fluctuation patterns according to their order from the ‘raw’ time series  $\{x_i\}_{i=1,\dots,N}$ . **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 (that is, 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 the identification of weighted holes when  $k$ -order patterns are gradually included. To maintain valid simplicial complexes at each step of the filtration, only  $k$ -order simplices respecting the simplicial closure condition can be included, while those simplices violating such a condition (e.g., a triangle entering the complex before its edges) are marked as simplicial violations ‘V’. For simplicity, we set  $k = 2$  for **c** and **d**.

map lattices (CMLs)<sup>47</sup>, which are high-dimensional dynamical systems defined on discrete time and space, with continuous state variables. We consider a ring lattice with  $N$  sites and 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

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

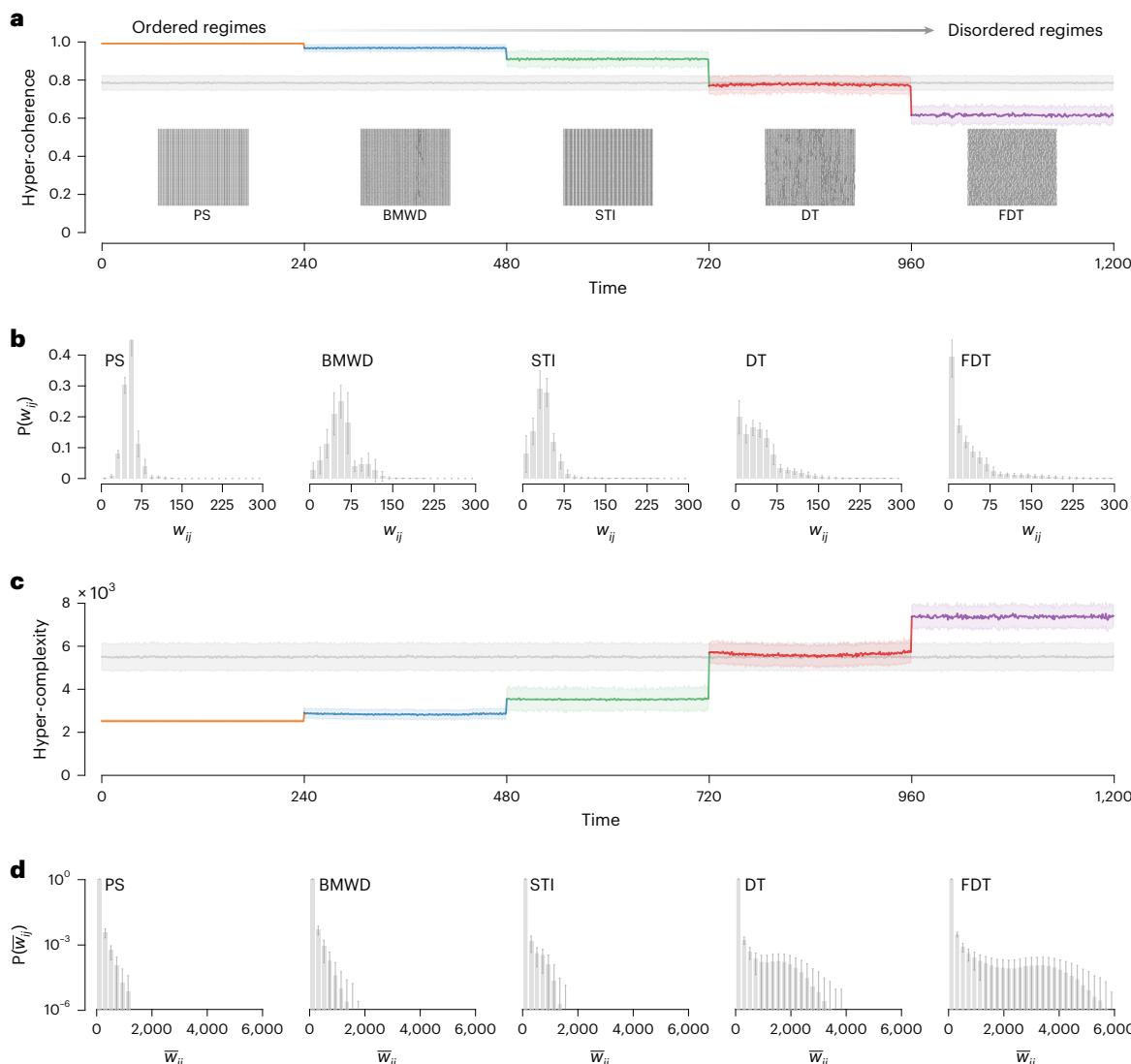
where  $i = 1, 2, \dots, N$ ,  $\varepsilon \in [0, 1]$  is the coupling strength and  $f[x]$  is generally a chaotic map. In our simulations, we considered the logistic map, that is  $f[x] = 1 - 1.75x^2$ . It is well established<sup>48</sup> that, by changing the values of  $\varepsilon$ , CMLs exhibit several 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 a randomly selected periodic attractor, reflecting quasi-periodic behaviours) and different forms of spatiotemporal intermittency (STI, chaotic pseudo-phases with low-dimensional attractors interpolating between FDT and PS). Moreover, we highlight two other phases: Brownian motion with defects (BMWD, a phase where defects fluctuate chaotically akin to Brownian motion) and defect turbulence (DT, a phase where many defects are generated and turbulently collide together)<sup>49</sup>.

Figure 2 summarizes the results of our higher-order approach when applied to these synthetic MTS with  $N = 119$  nodes and  $T = 1,200$ , obtained by concatenating five dynamical phases of CMLs with fixed

time length  $L = 240$ . Remarkably, the hyper-coherence indicator (Fig. 2a) clearly distinguishes the different phases of the CMLs, while preserving the ranking between ordered and disordered states. It assigns high values to fully and partially synchronized regimes, while chaotic regimes exhibit lower values of hyper-coherence. Interestingly, while this indicator provides only global information, refined information can be obtained by projecting the magnitudes of the list of violating triangles  $\Delta_v$  as a weighted graph (see Methods for the definition of downwards projections). 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 MTS, the edge weight distribution gradually changes its shape, with the limit case of the FDT chaotic series converging towards a fat-tailed distribution.

Notably, when investigating the temporal evolution of the hyper-complexity (Fig. 2c), the lowest value is assigned to periodic patterns, 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 regimes of CMLs, one might assume that the two higher-order indicators provide equivalent information as indicated by the negative correlation (that is, Spearman’s  $\rho \approx -0.95$ ). However, we will show that this is not generally true for real-world MTS.

Finally, Fig. 2d reports the edge weight distributions of the persistence homological scaffolds, graphs constructed from the persistent homology generators of  $H_1$  (see Methods and ref.<sup>8</sup> for details). These distributions quantify the topological importance of edges in the co-fluctuation landscape in terms of the persistence of their



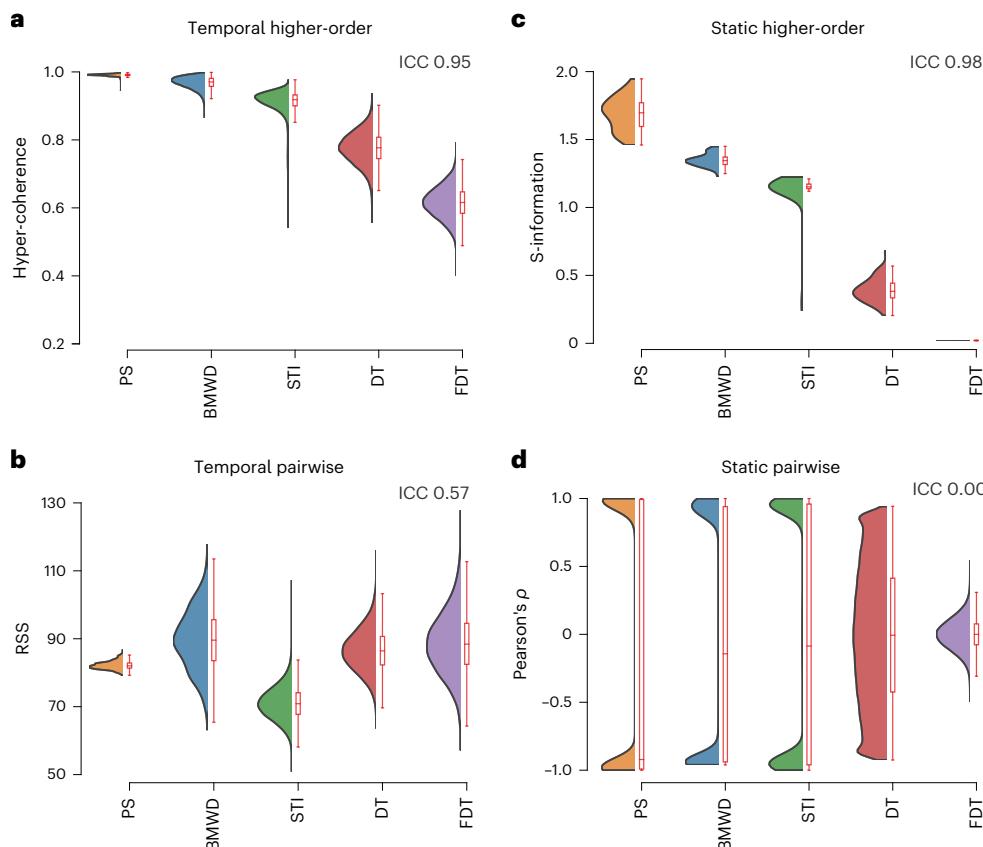
**Fig. 2 | Global and local higher-order indicators distinguish the dynamical regimes of coupled chaotic maps.** **a**, The temporal evolution of the hyper-coherence indicator for an MTS with  $N = 119$  nodes and  $T = 1,200$ , obtained by concatenating five different CML regimes with fixed time length  $L = 240$  (from order to disorder): PS at  $\varepsilon = 0.12$ , BMWD at  $\varepsilon = 0.08$ , STI II at  $\varepsilon = 0.3$ , DT at  $\varepsilon = 0.068$  and FDT at  $\varepsilon = 0.05$  (ref. <sup>48</sup>), for which a transient of  $10^5$  time points has been removed. A sample of such MTS is reported in Supplementary Fig. S2, while the effect of the z-score is studied in Supplementary Fig. S3. **b**, Notably, when projecting the list of violating triangles  $\Delta_v$  as a weighted graph (see Methods for

the definition of downwards projections), the edge weight distribution  $P(w_{ij})$  reflects the nature and the ‘ranking’ of the different dynamical regimes from ordered to disordered. **c,d**, The temporal evolution of the hyper-complexity indicator (**c**), and the distribution of weights  $P(\bar{w}_{ij})$  of the homological scaffold constructed from the persistent homology generators of  $H_1$  (**d**)<sup>8</sup> also mirror the different nature of the dynamical regimes. For comparison, **a** and **c** show the same indicators for a null model obtained when reshuffling without any constraint the MTS (grey curve). Shaded regions and error bars represent the s.d. across 100 independent realizations.

homological generators. Remarkably, also these distributions change their overall shape as we move from periodic to chaotic MTS but preserving the rank between order and disorder.

We quantitatively assess the capacity of our higher-order indicators to differentiate between dynamical regimes with the intraclass correlation coefficient (ICC)<sup>50</sup>, a statistical measure used to determine the agreement between ratings/scores of different groups. In other words, the stronger the agreement between groups, the higher the ICC value. Figure 3 reports the comparison of several approaches when differentiating the five regimes of CML. We find that both our higher-order measures (in Fig. 3a only hyper-coherence is shown) have high ICC values (that is, approximately 0.95 and 0.96 for hyper-coherence and hyper-complexity, respectively). For a commonly used temporal low-order measure, the root sum square (RSS) statistic<sup>41</sup>, which accounts for the

magnitude of peak amplitude in all the 1-order co-fluctuations, that is, the edge time series (see Methods for a formal definition), we find instead considerably smaller ICC values (approximately 0.57; Fig. 3b), implying that the higher-order effects are dominant. Indeed, our dynamic approach is comparable to ‘static’ higher-order information theoretic approaches<sup>35,51</sup>. That is, for S-information (Fig. 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 spatial information. Finally, (both static and dynamic) higher-order measures consistently outperform static lower-order methods based on Pearson’s correlation (that is,  $ICC \approx 0$ ; Fig. 3d). Hence, it appears clear that only higher-order approaches effectively distinguish the various spatiotemporal regimes, while lower-order statistics fail to capture the subtle differences between dynamical states.



**Fig. 3 | Higher-order approaches perform better in distinguishing the CML regimes.** **a–d**, Violin plots of the hyper-coherence (a), RSS (b)<sup>41</sup>, S-information (c)<sup>35,51</sup> and Person's correlation distributions (d), for the five dynamical regimes. Notably, only higher-order approaches are able to distinguish the five

dynamical regimes (that is,  $ICC > 0.9$ ). Also notice that static approaches can only be used when having prior knowledge on the position of each block. For detailed comparisons with other static and dynamic approaches<sup>37,41</sup>, see also Supplementary Figs. S4 and S5.

### Real-world complex systems exhibit non-trivial hyper-coherence structure

As examples of real-world applications, we consider our framework on fMRI signals from the Human Connectome Project (HCP)<sup>52</sup>, on prices of financial assets in the New York Stock Exchange (NYSE) and on historical data of infectious diseases in the United States<sup>53</sup>.

For the human brain data, we consider resting-state fMRI signals of 100 unrelated subjects, employing a cortical parcellation of 100 cortical<sup>54</sup> and 19 sub-cortical brain regions as provided by the HCP release<sup>55</sup>, for a total of  $N = 119$  regions of interest (ROIs). For the financial data, we analyse the daily evolution of  $N = 119$  U.S. stock prices from the NYSE over the period 2000–2021. Finally, for the epidemiological data set, we investigate the weekly number of cases at the U.S. state level ( $N = 50$ ) for seven infectious diseases (see Methods for details on the data sets).

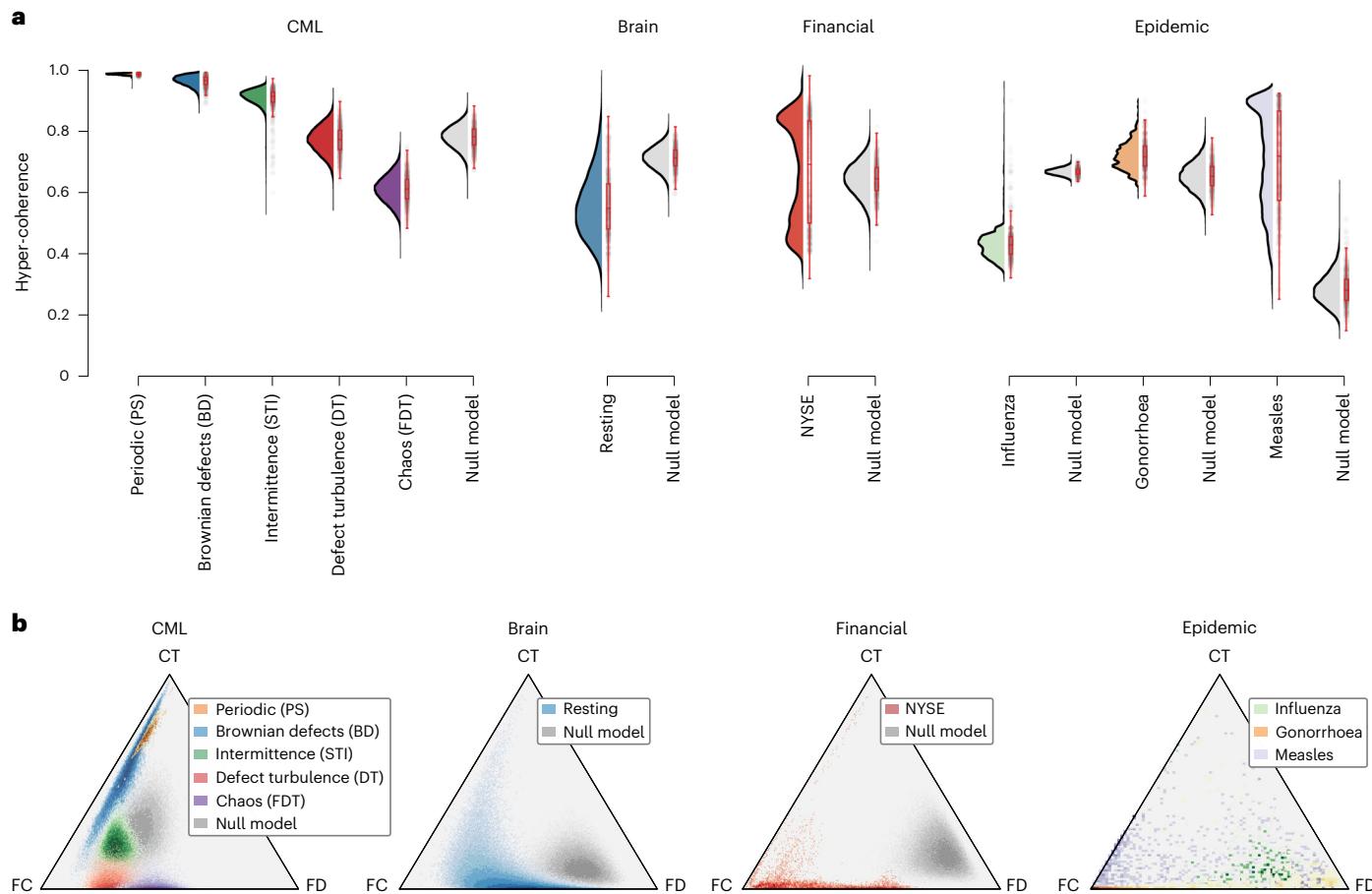
Figure 4 reports the distributions of the hyper-coherence for the CML regimes and the three data sets. For comparison, we plot the null models obtained by independently reshuffling synthetic and real-world MTS. Interestingly, these distributions are all statistically distinct from the corresponding null models ( $P < 10^{-10}$ , with the Kolmogorov–Smirnov test), yet they exhibit specific profiles that strongly differ from each other. In the epidemiological data, it is 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 synchronization, whereas periods of financial stability seem

to unfold in a more chaotic fashion. Moreover, armed with the CML 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<sup>56–58</sup>.

### Hyper-complexity decomposition provides detailed information about dynamical regimes

To characterize the evolution of 1D homological generators in the landscape of co-fluctuations, it is possible to decompose the hyper-complexity indicator into three 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, that is, edges and triangles having weights larger than zero, which we denote as a full coherence (FC) contribution; 1D cycles formed by coherent structures and closed by decoherent ones (that is, edges and triangles with weights smaller than zero), which we denote as a coherence transition (CT) contribution; and finally, 1D cycles created only by the fully decoherent structures, which we denote as a full decoherence (FD) contribution. By construction, these three contributions sum up to the total hyper-complexity. We show an illustrative example in Supplementary Fig. S1.

Figure 4b shows the three fractional contributions in a triangular representation. In this space, a point is placed at the bottom left corner if all the 1D cycles are formed and closed only by FC structures. Likewise, the bottom right corner (resp. top corner) corresponds to an exclusive contribution from FC (resp. CT) structures. Whenever the hyper-coherence indicator splits into similar FC and FD contributions, the point is placed between the corresponding corners,



**Fig. 4 | Higher-order indicators for real-world MTS.** **a**, Violin plots showing the distribution of hyper-coherence for three real-world data sets, namely resting-state fMRI data ( $N = 119$  brain regions), financial prices of 119 assets on the NYSE, and historical U.S. data for several infectious diseases at the U.S. state level ( $N = 50$ ). The real distributions are compared against the five CML dynamical regimes, as well as the corresponding null models (NMs) obtained when independently reshuffling without any constraint synthetic and real-world MTS (see Supplementary Figs. S6–S9 for the behaviour of the higher-order indicators in more conservative null models). Note how the distributions of the three real-world data sets employed exhibit noticeable differences in their profile

yet are always statistically distinct from the corresponding null models. See also Supplementary Fig. S10 for a similar analysis with edge-based indicators. **b**, Two-dimensional histograms 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 FC, CT and FD, while a corner position is reserved for points whose mainly contributions come from FC, CT or FD.

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 were split into three equal contributions from FC, CT and FD. Note that such decomposition carries completely different information with respect to the hyper-coherence indicator, and yet we draw analogies similar to the results presented in the previous section.

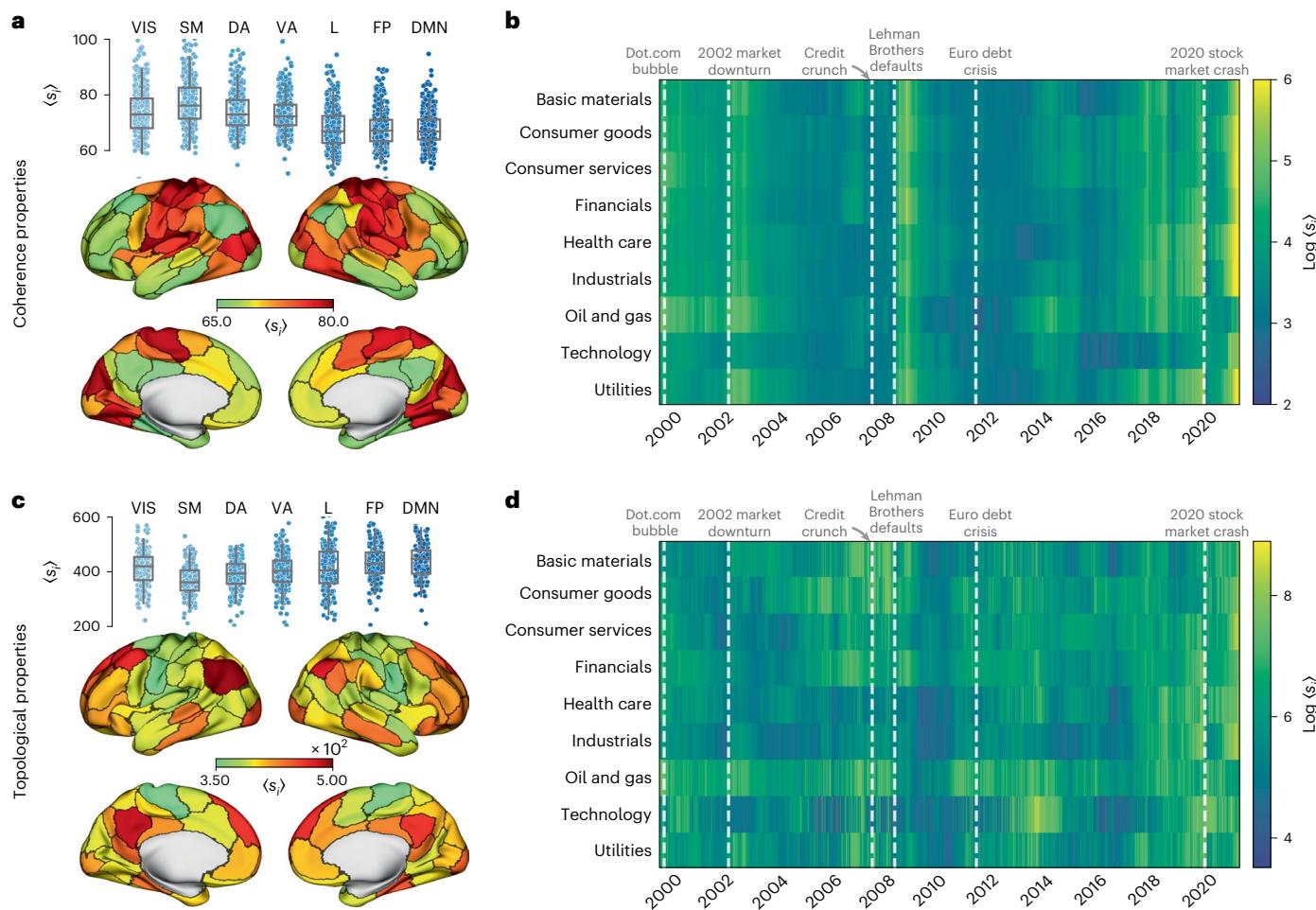
Indeed, when examining the different contributions of the hyper-complexity indicator in synthetic signals, we find that all CML regimes separate into different clusters. Partially synchronized signals are characterized by a mixture of FC and CT contributions, while chaotic signals are mainly determined by FC and FD (Fig. 4b, left). 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 a different nature. That is, it is only based on the number of simplicial violations.

#### Real-world applications of higher-order topological markers

We now report some representative applications when considering higher-order measures on a more local level, which are not deducible

from lower-order approaches (see Supplementary Figs. S11 and S12 and Supplementary Table S1 for comparison). 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% most coherent frames, which are those associated with a more synchronized dynamical phase. Figure 5a reports a brain map of the most discriminative nodes by projecting the magnitudes of the violating triangles  $\Delta_v$  on a nodal level (see Methods for details and Supplementary Fig. S13 for comparisons at other percentages). This is equivalent to considering the nodal strength extracted from  $\Delta_v$ . In other words, regions with the highest absolute value are the ones belonging to the most coherent higher-order structures. We find that 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<sup>59</sup>. This is confirmed when considering the histogram reporting the mean coherence within the seven canonical functional networks<sup>60</sup> (see also Supplementary Fig. S14 for the effects of higher-order indicators between the functional networks).

Figure 5b reports the temporal evolution of the nodal strength extracted from  $\Delta_v$ , but aggregated at the level of industrial sectors, for



**Fig. 5 | Projection of higher-order measures provides local spatiotemporal information.** The nodal strength extracted from the violating triangles  $\Delta_v$  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% most coherent frames, which are the ones associated with a more synchronized dynamics during rest. The box plots report the mean coherence within Yeo's seven canonical functional networks<sup>60</sup>: visual (VIS), somatomotor (SM), dorsal attention (DA), ventral attention (VA), limbic (L), frontoparietal (FP) and default mode network (DMN), where the box denotes the interquartile range (IQR), the horizontal bar indicates the median and the whiskers include

points that are within  $1.5 \times$  the IQR of its upper and lower bounds (the 25th and 75th percentiles). **b**, The temporal evolution of the nodal strength of violating triangles  $\Delta_v$  at the level of industrial sectors discriminates crises from periods of financial stability. **c**, The brain map obtained when selecting the 15% of the frames with lowest hyper-complexity mainly encompasses the DMN, 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 in 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.

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 distinguishes them from the unsynchronized intervals of 2002–2007 and 2013–2018, which in turn corresponds to more stable periods 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% of frames with lowest hyper-complexity, which are those associated with more synchronized dynamical phases. 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<sup>61</sup>.

By contrast, for the financial data in Fig. 5d, the temporal evolution of the nodal strength of the homological scaffold provides fine details

on the downturns in certain economic sectors. For instance, consumer goods, basic materials and oil/gas are the main sectors affected by the great recession of 2007.

Finally, we highlight that higher-order indicators can be further used to classify disease type from spatial spreading patterns. Owing to space limitations, we report these results only in Supplementary Fig. S15 and Supplementary Table S2.

## Discussion

Inferring the dynamics of higher-order structures in MTS is of utmost importance in many complex systems. However, higher-order network measurements are often inaccessible<sup>15</sup>. Our approach provides the first flexible method to dynamically reconstruct higher-order interactions from nodal-based MTS.

As a starting benchmark, we have validated our method against signals whose underlying dynamics is well known. In particular, differently from various lower-order statistics<sup>37,41</sup>, the proposed higher-order indicators robustly separate several dynamical phases in high-dimensional coupled chaotic maps, which appear to be distinguishable only through

methods based on higher-order statistics<sup>35</sup> (Supplementary Figs. S4 and S5). This provides further empirical evidence on the need for higher-order approaches to identify higher-order behaviours<sup>33</sup>.

When analysing real-world data, we found that, during rest, the human brain higher-order dynamics mainly oscillates between turbulence and partial synchronization. This is in agreement with recent studies supporting that the human brain operates in a turbulent regime<sup>36</sup>, at the edge of criticality<sup>62</sup>, which is thought to confer significant information processing advantages<sup>63</sup>. Moreover, when analysing brain states on a finer scale, we found two notable 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<sup>59</sup>. On the other hand, when examining the hyper-complexity marker at its lowest points, we found that the nodal projection coarsely captures the DMN, which is known to integrate high- and low-order information in human brain networks<sup>61</sup>. Hence, our two proposed markers provide complementary insights, which are not trivially deducible from an edge-wise approach (Supplementary Figs. S11 and S12), on how the brain network segregates and integrates higher-order information over time<sup>42,64</sup>.

In 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 (Supplementary Fig. S8). In particular, maximally coherent higher-order structures emerge in correspondence with major financial crises, mirroring the increase of synchronous co-activation patterns. While this is not new in the literature<sup>27,49,65</sup>, we stress that, unlike our method, most of those previous approaches rely on correlation matrices estimated over sliding time windows<sup>26</sup>, therefore neglecting the information that one might want to capture at the level of individual frames (for example, in high-frequency trading<sup>66</sup>). 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<sup>67</sup>, which suggests their potential to identify the building up of systemic risk<sup>68</sup>.

Taken together, here we have developed a new flexible tool to provide framewise estimates of higher-order structures in MTS. We believe that our framework can be used effectively in all situations where the dynamics of signals is poorly understood or unknown, paving the way towards further applications in the fields of biology, social sciences or clinical neuroscience.

## Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41567-022-01852-0>.

## References

- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D.-U. Complex networks: structure and dynamics. *Phys. Rep.* **424**, 175–308 (2006).
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925–979 (2015).
- Arenas, A., Díaz-Guilera, A., Kurths, J., Moreno, Y. & Zhou, C. Synchronization in complex networks. *Phys. Rep.* **469**, 93–153 (2008).
- Barrat, A., Barthélémy, M. & Vespignani, A. *Dynamical Processes on Complex Networks* (Cambridge Univ. Press, 2008).
- Watts, D. J. & Dodds, P. S. Influentials, networks, and public opinion formation. *J. Consum. Res.* **34**, 441–458 (2007).
- Latora, V., Nicosia, V. & Russo, G. *Complex Networks: Principles, Methods and Applications* (Cambridge Univ. Press, 2017).
- Benson, A. R., Gleich, D. F. & Leskovec, J. Higher-order organization of complex networks. *Science* **353**, 163–166 (2016).
- Petri, G. et al. Homological scaffolds of brain functional networks. *J. R. Soc. Interface* **10**, 186–198 (2014).
- Giusti, C., Pastalkova, E., Curto, C. & Itskov, V. Clique topology reveals intrinsic geometric structure in neural correlations. *Proc. Natl Acad. Sci. USA* **112**, 13455–13460 (2015).
- Sizemore, A. E. et al. Cliques and cavities in the human connectome. *J. Comput. Neurosci.* **44**, 115–145 (2018).
- Grilli, J., Barabás, G., Michalska-Smith, M. J. & Allesina, S. Higher-order interactions stabilize dynamics in competitive network models. *Nature* **548**, 210–213 (2017).
- Sanchez-Gorostiaga, A., Bajić, D., Osborne, M. L., Poyatos, J. F. & Sanchez, A. High-order interactions distort the functional landscape of microbial consortia. *PLOS Biol.* **17**, 3000550 (2019).
- Battiston, F. et al. Networks beyond pairwise interactions: structure and dynamics. *Phys. Rep.* **874**, 1–92 (2020).
- Torres, L., Blevins, A. S., Bassett, D. & Eliassi-Rad, T. The why, how, and when of representations for complex systems. *SIAM Rev.* **63**, 435–485 (2021).
- Battiston, F. et al. The physics of higher-order interactions in complex systems. *Nat. Phys.* **17**, 1093–1098 (2021).
- Battiston, F. & Petri, G. *Higher-Order Systems* (Springer, 2022).
- Millán, A. P., Torres, J. J. & Bianconi, G. Synchronization in network geometries with finite spectral dimension. *Phys. Rev. E* **99**, 022307 (2019).
- Schaub, M. T., Benson, A. R., Horn, P., Lippner, G. & Jadbabaie, A. Random walks on simplicial complexes and the normalized Hodge 1-Laplacian. *SIAM Rev.* **62**, 353–391 (2020).
- Carletti, T., Battiston, F., Cencetti, G. & Fanelli, D. Random walks on hypergraphs. *Phys. Rev. E* **101**, 022308 (2020).
- Iacobini, I., Petri, G., Barrat, A. & Latora, V. Simplicial models of social contagion. *Nat. Commun.* **10**, 1–9 (2019).
- de Arruda, G. F., Petri, G. & Moreno, Y. Social contagion models on hypergraphs. *Phys. Rev. Res.* **2**, 023032 (2020).
- Sahasrabuddhe, R., Neuhäuser, L. & Lambiotte, R. Modelling non-linear consensus dynamics on hypergraphs. *J. Phys. Complex.* **2**, 025006 (2021).
- Alvarez-Rodriguez, U. et al. Evolutionary dynamics of higher-order interactions in social networks. *Nat. Hum. Behav.* **5**, 586–595 (2021).
- Deco, G., Jirsa, V. K. & McIntosh, A. R. Emerging concepts for the dynamical organization of resting-state activity in the brain. *Nat. Rev. Neurosci.* **12**, 43–56 (2011).
- Avena-Koenigsberger, A., Masic, B. & Sporns, O. Communication dynamics in complex brain networks. *Nat. Rev. Neurosci.* **19**, 17–33 (2018).
- Mantegna, R. N. & Stanley, H. E. *Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge Univ. Press, 1999).
- Peron, T. K. D. & Rodrigues, F. A. Collective behavior in financial markets. *Europhys. Lett.* **96**, 48004 (2011).
- Brugere, I., Gallagher, B. & Berger-Wolf, T. Y. Network structure inference, a survey: motivations, methods, and applications. *ACM Comput. Surv.* **51**, 24–12439 (2018).
- Young, J. G., Petri, G. & Peixoto, T. P. Hypergraph reconstruction from network data. *Commun. Phys.* **4**, 1–11 (2021).
- Muscicotto, F., Battiston, F. & Mantegna, R. N. Detecting informative higher-order interactions in statistically validated hypergraphs. *Commun. Phys.* **4**, 1–9 (2021).
- Wang, H., Ma, C., Chen, H.-S., Lai, Y.-C. & Zhang, H.-F. Full reconstruction of simplicial complexes from binary contagion and ising data. *Nat. Commun.* **13**, 3043 (2022).

32. Lizotte, S., Young, J.-G., Allard, A. Hypergraph reconstruction from noisy pairwise observations. Preprint at <https://arxiv.org/abs/2208.06503> (2022).
33. Rosas, F. E. et al. Disentangling high-order mechanisms and high-order behaviours in complex systems. *Nat. Phys.* **18**, 476–477 (2022).
34. Faes, L. et al. A New Framework for the Time- and Frequency-Domain Assessment of High-Order Interactions in Networks of Random Processes. In *IEEE Transactions on Signal Processing* 1–12 (IEEE, 2022).
35. Rosas, F. E., Mediano, P. A. M., Gastpar, M. & Jensen, H. J. Quantifying high-order interdependencies via multivariate extensions of the mutual information. *Phys. Rev. E* **100**, 32305 (2019).
36. Stramaglia, S., Scagliarini, T., Daniels, B. C. & Marinazzo, D. Quantifying dynamical high-order interdependencies from the O-information: an application to neural spiking dynamics. *Front. Physiol.* **11**, 1–11 (2021).
37. MacMahon, M. & Garlaschelli, D. Community detection for correlation matrices. *Phys. Rev. X* **5**, 021006 (2015).
38. Tagliazucchi, E., Balenzuela, P., Fraiman, D. & Chialvo, D. Criticality in large-scale brain fMRI dynamics unveiled by a novel point process analysis. *Front. Physiol.* **3**, 15 (2012).
39. Liu, X. & Duyn, J. H. Time-varying functional network information extracted from brief instances of spontaneous brain activity. *Proc. Natl Acad. Sci. USA* **110**, 4392–4397 (2013).
40. Faskowitz, J., Esfahlani, F. Z., Jo, Y., Sporns, O. & Betzel, R. F. Edge-centric functional network representations of human cerebral cortex reveal overlapping system-level architecture. *Nat. Neurosci.* **23**, 1644–1654 (2020).
41. Esfahlani, F. Z. et al. High-amplitude cofluctuations in cortical activity drive functional connectivity. *Proc. Natl Acad. Sci. USA* **117**, 28393–28401 (2020).
42. Van De Ville, D., Farouj, Y., Preti, M. G., Liégeois, R. & Amico, E. When makes you unique: temporality of the human brain fingerprint. *Sci. Adv.* **7**, 0751 (2021).
43. Petri, G., Scolamiero, M., Donato, I. & Vaccarino, F. Topological strata of weighted complex networks. *PLoS ONE* **8**, e66506 (2013).
44. Zomorodian, A. & Carlsson, G. Computing persistent homology. *Discrete Comput. Geom.* **33**, 249–274 (2005).
45. Ghrist, R. Barcodes: the persistent topology of data. *Bull. Am. Math. Soc.* **45**, 61–75 (2008).
46. Carrière, M., Cuturi, M., Oudot, S. Sliced Wasserstein kernel for persistence diagrams. In *Proceedings of the 34th International Conference on Machine Learning* **70**, 664–673. Proceedings of Machine Learning Research (PMLR, 2017).
47. Kaneko, K. Overview of coupled map lattices. *Chaos* **2**, 279–282 (1992).
48. Kaneko, K. Pattern dynamics in spatiotemporal chaos. *Physica D* **34**, 1–41 (1989).
49. Lacasa, L., Nicosia, V. & Latora, V. Network structure of multivariate time series. *Sci. Rep.* **5**, 1–9 (2015).
50. Shrout, P. E. & Fleiss, J. L. Intraclass correlations: uses in assessing rater reliability. *Psychol. Bull.* **86**, 420–428 (1979).
51. Gatica, M. et al. High-order interdependencies in the aging brain. *Brain Connect.* **11**, 734–744 (2021).
52. Van Essen, D. C. et al. The WU-Minn Human Connectome Project: an overview. *NeuroImage* **80**, 62–79 (2013).
53. Scarpino, S.V. & Petri, G. On the predictability of infectious disease outbreaks. *Nat. Commun.* **10**, 1–8 (2019).
54. Schaefer, A. et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* **28**, 3095–3114 (2018).
55. Glasser, M. F. et al. The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80**, 105–124 (2013).
56. Deco, G. & Kringelbach, M. L. Turbulent-like dynamics in the human brain. *Cell Rep.* **33**, 108471 (2020).
57. Deco, G., Kringelbach, M. L., Jirsa, V. K. & Ritter, P. The dynamics of resting fluctuations in the brain: metastability and its dynamical cortical core. *Sci. Rep.* **7**, 3095 (2017).
58. Chialvo, D. R. Emergent complex neural dynamics. *Nat. Phys.* **6**, 744–750 (2010).
59. Smith, S. M. et al. Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl Acad. Sci. USA* **106**, 13040–13045 (2009).
60. Yeo, B. T. et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
61. Raichle, M. E. et al. A default mode of brain function. *Proc. Natl Acad. Sci. USA* **98**, 676–682 (2001).
62. Perl, Y.S., Escrichs, A., Tagliazucchi, E., Kringelbach, M.L., Deco, G. On the edge of criticality: strength-dependent perturbation unveils delicate balance between fluctuation and oscillation in brain dynamics. Preprint at bioRxiv <https://doi.org/10.1101/2021.09.23.461520> (2022).
63. Deco, G. et al. Rare long-range cortical connections enhance human information processing. *Curr. Biol.* **31**, 4436–44485 (2021).
64. Amico, E., Arenas, A. & Goñi, J. Centralized and distributed cognitive task processing in the human connectome. *Netw. Neurosci.* **3**, 455–474 (2019).
65. Kutner, R. et al. Econophysics and sociophysics: their milestones & challenges. *Physica A* **516**, 240–253 (2019).
66. Musciotto, F., Pillo, J. & Mantegna, R. N. High-frequency trading and networked markets. *Proc. Natl Acad. Sci. USA* **118**, 2015573118 (2021).
67. Musmeci, N., Nicosia, V., Aste, T., Di Matteo, T. & Latora, V. The multiplex dependency structure of financial markets. *Complexity* **2017**, 9586064 (2017).
68. Squartini, T., van Lelyveld, I. & Garlaschelli, D. Early-warning signals of topological collapse in interbank networks. *Sci. Rep.* **3**, 3357 (2013).

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

© The Author(s), under exclusive licence to Springer Nature Limited 2023, corrected publication 2023

## Methods

### Higher-order topology of MTS

Consider an  $N$ -dimensional real-valued time series  $\{\mathbf{x}(t)\}_{t=1}^N$ , 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 an  $N$ -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<sup>69–72</sup>. 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.<sup>40</sup>, 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 generalize such a concept to the case of higher-order interactions, that is 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  $\mu[\cdot]$  and  $\sigma[\cdot]$  are the time-averaged mean and s.d. We can then calculate the generic element at time  $t$  of the z-scored  $k$ -order co-fluctuations between  $(k+1)$  time series as

$$\xi_{0\dots k}(t) = \frac{\prod_{p=0}^k z_p(t) - \mu\left[\prod_{p=0}^k z_p\right]}{\sigma\left[\prod_{p=0}^k z_p\right]}, \quad (1)$$

where also in this case  $\mu[\cdot]$  and  $\sigma[\cdot]$  are the time-averaged mean and s.d. functions. 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,

$$\text{sign}[\xi_{0\dots k}(t)] := (-1)^{\sum_0^k \text{sgn}[z_i(t)]}, \quad (2)$$

where  $\text{sgn}[\cdot]$  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

$$w_{0\dots k}(t) = \text{sign}[\xi_{0\dots k}(t)]|\xi_{0\dots k}(t)|. \quad (3)$$

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  $\sigma$  is defined as the set of  $d$  vertices, that is,  $\sigma = [p_0, p_1, \dots, p_{d-1}]$ . A collection of simplices is a simplicial complex  $\mathcal{K}$  if, for each simplex  $\sigma$ , all its possible subsurfaces (defined as subsets of  $\sigma$ ) are themselves contained in  $\mathcal{K}$  (ref.<sup>73</sup>). 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  $\mathcal{K}$  while weights on the simplices, that is,  $w_{ij}$  and  $w_{ijk}$ , represent the magnitude of edge and triangle co-fluctuations.

Note finally that, to compare our approach with the edge-based approach, we employed the RSS of the edge time series<sup>41</sup>, 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, that is,  $\text{RSS}(t) = \sqrt{\sum_{ij>i} e_{ij}(t)^2}$ , where the vector  $\mathbf{e}_{ij} = \mathbf{z}_i \mathbf{z}_j$  is the 1-order co-fluctuation (that is, 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<sup>74</sup>, which has been shown to unveil new dynamical properties of different complex systems<sup>75–78</sup>. In particular, we rely on persistent homology<sup>44,79</sup>, which is a recent technique in computational topology that has been largely used for the analysis of high-dimensional data sets<sup>80,81</sup> and in disparate applications<sup>82–84</sup>. The central idea is the construction of a sequence of successive simplicial complexes that approximates with increasing precision the original weighted simplicial complex. This sequence of simplicial complexes, that is,  $\emptyset = \mathcal{S}_0 \subset \mathcal{S}_1 \subset \dots \subset \mathcal{S}_l \subset \dots \subset \mathcal{S}_n$  is such that  $S_i \subset S_j$  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 sequence. Equivalently,  $\epsilon_l$  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, that is,  $\exists i, j: w_{ij} < w_{jk}$ . Such triangles are considered as a violation and inserted, along with the corresponding weights, into the list of violations  $\Delta_v = \{(i, j, k, w_{ijk})\}$ . The remaining links and triangles with a weight larger than  $\epsilon_l$  belong to the simplicial complex  $\mathcal{S}_l$

We then define the hyper-coherence indicator as the fraction of violating coherent triangles (that is, violating triangles with a weight greater than zero) over all the possible coherent triangles (that is, triangles with a weight greater than zero). Notice also that, when identifying each violating triangle (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 1D holes (that is, loops), therefore analysing the homology group  $H_1$ . More precisely, at each step of the filtration process, a generator  $g$  uniquely identifies a 1D cycle by its constituent elements. The importance of the one-dimensional hole  $g$  is encoded in the form of ‘time stamps’ recording its birth  $b_g$  and death  $d_g$  along the filtration  $\{\mathcal{S}_l\}$  (ref.<sup>8</sup>). 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 the 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 (that is, the space without  $(k+1)$ -dimensional holes). In our case, we define the hyper-complexity indicator as the Wasserstein distance<sup>46</sup> 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 Supplementary Fig. S16, we briefly investigate the presence of 1D cycles and 3D cavities in the context of CML when extending our framework to four-body interactions.

### Homological scaffold and lower-order projections

To obtain a finer description of the topological features present in the persistence diagram, we consider the persistence homological scaffold

as proposed in ref.<sup>8</sup>. In a nutshell, this object is a weighted network composed of all the cycle paths corresponding to generators  $g_i$  weighted by their persistence  $\pi_{g_i}$ . In other words, if an edge  $e$  belongs to multiple one-dimensional cycles  $g_0, g_1, \dots, g_s$ , its weight  $\bar{w}_e^{\pi}$  is defined as the sum of the generators' persistence:

$$\bar{w}_e^{\pi} = \sum_{g_i|e \in g_i} \pi_{g_i}. \quad (4)$$

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  $\bar{w}_e^{\pi}$  for a link  $e$  implies that such link acts as a locally strong bridge in the space of coherent and decoherent co-fluctuations<sup>8</sup>.

Lastly, to analyse the information provided by the list of violations  $\Delta_e$  on an edge/node level, we rely on downwards 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, that is, triangles of the form  $(i, j, \cdot)$  with a weight  $w_{ij}$ , and the average is computed over the number of triangles  $n_{ij}$  defined by that edge. Similarly, we define the nodal strength  $w_i$  of node  $i$  as the sum of weights of the triangles connected to node  $i$  after the edge projection. 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<sup>6,85–87</sup>, as the sum of the weights of edges connected to the node  $i$ .

### Real-world data sets

We analysed three data sets belonging to different domains. Specifically, we considered fMRI resting-state data from the HCP (<http://www.humanconnectome.org/>), the stock prices of the NYSE financial market obtained from the Yahoo! Finance application programming interface<sup>88</sup>, and the historical data of several infectious diseases in the United States<sup>53,89</sup>.

The fMRI data set used in this work consists of resting-state data from 100 unrelated subjects (54 females and 46 males; mean age  $29.1 \pm 3.7$  years) as provided at the HCP 900 subjects data release<sup>52,90</sup>. We added some extra steps to the HCP minimal preprocessing pipeline<sup>55,91</sup>. 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 0.01–0.15 Hz. Last, the voxel-wise fMRI time series were averaged into the  $N=100$  corresponding brain nodes of the Schaefer cortical atlas<sup>54</sup> and then z-scored. For completeness, 19 sub-cortical regions were added, as provided by the HCP release<sup>55</sup>. The interested reader can refer to ref.<sup>42</sup> for details on these steps.

The financial data set used in this study was obtained from the Yahoo! finance historical data application programming interface (via the Python library 'yfinance'<sup>88</sup>). We collected the daily prices of 119 U.S. companies on the NYSE from Yahoo! Finance in the period from 1 January 2000 to 17 June 2021.

We considered the weekly historical data at the U.S. state level of several infectious diseases including chlamydia, gonorrhoea, influenza, measles, polio and pertussis. This data set was previously used in ref.<sup>53</sup> 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 (that is, how much previous time points affect the next ones in terms of the proposed topological markers). Note also that our dynamical higher-order approach, as for many other

existing pairwise dynamical methods<sup>38,39,41,92–94</sup>, can be heavily affected by noisy fluctuations in the time series. However, this issue can be smoothed out by analysing statistics averaged over multiple time frames, as we have done in this work. Furthermore, the higher-order brain maps reported in Fig. 5 appear to be robust after accounting for the presence of head motion volumes in the fMRI data (see also Supplementary Fig. S17). Finally, we stress that our framework, with the exception of the hyper-complexity indicator, mainly detects coherent synchronous patterns but 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

All data needed to evaluate the conclusions of this paper are available on Zenodo ([10.5281/zenodo.7210075](https://doi.org/10.5281/zenodo.7210075)).

### Code availability

The code used in this work is available at: <https://github.com/andresanton/RHOSTS> as well as a maintained version on E.A.'s GitHub page (<https://github.com/eamico>).

### References

69. Wei, W. W. *Time Series Analysis* (Addison Wesley, 2005).
70. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**, 186–198 (2009).
71. Sporns, O. *Networks of the Brain* (MIT Press, 2010).
72. Wei, W. W. *Multivariate Time Series Analysis and Applications* (Wiley, 2019).
73. Hatcher, A. *Algebraic Topology* (Cambridge Univ. Press, 2005).
74. Wasserman, L. Topological data analysis. *Annu. Rev. Stat. Appl.* **5**, 501–532 (2016).
75. Lum, P. Y. et al. Extracting insights from the shape of complex data using topology. *Sci. Rep.* **3**, 1236 (2013).
76. Nicolau, M., Levine, A. J. & Carlsson, G. Topology based data analysis identifies a subgroup of breast cancers with a unique mutational profile and excellent survival. *Proc. Natl Acad. Sci. USA* **108**, 7265–7270 (2011).
77. Saggar, M. et al. Towards a new approach to reveal dynamical organization of the brain using topological data analysis. *Nat. Commun.* **9**, 1399 (2018).
78. Saggar, M., Shine, J. M., Liégeois, R., Dosenbach, N. U. F. & Fair, D. Precision dynamical mapping using topological data analysis reveals a hub-like transition state at rest. *Nat. Commun.* **13**, 4791 (2022).
79. Edelsbrunner, H., Letscher, D., Zomorodian, A. Topological persistence and simplification. In *Proc. of 41st Annual Symposium on Foundations of Computer Science* 454–463 (2000).
80. Ghrist, R. Barcodes: the persistent topology of data. *Bull. Am. Math. Soc.* **45**, 61–75 (2008).
81. Carlsson, G., Ishkhanov, T., de Silva, V. & Zomorodian, A. On the local behavior of spaces of natural images. *Int. J. Comput. Vis.* **76**, 1–12 (2008).
82. Lee, H., Chung, M.K., Kang, H., Kim, B.-N., Lee, D.S.: Discriminative persistent homology of brain networks. In *Proc. of 2011 IEEE International Symposium on Biomedical Imaging: from Nano to Macro* 841–844 (IEEE, 2011).
83. Carstens, C. J. & Horadam, K. J. Persistent homology of collaboration networks. *Math. Probl. Eng.* **2013**, 815035 (2013).
84. Horak, D., Maletić, S. & Rajković, M. Persistent homology of complex networks. *J. Stat. Mech. Theory Exp.* **2009**, 03034 (2009).
85. Barrat, A., Barthélémy, M., Pastor-Satorras, R. & Vespignani, A. The architecture of complex weighted networks. *Proc. Natl Acad. Sci. USA* **101**, 3747–3752 (2004).

86. Barabási, A.-L. *Network Science* (Cambridge Univ. Press, 2016).
87. Newman, M. *Networks* (Oxford Univ. Press, Oxford, 2018).
88. Aroussi, R., et al. Yahoo! Finance market data downloader. *GitHub* <https://github.com/ranaroussi/yfinance> (2022).
89. van Panhuis, W. G. et al. Contagious diseases in the United States from 1888 to the present. *N. Engl. J. Med.* **369**, 2152–2158 (2013).
90. Van Essen, D. C. et al. The Human Connectome Project: a data acquisition perspective. *NeuroImage* **62**, 2222–2231 (2012).
91. Smith, S. M. et al. Resting-state fMRI in the Human Connectome Project. *NeuroImage* **80**, 144–168 (2013).
92. Karahanoğlu, F. I. & Van De Ville, D. Transient brain activity disentangles fMRI resting-state dynamics in terms of spatially and temporally overlapping networks. *Nat. Commun.* **6**, 7751 (2015).
93. Preti, M. G., Bolton, T. A. & Van De Ville, D. The dynamic functional connectome: state-of-the-art and perspectives. *NeuroImage* **160**, 41–54 (2017).
94. Liu, X., Zhang, N., Chang, C. & Duyn, J. H. Co-activation patterns in resting-state fMRI signals. *NeuroImage* **180**, 485–494 (2018).

## Acknowledgements

We thank L. Lacasa and J. Goñi for feedback on an earlier version of the manuscript. Data were provided (in part) by the HCP, the WU-Minn Consortium (principal investigators D. V. Essen and K. Ugurbil; 1U54MH091657) funded by the 16 NIH Institutes and Centers that support the NIH Blueprint for Neuroscience Research and by the McDonnell Center for Systems Neuroscience at Washington University. E.A. acknowledges financial support from the SNSF Ambizione project ‘Fingerprinting the brain: network science to extract features

of cognition, behaviour and dysfunction’ (grant no. PZ00P2\_185716). A.S. and E.A. acknowledge support from the SNSF COST project ‘Mapping the higher-order dynamics of neurodegeneration in human brain networks’ (grant no. IZCOZO\_198144). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author contributions

A.S., F.B., G.P. and E.A. conceptualized the study. A.S. performed the numerical analysis. A.S., F.B., G.P. and E.A. wrote the paper.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41567-022-01852-0>.

**Correspondence and requests for materials** should be addressed to Enrico Amico.

**Peer review information** *Nature Physics* thanks Giulio Cimini, Bratislav Misic and the other, anonymous, reviewer(s) for their contribution to the peer review of this work

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).