Higher-order Laplacian renormalization

Higher-order Laplacian Renormalization

Marco Nurisso,[1,2,3] Marta Morandini,[2,4] Maxime Lucas,[2] Francesco Vaccarino,[1,3] Tommaso Gili,[5,*] and Giovanni Petri[6,2,7,†]

[1] *Dipartimento di Scienze Matematiche, Politecnico di Torino, Turin, 10129, Italy*
[2] *CENTAI Institute, Turin, 10138, Italy*
[3] *SmartData@PoliTO Center, Politecnico di Torino, Turin, 10129, Italy*
[4] *Institut de Neurosciences de La Timone, UMR 7289, CNRS, Aix-Marseille Université, Marseille 13005, France*
[5] *Networks Unit, IMT Scuola Alti Studi Lucca, Lucca, 55100, Italy*
[6] *NPLab, Network Science Institute, Northeastern University London, London, UK*
[7] *Department of Physics, Northeastern University, Boston, MA 02115, USA*

We propose a cross-order Laplacian renormalization group (X-LRG) scheme for arbitrary higher-order networks. The renormalization group (RG) is a pillar of the theory of scaling, scale-invariance, and universality in physics. An RG scheme based on diffusion dynamics was recently introduced for complex networks with dyadic interactions. Despite mounting evidence of the importance of polyadic interactions, we still lack a general RG scheme for higher-order networks. Our approach uses a diffusion process to group nodes or hyperedges, where information can flow between nodes and between hyperedges (higher-order interactions). This approach allows us (i) to probe higher-order structures, defining scale-invariance at various orders, and (ii) to propose a coarse-graining scheme. We demonstrate our approach on controlled synthetic higher-order systems and then use it to detect the presence of order-specific scale-invariant profiles of real-world complex systems from multiple domains.

I. INTRODUCTION

The renormalization group (RG) [1] is a cornerstone of modern theoretical physics because it allows us to study how a physical system depends on the scale of observation, defining universality classes and, importantly, formalizing the concept of scale-invariance. While the RG has been a powerful tool for understanding a broad class of physical systems, extending its framework to complex networks has posed a recent and significant challenge, mainly due to the correlations between scales caused by small-world effects [2]. This challenge has gained substantial attention [3–10] due to its potential to provide insights into the multiscale structural organization of complex networks.

Notable approaches [6, 11] are based on the hypothesis that an embedding space exists and is responsible for the network structure. This underlying geometry provides a natural way to identify groups of nearby nodes, reminiscent of spin blocks in the traditional real-space RG process [12]. These groups can then be collapsed into "super-nodes", providing a coarse-grained network description. However, this perspective encounters a fundamental limitation: networks are inherently topological structures, devoid of geometry, and thus need a topological notion of RG [7, 13].

Diffusion provides such a notion: it is a dynamical process that one can define on any combinatorial structure and depends only on the structure's topology. Diffusion on graphs describes the dynamics of information flowing from node to node through edges, eventually becoming uniformly distributed on its connected components. This process is formalized as a first-order system of linear differential equations specified by the graph Laplacian matrix $\mathbf{L}$. One can then see the diffusion time as a *resolution parameter*: at short times, information only diffuses to neighboring nodes, revealing the local structure; at longer times, diffusion reaches nodes further apart and reveals the global network structure. A recent proposal, the Laplacian renormalization group (LRG) scheme [10], leverages this observation to produce coarser descriptions of a network's structure by identifying groups of nodes that are strongly linked by diffusion at a given scale. Moreover, one can adopt this same approach to define an *informational* notion of *scale-invariance*, based on the properties of the diffusion process via the Laplacian spectrum [14], and thus different from the canonical concept of scale-freeness, which instead depends directly on the degree distribution. Both of these results hinge on the formalism of network density matrices [15] to describe the complete behavior of information diffusion at a given scale.

Networks, however, are only part of the story. Great attention has recently been devoted to the study of higher-order networks: networks that encode multi-node interactions, going beyond the pairwise interactions of traditional networks [16–18]. Higher-order interactions are present in many natural systems and drastically impact most dynamical processes, such as random walks [19–21], diffusion [22, 23], spreading [24–30], coordination [31–37], and synchronization [38–43]. A higher-order interaction between $k+1$ nodes is typically called a hyperedge of order k. Systems with such interactions are formalized using simplicial complexes or hypergraphs, with

* tommaso.gili@imtlucca.it
† giovanni.petri@nulondon.ac.uk

the former being more structured, as the presence of interaction also requires the presence of interactions between all of its node subsets.

However, little work exists on RG approaches to higher-order networks despite their importance. A direct generalization of the Laplacian RG approach [10], based on the multiorder Laplacian [39], was recently proposed [44]. This proposal is, however, *node-centric*: it only considers the diffusion of information from node to node. A parallel research line [14, 45, 46] focused on specific families of simplicial complexes and used renormalization group techniques to compute some notable statistical properties.

Here, we propose a general renormalization group scheme for arbitrary higher-order networks. Our approach uses a higher-order notion of diffusion that we formally define by introducing the *cross-order Laplacian*. In this new diffusion process, information can flow through hyperedges of any order k via hyperedges of any other order m. This proposal provides a natural generalization of previous ones [44, 47], which are restricted to node-node diffusion. By studying the properties of this diffusion via the cross-order Laplacians, our approach allows us to probe the existence of characteristic scales, or—crucially—their absence (*scale-invariance*) in higher-order networks at each order. In particular, we first define the appropriate Laplacian matrices to describe generic higher-order diffusion. We then leverage them to define (i) a higher-order notion of informational scale-invariance through the von Neumann entropy and entropic susceptibility, and (ii) an explicit RG scheme informed by a chosen higher-order diffusion process. Using these tools, we extract a cross-order scale signature in higher-order networks obtained from synthetic models and real-world data and show that in most cases, scale-invariance is found only under the lens of specific orders, suggesting the existence of underlying order-specific processes.

II. HIGHER-ORDER NETWORKS AND THEIR STRUCTURE

a. ***Higher-order networks.*** Let Δ be a higher-order network (also named hypergraph) on a finite set of vertices V, i.e., a family of subsets of V. Any element of Δ is called *hyperedge* and its *order* is defined as its cardinality reduced by one. A vertex will thus be a 0-hyperedge, an edge a 1-hyperedge, a triangle a 2-hyperedge, and so on. If $\eta \in \Delta$ is a subset of a k-hyperedge σ, then we say that η is a *face* of σ. We write Δ_k to denote the set of k-hyperedges of Δ and call n_k its cardinality. The object Δ is a simplicial complex when closed under inclusion, that is, $\sigma \in \Delta,\ \eta \subseteq \sigma \implies \eta \in \Delta$. In that case, we call the hyperedges *simplices*.

Having fixed an order $k \in \mathbb{N}$, we want to find whether the hypergraph possesses characteristic scales or is scale-invariant from the perspective of k-hyperedges. In Ref. [10], it is argued that a diffusion process can be seen as a telescopic scanner of a (pairwise) network capable of extracting multiscale information about its structure. This fact suggests studying our hypergraph's k-th order properties through a diffusion process on the k-hyperedges.

b. ***Cross-order Laplacians.*** A diffusion-like process on the hyperedges of a higher-order network can be defined in multiple ways, each associated with a different Laplacian matrix. In the case of simplicial complexes, the k-th order combinatorial Hodge Laplacian, proposed by Eckmann [48], is the one most commonly considered [49, 50], mainly for its deep connections with topology [49]. A diffusion process on the k-hyperedges [22] can indeed be defined with the Hodge Laplacian, making it a natural candidate for our approach. However, it is possible to see that diffusion through the Hodge Laplacian does not correspond to a "standard" diffusion process as, for instance, the total amount of information flowing between hyperedges is not conserved (more details can be found in the Supplementary Information Section 1). Due to this lack of clear physical interpretability, it is difficult to directly employ it in the Laplacian renormalization framework proposed in Ref. [47]. Moreover, the Hodge Laplacian cannot be naively extended to the general hypergraph setting.

Thus, we define a new family of Laplacian matrices that can describe a plethora of higher-order relations while maintaining a form analogous to the canonical graph Laplacian. We do this by taking inspiration from the general theory of combinatorial complexes [51] and the hypergraph Laplacian [52].

Formally, we fix a number $k \in \mathbb{N}$, which we call *diffusion order*, to describe a diffusion process on the k-hyperedges of the hypergraph. We then need to decide how the process occurs, specifically, how hyperedges are "connected" so that information can flow between them. The most natural approach is to extract this information from the structure of Δ by employing a notion of *adjacency* among hyperedges. In general, two k-hyperedges σ and η can be adjacent in two ways:

- σ, η are *m-adjacent from above*, when there is an m-hyperedge ξ, with $m > k$, containing both of them $\xi \supseteq \sigma \cup \eta$;
- σ, η are *m-adjacent from below*, when they share a common m-face $\xi \subseteq \sigma \cap \eta$.

Combining these two definitions, we can define the *adjacency number* $a_{(k,m)}$ of a pair of *distinct* k-hyperedges $\sigma \neq \eta$ as

$$a_{(k,m)}(\sigma,\eta) = \begin{cases} |\{\lambda \in \Delta_m : \lambda \subseteq \sigma \cap \eta\}| \text{ if } m < k \\ |\{\lambda \in \Delta_m : \sigma \cup \eta \subseteq \lambda\}| \text{ if } m > k \\ 0 \text{ if } m = k \end{cases} \tag{1}$$

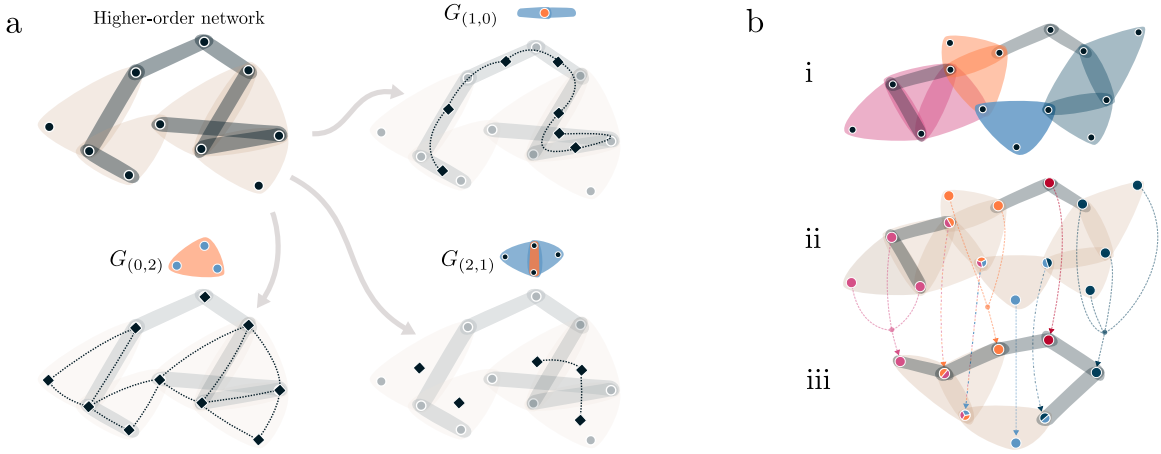

FIG. 1. **Cross-order Laplacian renormalization scheme: partition and coarse-grain. a.** Schematic representation of how the adjacency graphs are built for a hypergraph. Notice how there can be cases like $G_{(0,2)}$, where the edges are weighted (here represented with multi-edges) because two vertices can be connected by more than one triangle. **b.** Pictorial representation of our higher-order coarse-graining scheme with $k = 2$: (i) a partition of the 2-hyperedges of a hypergraph, here represented in color, is obtained from $\boldsymbol{\rho}(\tau)$, (ii) each vertex inherits a signature containing the labels of all 2-hyperedges it belongs to, (iii) vertices with the same signature are glued together and hyperedges are induced from the starting hypergraph.

so that we may consider η and σ to be m-adjacent $\sigma \overset{m}{\sim} \eta$ when $a_{(k,m)}(\sigma,\eta) > 0$. Intuitively, $a_{(k,m)}(\sigma,\eta)$ counts the number of m-hyperedges connecting σ and η, allowing us to differentiate between different "strengths" of adjacency. For example, $a_{(2,0)} = 1$ when two triangles share a single vertex and $a_{(2,0)} = 2$ when they share an edge.

The (k,m)-adjacency relations can be formalized into different *adjacency matrices*, analogous to those defined in Refs. [51, 53]. If we index the k-hyperedges as $\sigma_1,\ldots,\sigma_{n_k}$, we can define the (k,m)-adjacency matrix with *diffusion order* k and *interaction order* m as the square $n_k \times n_k$ matrix with elements

$$(\mathbf{A}_{(k,m)})_{ij} = a_{(k,m)}(\sigma_i,\sigma_j). \tag{2}$$

The adjacency matrix $\mathbf{A}_{(1,2)}$, for example, describes how edges (1-hyperedges) are connected through triangles (2-hyperedges), while $\mathbf{A}_{(3,0)}$ tells us how tetrahedra (3-hyperedges) are attached to one another through vertices (0-hyperedges). Notice how (k,m)-adjacency is equivalent to a weighted version of the q-nearness described in [54–56] when Δ is a simplicial complex and we restrict our attention only to k-simplices.

The matrix $\mathbf{A}_{(k,m)}$ can be seen as the adjacency matrix of a (weighted) graph, which we call *adjacency graph* $G_{(k,m)}$ (see Figure 1a), whose nodes are the k-hyperedges and the edges are their adjacency relations given by m-hyperedges. In particular, $G_{(0,1)}$ is the graph underlying the higher-order network (i.e. $\Delta_0 \cup \Delta_1$) and $G_{(1,0)}$ corresponds to its line graph [57]. To each weighted adjacency matrix, we can associate its (weighted) Laplacian, which we name *cross-order Laplacian*, through the usual formula

$$\mathbf{L}^{\times}_{(k,m)} = \mathrm{diag}(\deg_{(k,m)}) - \mathbf{A}_{(k,m)}, \tag{3}$$

where $\deg_{(k,m)}$ is the vector containing the higher-order (k,m)-degrees, defined as the row-sums of $\mathbf{A}_{(k,m)}$. More specifically,

$$\deg_{(k,m)}(\sigma) = \sum_{\eta\in\Delta_k} a_{(k,m)}(\sigma,\eta). \tag{4}$$

The matrix $\mathbf{L}^{\times}_{(k,m)}$, being a weighted graph Laplacian, is symmetric, positive semidefinite and has an eigenvalue 0 with multiplicity given by the number of connected components of its underlying graph $G_{(k,m)}$. We note that our definition includes two existing families of Laplacians as particular cases: (i) the "vertex-centric" higher-order Laplacians describing how vertices can exchange information through hyperedges (e.g., the generalized Laplacians of Ref. [39]), as cross-order Laplacians of the form $\mathbf{L}^{\times}_{(0,m)}$ and (ii) Hodge-like Laplacians where diffusion happens between hyperedges through hyperedges of a directly adjacent order, as $\mathbf{L}^{\times}_{(k,k\pm1)}$. Cross-order Laplacians generalize these two notions, allowing the description of diffusion processes which, as with Hodge Laplacians, can happen on hyperedges of any order and, as with generalized Laplacians, can "jump" orders, connecting them with hyperedges of any other order.

III. STATISTICAL PHYSICS OF HIGHER-ORDER DIFFUSION PROCESSES

a. ***Cross-order diffusion.*** From now on, we will consider the high-order diffusion process on k-hyperedges through m-hyperedges on a given Δ, and thus omit the (k,m) notation. Such a process can be easily written as the linear, first-order ODE

$$\dot{x}(\tau) = -\mathbf{L}^{\times}\, x(\tau), \tag{5}$$

where $x(\tau) \in \mathbb{R}^{n_k}$ is a real scalar function on the k-hyperedges. Equation (5) can be solved with the time propagation operator (also called *heat kernel*) at time $\tau > 0$,

$$\boldsymbol{\rho}(\tau) = e^{-\tau \mathbf{L}^{\times}}, \tag{6}$$

so that

$$x(\tau) = \boldsymbol{\rho}(\tau) x(0). \tag{7}$$

Due to linearity, the j-th column $\boldsymbol{\rho}(\tau)_{,j}$ of $\boldsymbol{\rho}(\tau)$ describes the distribution of information over the k-hyperedges at time τ, when the total information is concentrated in a single k-hyperedge σ_j at time $\tau = 0$.

It turns out that we can derive aggregate measures from the heat kernel, both for the standard Laplacian on networks [15] and for Hodge Laplacians on simplicial complexes [58]. These can help us probe the characteristic scales of the structure under consideration. First, $\boldsymbol{\rho}$ is normalized to a *density matrix* [15]

$$\hat{\boldsymbol{\rho}}(\tau) = \frac{e^{-\tau \mathbf{L}^{\times}}}{Z(\tau)} \tag{8}$$

where $Z(\tau) = \mathrm{Tr}\left(e^{-\tau \mathbf{L}^{\times}}\right)$ is called the *return probability*, describing how much of the information has remained "trapped" and did not diffuse at time τ [59]. With this density matrix, we can compute the von Neumann *entropy* associated to the diffusion process as

$$S(\tau) = -\,\mathrm{Tr}(\hat{\boldsymbol{\rho}}(\tau) \log \hat{\boldsymbol{\rho}}(\tau)) \tag{9}$$

and its *entropic susceptibility* [47]

$$C(\tau) = -\frac{\mathrm{d}S}{\mathrm{d}\log\tau}. \tag{10}$$

The density matrix formalism and the associated von Neumann entropy allow us to describe a network's transport properties by simultaneously considering all possible diffusion trajectories. Thus, we can directly apply it to the cross-order Laplacian to investigate our hypergraph's k-th order properties at different scales. This approach, in the $k = 0$ case, has been fruitfully explored in the literature to extract key information about the network's structural organization [15, 60–62].

In particular, the maxima and minima of the entropic susceptibility, associated with times of fast deceleration and acceleration of the diffusion process, were shown, in the case of networks, to correspond to the presence of characteristic scales [10, 47]. Most importantly, when $C(\tau)$ is constant over a time range $I = [\tau_{\min}, \tau_{\max}]$, we say that Δ exhibits informational *scale-invariance* [10] in I. Explicitly writing orders (k, m) again: it is known that the (network, $k = 0$ and $m = 1$) entropic susceptibility $C_{(0,1)}(\tau)$ shows a large plateau in the case of grid graphs, Barabási-Albert networks, and random trees, all of which are examples of self-similar structures. Interestingly, entropic susceptibility and the associated notion of scale-invariance can be related to the concept of spectral dimension [63–65], which intuitively measures the dimensionality "perceived" by a diffusion process taking place on a manifold or, in our case, a graph (see Supplementary Information Section 2).

b. ***Measuring scale-invariance.*** Following the considerations above, we want to quantify whether a higher-order network exhibits scale-invariance at order k via order m. To do so, we define the *scale-invariance parameter* (SIP) $P_{(k,m)}(\epsilon)$, as the logarithmic lifespan of the longest connected plateau of $C_{(k,m)}(\tau)$ w.r.t. a given tolerance $\epsilon > 0$ on its "flatness". In detail, given a value $y > 0$, we define the set

$$E_{(k,m)}(y;\epsilon) = \left\{\tau > 0 : |\log C_{(k,m)}(\tau) - y| < \epsilon\right\} \tag{11}$$

which, being the inverse image of the open interval $(y - \epsilon, y + \epsilon)$, will be given by a countable, disjoint union of open intervals (a_i, b_i)

$$E_{(k,m)}(y;\epsilon) = \coprod_i (a_i, b_i). \tag{12}$$

We then define the *scale-invariance parameter* as

$$P_{(k,m)}(\epsilon) = \max_{y>0} \max_i (\log b_i - \log a_i). \tag{13}$$

If $P_{(k,m)}(\epsilon)$ is large, we can say that the hypergraph is scale-invariant at order k via order m, while if it is close to 0 then $C_{(k,m)}(\tau)$ is peaked, thus signaling the presence of characteristic scales. Unless stated otherwise, in the numerical experiments we fix $\epsilon = 0.2$ and omit ϵ in the notation: $P_{(k,m)}(\epsilon) \equiv P_{(k,m)}$.

IV. CROSS-ORDER LAPLACIAN RENORMALIZATION SCHEME

Let us assume that a hypergraph Δ has been recognized as scale-invariant with respect to $C_{(k,m)}(\tau)$ (10). We would now like an algorithmic method to reduce it to a smaller, equivalent one that can still be recognized as scale-invariant. Similarly to the case of networks, we can devise a *higher-order* Laplacian renormalization scheme based on the relationship between k-hyperedges through m-hyperedges.

It consists of the following steps:

1. first, choose a diffusion order k and an interaction order m, resulting in the cross-order Laplacian matrix $\mathbf{L}^{\times}_{(k,m)}$;
2. choose a diffusion time $\tau^* > 0$ corresponding to the scale at which to "zoom out";
3. compute a partition of the k-hyperedges from the values of $\boldsymbol{\rho}_{(k,m)}(\tau^*)$ such that hyperedges in the same set are strongly linked by the diffusion process at time τ;

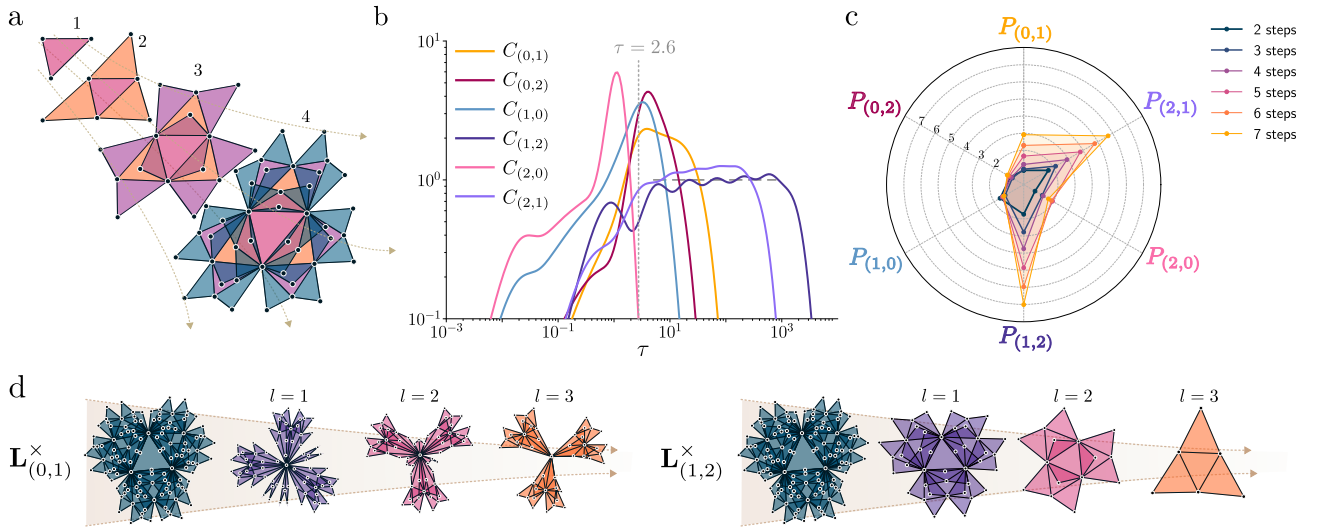

FIG. 2. **Scale-invariance and renormalization in pseudofractal simplicial complexes. a.** Graphical depiction of the first three steps of constructing the pseudofractal simplicial complex of dimension 2. **b.** Entropic susceptibility curves of all the non-zero cross-order Laplacian matrices, computed for the 2-dimensional pseudofractal simplicial complex built with six steps (1095 vertices). **c.** Values of the scale-invariance parameters as the number of steps with which the pseudofractal simplicial complex is built increases. **d.** On the left, the first three steps of the $\mathbf{L}^{\times}_{(0,1)}$ renormalization scheme with $\tau = 0.2$. On the right, the first three steps of the $\mathbf{L}^{\times}_{(1,2)}$ renormalization with $\tau = 2.6$.

4. coarse-grain Δ by merging its vertices according to the partition, to obtain a new, smaller Δ'.

Full details about steps 3 and 4 can be found in Section 3 of the Supplementary Information. The process is then repeated, resulting in a sequence of hypergraphs with a decreasing or constant number of vertices. We name such a sequence *renormalization flow*

Intuitively, step 3 identifies groups of k-hyperedges, which can be seen as generalizations of the spin blocks of Kadanoff's renormalization scheme [12]. In general, however, they will not be identical blocks of nodes, but their shape will reflect the structure of Δ (see Figure 1b-i).

Once the partition of the k-hyperedges is obtained, we perform a coarse-graining step to aggregate the hyperedges belonging to the same block. Notice that we cannot directly coarse grain the adjacency graph $G_{(k,m)}$ as we would loose all the relational information between hyperedges of Δ which is not contained in $G_{(k,m)}$. We therefore take a different approach and use the information in the partition of k-hyperedges to coarse grain the vertices of Δ. First, we move the problem to the domain of the vertices by letting each one inherit the labels of all the k-hyperedges to which they belong (see Figure 1b-ii). Afterward, we glue together the vertices that inherited the same label set and induce hyperedges from the starting Δ (see Figure 1b-iii). A visualization of step 3 can be found in Sections 4 of the Supplementary Information. A code implementation of the method can be found in [66].

Note that the choice of the time τ^* heavily influences the renormalization process. If τ is too low, no vertices will be merged, whereas if τ grows large enough, every family of connected k-hyperedges in $G_{(k,m)}$ will get the same label.

V. HIGHER-ORDER SCALE-INVARIANCE

We now show explicit examples of applications of the cross-order renormalization scheme. We first focus on synthetic models of simplicial complexes to confirm that, in controlled situations, the cross-order renormalization group recovers exactly the scale-invariant structure and order of the underlying system. After this confirmation, we extract the *cross-order scale signature* from some real-world datasets.

a. ***Pseudofractal simplicial complexes.*** As we mentioned above, there are situations in which the organization of a system is most evident when looking at it from a high-order point of view. A pretty interesting example is given by the family of scale-free pseudofractal simplicial complexes [67]. Simplicial complexes in this family are built starting with a single k-simplex and by iteratively attaching a k-simplex to each $(k-1)$-simplex already present in the complex (Figure 2a).

We expect the evident hierarchical nature of these objects to be visible in the entropic susceptibility curve of one or more diffusion processes defined on it.

Figure 2b illustrates how the different entropic susceptibilities $C_{(k,m)}$ display different non-trivial behaviors. Most importantly, the curves associated with $\mathbf{L}^{\times}_{(1,2)}$ and $\mathbf{L}^{\times}_{(2,1)}$ show oscillating plateaux that span multiple orders of magnitude in τ. These correspond to the different, well-separated scales resulting from the iterative construction process, which can be interpreted as an approxi-

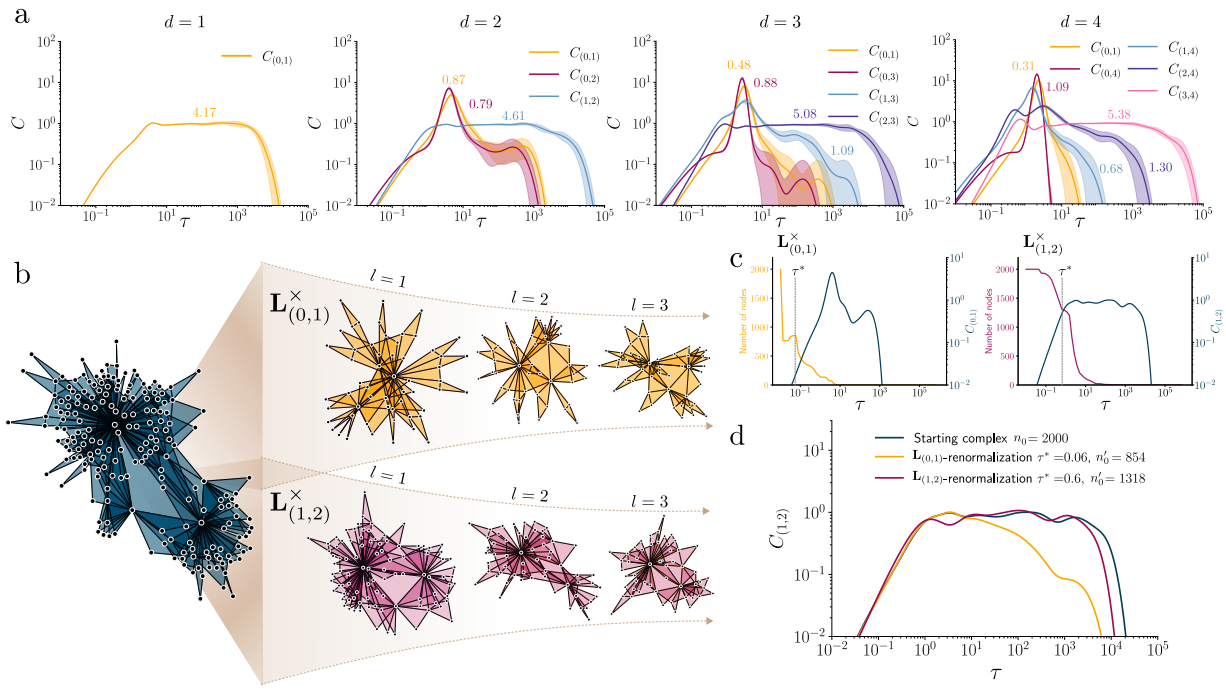

FIG. 3. **Higher-order Laplacian renormalization scheme applied to heterogeneous NGF simplicial complexes a.** Entropic susceptibility curves, together with their 95% CIs over 10 repetitions, of the NGF simplicial complexes of flavor $s = 1$, dimensions $d \in \{1, 2, 3, 4\}$, $\beta = 5$ and 3000 vertices. The numbers over the curve show the scale-invariance parameters. **b.** A small 2-dimensional NGF simplicial complex is renormalized using $\mathbf{L}^{\times}_{(0,1)}$ (top) and $\mathbf{L}^{\times}_{(1,2)}$ (bottom). **c.** In the left panel, $C_{(0,1)}$ together with the number of vertices of the complex after one step of $\mathbf{L}^{\times}_{(0,1)}$-renormalization as a function of τ. In the right panel, $C_{(1,2)}$ together with the number of vertices of the complex after one step of $\mathbf{L}^{\times}_{(1,2)}$-renormalization as a function of τ. **d.** Evolution of the entropic susceptibility $C_{(1,2)}$ over the first step of the two types of renormalization considered. The starting NGF simplicial complex has 2000 vertices and is reduced to 854 vertices by the flow of $\mathbf{L}^{\times}_{(0,1)}$ ($\tau^* = 0.06$) and to 1318 by $\mathbf{L}^{\times}_{(1,2)}$ ($\tau^* = 0.6$).

mate form of higher-order informational scale-invariance. Surprisingly, even if the network underlying the simplicial complex is scale-free [68], the curve of $\mathbf{L}^{\times}_{(0,1)}$ (the canonical graph Laplacian) is peaked, meaning that the self-similarity of the structure is not visible at the vertex-edge level.

These results are consistently observed when we increase the number of steps with which the simplicial complex is built. As expected for a growing fractal structure, we find that almost all the scale-invariance parameters $P_{(k,m)}$ (Equation (13)) for all possible cross-order Laplacians increase with the number of steps in the construction of the pseudofractal complex (Figure 2c). Crucially, we also find that the highest values of P are consistently obtained for $P_{(1,2)}$ and $P_{(2,1)}$, which correspond to the cross-order Laplacian naturally associated with the complex's growth (addition of triangles along edges). Consistently, we also find that the values of $P_{(2,0)}$, $P_{(1,0)}$ and $P_{(0,2)}$ have low values which remain approximately constant, while $P_{(0,1)}$ slowly increases, indicating the presence of a growing plateau (which was invisible in Figure 2b, where we show results for the realization with six steps). Finally, we observe that the observed behavior is also consistent with previous results [14] on the spectrum of $\mathbf{L}^{\times}_{(0,1)}$ of pseudofractal simplicial complexes, which has power-law behavior for small eigenvalues when the number of vertices is large, i.e. a plateau in the entropic susceptibility (see Methods of [47]). However, the scale-invariant behavior is much more easily and quickly detectable when considering the entropic susceptibility of Laplacians associated with the complex's intrinsic growth process (between edges and triangles), even when the number of vertices is still small.

We confirm this observation by applying the renormalization scheme (Section IV) to a pseudofractal simplicial complex of dimension 2 and visualizing its evolution at each step. We find that the renormalization based on $\mathbf{L}^{\times}_{(1,2)}$ preserves the structure of the simplicial complex, perfectly reversing its iterative construction process (Figure 2d, right). In contrast, a renormalization flow based on $\mathbf{L}^{\times}_{(0,1)}$ destroys the pseudofractal structure and rapidly collapses the central vertices, resulting in a "star-shaped" simplicial complex with no apparent relation with the original one (Figure 2d, left). In the Supplementary Information (Section 5), we thoroughly explore the renormalization of the pseudofractal simplicial complex under different Laplacians and values of τ.

b. ***Network Geometry with flavor.*** Up to now, we found the presence of higher-order scale-invariance only for pseudofractal simplicial complexes. While being a relevant consistency check, this result is not particularly striking due to the homogeneity of their structure and their evident hierarchical nature. We show here that analogous results can be found in the case of heterogeneous simplicial complexes, similar to what was previously observed in random scale-free networks and random trees [10]. To this end, a natural candidate is the family of simplicial complexes given by the *Network Geometry with flavor* (NGF) model [68, 69]. Indeed, the NGF model in dimension d can generate both hyperbolic manifolds and *scale-free* growing simplicial complexes by progressively attaching d-simplices to $(d-1)$-simplices in a stochastic manner (see Section 6A of the Supplementary Information for details on the model).

This attachment process is governed by three possible *flavors* $s \in \{-1, 0, 1\}$, each resulting in distinct structural properties. Additionally, the model incorporates a parameter β acting as an inverse temperature, influencing the randomness in the process.

Notably, certain higher-order degrees, depending on flavor and dimension, exhibit power-law distributions, indicating a scale-free structural organization. Specifically, $\deg_{(k,d)}$ (Equation (4)) follows a power-law when $k \leq d-3$ for $s=-1$, when $k \leq d-2$ for $s=0$, and for all $k \leq d-1$ for $s=1$ (see Section 6B of the Supplementary Information for details).

Given their scale-free nature *at all levels*, our focus is on NGF simplicial complexes with flavor $s=1$. We examine their entropic susceptibility curves to discern scale-invariance. To manage computational complexity, we calculate curves only for relations predicted to be scale-free, alongside the vertex-edge $(0,1)$ relation for comparison.

The situation differs from the pseudofractal case, as shown in Figure 3a. Although the simplicial complex follows a similar construction process, its randomness does not allow for a clear separation of the hierarchical scales. Unlike the oscillating plateau observed in Figure 2a, the curves of $C_{(d,d-1)}$ exhibit a small peak, which grows with higher dimensions, indicating the presence of a distinct microscopic scale. This is because the $(d-1,d)$-adjacency graph (of both d-dimensional NGFs and pseudofractals) is composed by $(d+1)$-cliques arranged in a tree-like structure, as d-simplices have $d+1$ $(d-1)$-faces. The first peak thus corresponds to the microscale associated with these cliques (see Section 4 of the Supplementary Information for a visualization of this fact). Afterward, a scale-invariant regime ensues, illustrated by the near-perfect plateau of $C_{(d-1,d)}$. However, despite being associated with power-law degree distributions, other curves lack plateaus, showcasing non-trivial structural organization at specific scales.

In Figure 3b, we visually demonstrate that a renormalization based on a higher-order relation ($\mathbf{L}^{\times}_{(1,2)}$) better preserves the structure of a 2-dimensional NGF simplicial complex than the standard vertex-edge Laplacian renormalization. Notably, Figure 3c shows that the $\mathbf{L}^{\times}_{(0,1)}$-renormalization drastically reduces the number of vertices, collapsing them into a single super-vertex at the first peak of $C_{(0,1)}$. Conversely, $\mathbf{L}^{\times}_{(1,2)}$-renormalization compresses the simplicial complex more gradually, revealing a clear transition point just before the first peak. This distinction is further emphasized by tracking the evolution of the entropic susceptibility after one step of the renormalization flow in Figure 3d, where the plateau of $C_{(1,2)}$ is preserved by $\mathbf{L}^{\times}_{(1,2)}$ but destroyed by $\mathbf{L}^{\times}_{(0,1)}$.

c. ***Higher-order scale-invariance in real data.*** The framework we established can be fruitfully employed to probe the structure of real-world networks. We can leverage scale-invariance parameters to provide effective higher-order signatures describing their hierarchical nature in a multifaceted way.

We take real higher-order network datasets (see Section 7D of the Supplementary Information for details) and consider their interactions up to the 4th order, i.e hyperedges of maximum 5 nodes. For each hypergraph, we compute all the $C_{(k,m)}$ with associated $P_{(k,m)}$ and build the null model by doing the same on their adjacency graphs randomized with a degree-preserving configuration model.

As shown in Figure 4a, we obtain an effective signature of the hypergraph encoding the amount of scale-invariance in each of its higher-order relations. We find different behaviors, like the "flower-like" shape for the *High school contacts* [70] hypergraph, the "fan-like" shape —with higher values of scale-invariance in some specific orders— of the *Senate bills* [71], *House bills* [71], the *Diseasome* [72] and the *NDC Substances* [73] hypergraphs. Others, like the *Enron Email* [73, 74], show a more uniform distribution of the SIPs values, without a clear dominance of orders. Note that when the k-hyperedges are too sparse information cannot flow through m-hyperedges, resulting in zero-valued scale-invariance parameters for those specific orders.

Interestingly, even when we do not have access to a genuine higher-order network, it is still possible to obtain meaningful information through our approach. To show this, we consider real pairwise network datasets (see Section 7C of the Supplementary Information for details) and compute their associated *clique complexes* up to 2nd order, i.e. build simplicial complexes by considering their cliques of three nodes as 2-hyperedges. In this way, notice, we artificially lift the pairwise network to a higher-order one and, as a consequence, any resulting higher-order properties should be interpreted as describing the structural organization of cliques of nodes and not of "genuine" higher-order interactions.

In Figure 4c are shown the scale-invariance parameters associated to the specific heats for each clique complex, together with the null model. We find a zoo of different behaviors, like the *E-road* network [75] and the network science *Co-authorship* network [76], which are strongly scale-invariant in the standard sense, i.e. in the relation between vertices and edges (high values of $P_{(0,1)}$

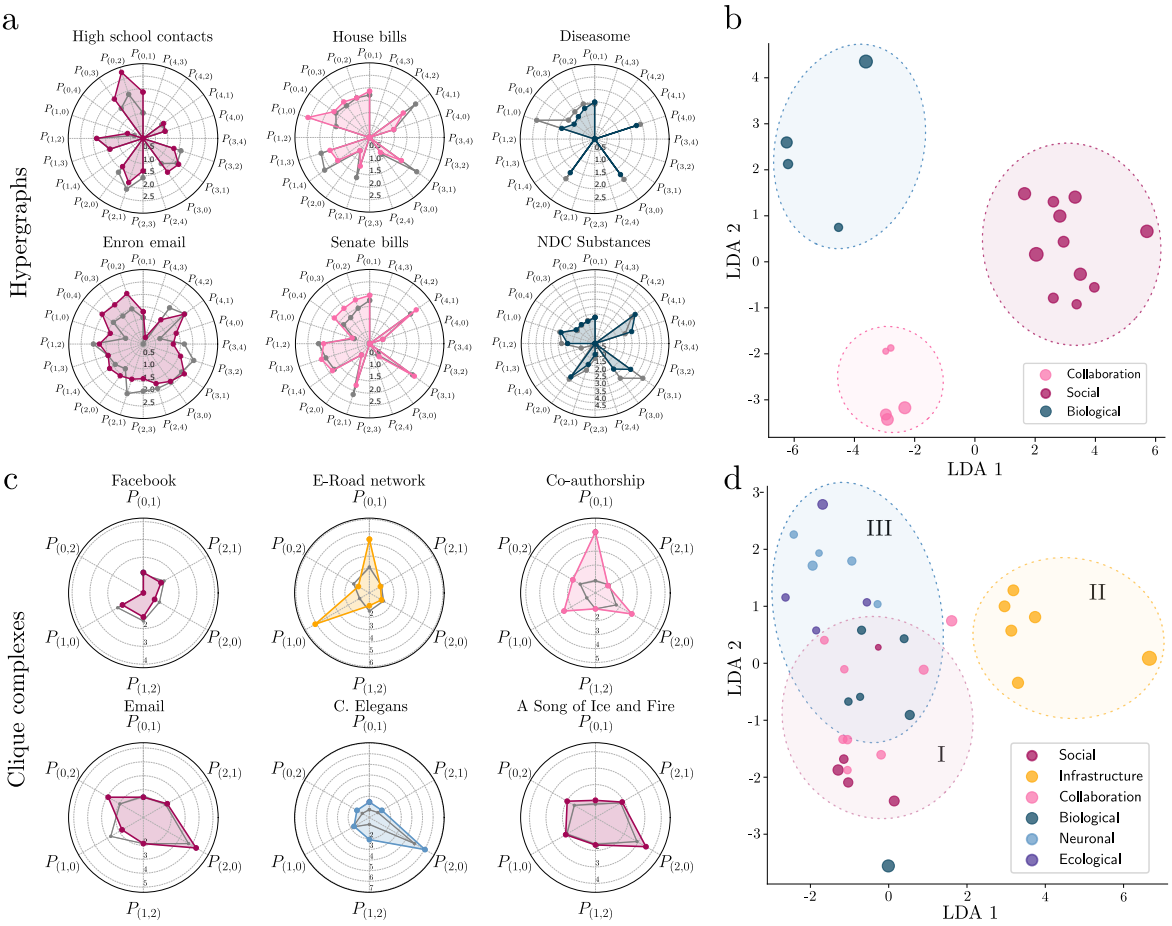

FIG. 4. **Higher-order scale-invariance in real data.** **a.** Scale-invariance parameters $P_{(k,m)}$ computed for $k \neq m \in \{0, 1, 2, 3, 4\}$, for six different hypergraphs–up to 4-hyperedges—obtained from real-world datasets. The gray line and the associated region represent the mean and 95% CI for the values of $P_{(k,m)}$ associated with the (k, m)-adjacency graphs randomized with the configuration model over 10 repetitions. **b.** LDA projection of a set of 20 real hypergraphs of different types. The color of each point represents the associated hypergraph's type, while the size is proportional to the value of the network's highest SIP. The colored regions highlight clusters of nearby points that belong to the same class of hypergraphs. **c.** Scale-invariance parameters $P_{(k,m)}$ computed for $k \neq m \in \{0, 1, 2\}$, for six different second-order clique complexes obtained from real-world network datasets. The gray line and the associated region represent the mean and 95% CI for the values of $P_{(k,m)}$ associated with the (k, m)-adjacency graphs randomized with the configuration model over 10 repetitions. **d.** LDA projection of a set of 34 real clique complexes of different types. The color of each point represents the associated network's type, while the size is proportional to the value of the network's highest SIP. The colored regions highlight clusters of nearby points that belong to the same class of networks: social origin (I), infrastructural (II), biological origin (III).

and $P_{(1,0)}$). Others, as the *Facebook* network [77] show negligible values of scale-invariance in each order. Interestingly, despite the inherent pairwise nature of the networks, we find a marked presence of high-order scale-invariance in some of them, like the University Rovira i Virgili *Email* network [78], the *C.Elegans* metabolic network [79–81] and the fictional *A Song of Ice and Fire* social network [82]. We stress again that this higher-order description is obtained without any assumptions on the presence of genuine higher-order interactions in these systems, but rather it is based only on nodes, edges, and the cliques they form. This observation validates our claim that some aspects of the hierarchical scale-invariant structure of a higher-order network may be hidden from the node-centric point of view, but can be evident when looking at how simplices (in this case, cliques) are related to one another. Additionally, the null models' scale-invariance parameters, displayed in gray, show high variability w.r.t. the real SIPs, attaining both higher and lower values. This suggests that the SIP is not merely a function of the high-order degree distribution but is instead dependent on the networks' specific structural organizations.

Finally, we use scale-invariance parameters as coordinates to embed them in $\mathbb{R}^2$ (for details, see Supplementary Information Section 7C and 7D). In Figure 4b and Figure 4d, the linear projection–Linear Discriminant Analysis (LDA)[83]–of the resulting point cloud are showing heterogeneity in the amount of higher-order scale-invariance present in each dataset. Surprisingly, despite

the coarseness of this measure, we see that higher-order networks of similar nature tend to be closer to each other, as shown by the colored regions, highlighting the presence of specific higher-order signatures associated with each type of data.

VI. DISCUSSION

By leveraging the Laplacian renormalization group, we developed a method to investigate the structure of higher-order networks via a robust renormalization procedure based on the connectivity structure of higher-order relations, as encoded in the proposed cross-order Laplacian. We showed that the renormalization scheme can revert exactly the construction of higher-order structures that are self-similar by construction (the pseudofractal) and detect the correct order of the dominating growth mechanism in scale-invariant complexes with induced heterogeneous lower order structures (the NGF model). Armed with these results on controlled synthetic systems, we employed the entropic susceptibilities obtained in our scheme to build *scale-invariance profiles* for a set of real-world systems, revealing both different unexpected kinds of characteristic scales and scale-invariance at various orders and commonality of such profiles across systems belonging to the same domain.

From a technical point of view, the cross-order Laplacians bear both similarities and differences with previously defined higher-order Laplacians, notably the Hodge (or combinatorial) Laplacian [22] $\boldsymbol{\mathcal{L}}_k$, and the multi-order Laplacian $\mathbf{L}^{(\mathrm{mul})}$ [39]. In fact, both the k-order Hodge Laplacian and cross-order Laplacians $\mathbf{L}^{\times}_{(k,m)}$ (for any m) are defined on the hyperedges of order k, and thus are $n_k \times n_k$ matrices, where n_k is the number of k-hyperedges. However, the Hodge Laplacian is limited to adjacencies defined by boundary and coboundary relations, and thus, with $m = k \pm 1$, while the cross-order can capture arbitrary adjacencies through any m-hyperedges. Conversely, the multi-order Laplacian $\mathbf{L}^{(\mathrm{mul})}$ is defined as the weighted sum of Laplacians defined on nodes and adjacent via hyperedges of any m order. It can be rewritten in our notation as $\mathbf{L}^{(\mathrm{mul})} = \sum_{m=1}^{M} \omega_m \mathbf{L}^{\times}_{(0,m)}$, where the ω_m are an arbitrary weighting scheme and M the maximal order considered. Similarly, we note that, while in this work we focused on renormalization based on specific pairs (k, m), important future work should understand the effect of extending the cross-order renormalization scheme to a *multicross*-order scheme, which could be achieved by considering combinations of hyperedges of dimension higher or lower than a certain threshold on the interaction order (e.g., $\mathbf{L}_k^{(\mathrm{mul}\times)} = \sum_{m\in\{\mathbf{m}\}} \omega_m \mathbf{L}^{\times}_{(k,m)}$ over a set $\{\mathbf{m}\}$). Also, while this work focused on simplicial complexes for convenience and clearness of exposition, any combinatorial structure [51] built with a set of ranked sets is amenable to our scheme.

Finally, our results provide a new lens to address questions on the origin of different higher-order invariant structures in various domains [84, 85], on their effects on dynamical processes taking place on them [25, 28] and their reducibility [86], as well as to the limits they might pose on the predictability and reconstruction of complex systems [87].

VII. ACKNOWLEDGEMENTS

M.N. acknowledges the project PNRR-NGEU, which has received funding from the MUR – DM 352/2022. T.G. thankfully acknowledges financial support by the European Union - NextGenerationEU - National Recovery and Resilience Plan (Piano Nazionale di Ripresa e Resilienza, PNRR), project 'SoBigData.it - Strengthening the Italian RI for Social Mining and Big Data Analytics' - Grant IR0000013 (n. 3264, 28/12/2021). We also thank P. Villegas, A. Gabrielli, C. Agostinelli, M. Neri and T. Robiglio for extremely valuable suggestions on preliminary versions of the manuscript.

[1] M. E. Fisher, Reviews of Modern Physics **46**, 597 (1974).
[2] Y. Tu, Nature Physics **19**, 1536 (2023).
[3] C. Song, S. Havlin, and H. A. Makse, Nature **433**, 392 (2005).
[4] K.-I. Goh, G. Salvi, B. Kahng, and D. Kim, Physical Review Letters **96**, 018701 (2006).
[5] H. D. Rozenfeld, C. Song, and H. A. Makse, Physical Review Letters **104**, 025701 (2010).
[6] G. García-Pérez, M. Boguñá, and M. Á. Serrano, Nature Physics **14**, 583 (2018).
[7] E. Garuccio, M. Lalli, and D. Garlaschelli, arXiv preprint arXiv:2009.11024 (2020), arxiv:2009.11024.
[8] M. Zheng, G. García-Pérez, M. Boguñá, and M. Serrano, arXiv preprint arXiv:2307.00879 (2023), arxiv:2307.00879.
[9] M. d. C. Loures, A. A. Piovesana, and J. A. Brum, arXiv 10.48550/arxiv.2302.07093 (2023), 2302.07093.
[10] P. Villegas, T. Gili, G. Caldarelli, and A. Gabrielli, Nature Physics **19**, 445 (2023).
[11] M. Á. Serrano, D. Krioukov, and M. Boguná, Physical Review Letters **100**, 078701 (2008).
[12] L. P. Kadanoff, Physics Physique Fizika **2**, 263 (1966).
[13] M. Boguna, I. Bonamassa, M. De Domenico, S. Havlin, D. Krioukov, and M. Á. Serrano, Nature Reviews Physics **3**, 114 (2021).
[14] G. Bianconi and S. N. Dorogovstev, Journal of Statistical Mechanics: Theory and Experiment **2020**, 014005 (2020).
[15] M. De Domenico and J. Biamonte, Physical Review X **6**, 041062 (2016).

[16] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, Physics Reports **874**, 1 (2020).
[17] F. Battiston, E. Amico, A. Barrat, G. Bianconi, G. Ferraz de Arruda, B. Franceschiello, I. Iacopini, S. Kéfi, V. Latora, Y. Moreno, *et al.*, Nature Physics **17**, 1093 (2021).
[18] C. Bick, E. Gross, H. A. Harrington, and M. T. Schaub, SIAM Review **65**, 686 (2023).
[19] M. T. Schaub, A. R. Benson, P. Horn, G. Lippner, and A. Jadbabaie, SIAM Review **62**, 353 (2020).
[20] S. Mukherjee and J. Steenbergen, Random structures & algorithms **49**, 379 (2016).
[21] O. Parzanchevski and R. Rosenthal, Random Structures & Algorithms **50**, 225 (2017).
[22] A. Muhammad and M. Egerstedt, in *Proc. of 17th International Symposium on Mathematical Theory of Networks and Systems* (Citeseer, 2006) pp. 1024–1038.
[23] J. J. Torres and G. Bianconi, Journal of Physics: Complexity **1**, 015002 (2020).
[24] M. Lucas, I. Iacopini, T. Robiglio, A. Barrat, and G. Petri, Physical Review Research **5**, 013201 (2023).
[25] I. Iacopini, G. Petri, A. Barrat, and V. Latora, Nature Communications **10**, 2485 (2019).
[26] S. Chowdhary, A. Kumar, G. Cencetti, I. Iacopini, and F. Battiston, Journal of Physics: Complexity **2**, 035019 (2021).
[27] G. St-Onge, I. Iacopini, V. Latora, A. Barrat, G. Petri, A. Allard, and L. Hébert-Dufresne, Communications Physics **5**, 25 (2022).
[28] G. Ferraz de Arruda, G. Petri, P. M. Rodriguez, and Y. Moreno, Nature Communications **14**, 1375 (2023).
[29] Y. Chen, Y. R. Gel, M. V. Marathe, and H. V. Poor, Proceedings of the National Academy of Sciences **121**, e2313171120 (2024).
[30] I. Z. Kiss, I. Iacopini, P. L. Simon, and N. Georgiou, Journal of Complex Networks **11**, cnad044 (2023).
[31] U. Alvarez-Rodriguez, F. Battiston, G. F. de Arruda, Y. Moreno, M. Perc, and V. Latora, Nature Human Behaviour **5**, 586 (2021).
[32] I. Iacopini, M. Karsai, and A. Barrat, arXiv preprint arXiv:2306.09967 (2023), arxiv:2306.09967.
[33] M. Mancastroppa, I. Iacopini, G. Petri, and A. Barrat, Nature Communications **14**, 6223 (2023).
[34] Y. Shang, Proceedings of the Royal Society A **478**, 20210564 (2022).
[35] L. Neuhäuser, A. Mellor, and R. Lambiotte, Physical Review E **101**, 032310 (2020).
[36] L. Neuhäuser, R. Lambiotte, and M. T. Schaub, Physical Review E **104**, 064305 (2021).
[37] R. Sahasrabuddhe, L. Neuhäuser, and R. Lambiotte, Journal of Physics: Complexity **2**, 025006 (2021).
[38] P. S. Skardal and A. Arenas, Physical Review Letters **122**, 248301 (2019).
[39] M. Lucas, G. Cencetti, and F. Battiston, Physical Review Research **2**, 033410 (2020).
[40] Y. Zhang, M. Lucas, and F. Battiston, Nature Communications **14**, 1605 (2023).
[41] L. V. Gambuzza, F. Di Patti, L. Gallo, S. Lepri, M. Romance, R. Criado, M. Frasca, V. Latora, and S. Boccaletti, Nature Communications **12**, 1255 (2021).
[42] A. P. Millán, J. J. Torres, and G. Bianconi, Physical Review Letters **124**, 218301 (2020).
[43] M. Nurisso, A. Arnaudon, M. Lucas, R. L. Peach, P. Expert, F. Vaccarino, and G. Petri, Chaos **34**, doi.org/10.1063/5.0169388 (2024).
[44] A. Cheng, P. Sun, and Y. Tian, arXiv preprint arXiv:2305.01895 (2023), arxiv:2305.01895.
[45] M. Reitz and G. Bianconi, J. Phys. A Math. Theor. **53**, 295001 (2020).
[46] H. Sun, R. M. Ziff, and G. Bianconi, Physical Review E **102**, 012308 (2020).
[47] P. Villegas, A. Gabrielli, F. Santucci, G. Caldarelli, and T. Gili, Physical Review Research **4**, 033196 (2022).
[48] B. Eckmann, Commentarii Mathematici Helvetici **17**, 240 (1944).
[49] L.-H. Lim, Siam Review **62**, 685 (2020).
[50] S. Krishnagopal and G. Bianconi, Physical Review E **104**, 064303 (2021).
[51] M. Hajij, G. Zamzmi, T. Papamarkou, N. Miolane, A. Guzmán-Sáenz, K. N. Ramamurthy, T. Birdal, T. K. Dey, S. Mukherjee, S. N. Samaga, *et al.*, arXiv preprint 10.48550/arXiv.2206.00606 (2023).
[52] M. E. Aktas and E. Akbas, in *Complex Networks & Their Applications X: Volume 2, Proceedings of the Tenth International Conference on Complex Networks and Their Applications COMPLEX NETWORKS 2021 10* (Springer, 2022) pp. 277–288.
[53] E. Estrada and G. J. Ross, Journal of theoretical biology **438**, 46 (2018).
[54] R. H. Atkin, International journal of man-machine studies **4**, 139 (1972).
[55] R. Atkin, International Journal of Man-Machine Studies **8**, 483 (1976).
[56] R. H. Atkin, *Mathematical structure in human affairs* (Heinemann Educational London, 1974).
[57] H. Whitney, Hassler Whitney Collected Papers , 61 (1992).
[58] F. Baccini, F. Geraci, and G. Bianconi, Physical Review E **106**, 034319 (2022).
[59] A. Ghavasieh and M. De Domenico, Journal of Physics: Complexity **3**, 011001 (2022).
[60] A. Ghavasieh, C. Nicolini, and M. De Domenico, Phys. Rev. E **102**, 052304 (2020).
[61] A. Ghavasieh and M. De Domenico, Phys. Rev. E **107**, 044304 (2023).
[62] A. Ghavasieh and M. De Domenico, Nature Physics , 1 (2024).
[63] R. Rammal and G. Toulouse, Journal de Physique Lettres **44**, 13 (1983).
[64] J. Ambjørn, J. Jurkiewicz, and R. Loll, Physical Review Letters **95**, 171301 (2005).
[65] G. Calcagni, D. Oriti, and J. Thürigen, Classical and Quantum Gravity **31**, 135014 (2014).
[66] M. Nurisso, M. Morandini, M. Lucas, F. Vaccarino, T. Gili, and G. Petri, Code implementation of the higher-order laplacian renormalization method, `https://github.com/nplresearch/higher_order_LRG` (2024).
[67] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Physical Review E **65**, 066122 (2002).
[68] G. Bianconi and C. Rahmede, Physical Review E **93**, 032315 (2016).
[69] G. Bianconi and C. Rahmede, Scientific reports **7**, 41974 (2017).
[70] R. Mastrandrea, J. Fournet, and A. Barrat, PLoS One **10**, e0136497 (2015).
[71] N. W. Landry, M. Lucas, I. Iacopini, G. Petri, A. Schwarze, A. Patania, and L. Torres, Journal of Open Source Software **8**, 5162 (2023).

[72] K.-I. Goh, M. E. Cusick, D. Valle, B. Childs, M. Vidal, and A.-L. Barabási, Proceedings of the National Academy of Sciences of the United States of America **104**, 8685 (2007).

[73] A. R. Benson, R. Abebe, M. T. Schaub, A. Jadbabaie, and J. Kleinberg, Proc Natl Acad Sci U.S.A **115**, E11221 (2018).

[74] Enron email dataset (2024).

[75] L. Šubelj and M. Bajec, The European Physical Journal B **81**, 353 (2011).

[76] M. E. Newman, Physical Review E **74**, 036104 (2006).

[77] J. Leskovec and J. Mcauley, Advances in neural information processing systems **25** (2012).

[78] R. Guimera, L. Danon, A. Diaz-Guilera, F. Giralt, and A. Arenas, Physical Review E **68**, 065103 (2003).

[79] J. Duch and A. Arenas, Physical Review E **72**, 027104 (2005).

[80] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, Nature **407**, 651 (2000).

[81] R. Overbeek, N. Larsen, G. D. Pusch, M. D'Souza, E. S. Jr, N. Kyrpides, M. Fonstein, N. Maltsev, and E. Selkov, Nucleic acids research **28**, 123 (2000).

[82] A. Beveridge and M. Hunger, Asoiaf (2013).

[83] G. James, D. Witten, T. Hastie, R. Tibshirani, *et al.*, *An introduction to statistical learning*, Vol. 112 (Springer, 2013).

[84] V. Thibeault, A. Allard, and P. Desrosiers, Nature Physics , 1 (2024).

[85] H. Lyu, Y. H. Kureh, J. Vendrow, and M. A. Porter, Nature Communications **15**, 224 (2024).

[86] M. Lucas, L. Gallo, A. Ghavasieh, F. Battiston, and M. De Domenico, arXiv:2404.08547 10.48550/arXiv.2404.08547 (2024).

[87] C. Murphy, V. Thibeault, A. Allard, and P. Desrosiers, arXiv preprint arXiv:2206.04000 (2022), arxiv:2206.04000.

CONTENTS

I. RELATION WITH THE HODGE LAPLACIAN MATRICES

The combinatorial Hodge Laplacian matrix [1], first introduced by Eckmann [2], is the combinatorial analogue of the Hodge Laplacian acting on p-forms of differential geometry. It is a $n_p \times n_p$ matrix which describes a particular relation between p-simplices which depends on $(p-1)$ and $(p+1)$-simplices. To build it, we first need to assign an *orientation* to each simplex $\sigma = \{v_0, \dots, v_k\} \in \Delta$, i.e. an ordering of its vertices $\sigma = [v_0, \dots, v_k]$. When the simplices are oriented, it is possible to define a notion of coherence and incoherence of orientation between $(p+1)$ and $(p-1)$-adjacent p-simplices (see [3] for details). We write $\sigma \sim \eta$ when σ, η are coherently oriented and $\sigma \nsim \eta$ when they are incoherently oriented. The Hodge-Laplacian is then defined as

$$(\boldsymbol{\mathcal{L}}_p)_{ij} = \begin{cases} \mathbf{k}_{p,p+1}(\sigma_i) + p + 1 & \text{if } i = j \\ 1 & \text{if } i \neq j,\ a_{(p,p-1)}(\sigma_i, \sigma_j) \neq 0,\ a_{(p,p+1)}(\sigma_i, \sigma_j) = 0 \ \text{ and } \sigma_i \sim \sigma_j \\ -1 & \text{if } i \neq j,\ a_{(p,p-1)}(\sigma_i, \sigma_j) \neq 0,\ a_{(p,p+1)}(\sigma_i, \sigma_j) = 0 \ \text{ and } \sigma_i \nsim \sigma_j \end{cases}, \tag{1}$$

where $\mathbf{k}_{p,p+1}(\sigma_i)$ is the number of $(p+1)$-simplices which contain σ_i and $a_{(p,p\pm1)}$ is the adjacency number as defined in the main text. The pattern of non-zero elements of Equation (1) makes it clear how $\boldsymbol{\mathcal{L}}_p$ describes relations among p-simplices which are adjacent from below but are *not* adjacent from above. When two p-simplices σ, η satisfy this condition, they are said to be *parallel neighbors* [4] and we write $\sigma \parallel \eta$.

Moreover, there are other notable differences between this matrix and the standard graph Laplacian. First, the oriented nature of the simplices is such that the extra-diagonal elements can be both -1 and $+1$, instead of being all -1. Second, the rows in general do not sum to zero. This last property comes from the fact that the vector $\mathbb{1} = (1, \dots, 1)^\top$ does not belong in general to $\ker \boldsymbol{\mathcal{L}}_p$, which instead contains the so-called *harmonic vectors* corresponding to the p-th homology classes of Δ.

These two properties make it so that diffusion through the Hodge-Laplacian is not a standard "intuitive" diffusion process. The fact that $\mathbb{1} \notin \ker \boldsymbol{\mathcal{L}}_p$, for instance, means that the total information present in the p-simplices is *not* conserved in time, as, under the dynamics $\dot{x}(\tau) = -\boldsymbol{\mathcal{L}}_p x(\tau)$

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_i x_i(\tau) = \frac{\mathrm{d}}{\mathrm{d}t} \mathbb{1}^\top x(\tau) = \mathbb{1}^\top \dot{x}(\tau) = \mathbb{1}^\top \boldsymbol{\mathcal{L}}_p x(\tau) \neq 0.$$

The Hodge Laplacian and the graph-like cross-order Laplacian defined in the main text, however, can be related in the following way. In Ref. [4], Forman introduces a combinatorial version of the well-known Weitzenböck identity of Riemannian geometry, which states that the Hodge Laplacian can be decomposed as the sum of the Bochner (or rough) Laplacian and a term depending only on curvature. In our discrete setting this amounts to

$$\boldsymbol{\mathcal{L}}_p = \boldsymbol{\mathcal{L}}_p^B + \mathbf{F}_p, \tag{2}$$

where $\boldsymbol{\mathcal{L}}_p^B$ is the p-th *Bochner Laplacian*, a positive semidefinite matrix defined as

$$(\boldsymbol{\mathcal{L}}_p^B)_{ij} = \begin{cases} \sum_{l \neq i} |\boldsymbol{\mathcal{L}}_p|_{il} & \text{if } i = j \\ (\boldsymbol{\mathcal{L}}_p)_{ij} & \text{if } i \neq j \end{cases}, \tag{3}$$

and $\mathbf{F}_p$, named *Forman curvature*, is a diagonal matrix containing the combinatorial curvatures of the p-simplices. This particular decomposition is such that $\mathbf{F}_p$ can be thought of as a "correction" to the fact that the absolute values of the row elements of $\boldsymbol{\mathcal{L}}_p$ do not sum to 0. We now have that the diagonal elements of the Bochner Laplacian correspond to the number of non-zero elements in each row, i.e. the number of parallel neighbors

$$(\boldsymbol{\mathcal{L}}_p^B)_{ii} = |\{\sigma_j \in \Delta_p \mid \sigma_i \parallel \sigma_j\}| \tag{4}$$

which, given that in a simplicial complex simplices adjacent from above are also adjacent from below, can in turn be related to the higher-order degrees defined in the main text:

$$(\boldsymbol{\mathcal{L}}_p^B)_{ii} = \deg_{(p,p-1)}(\sigma_i) - \deg_{(p,p+1)}(\sigma_i). \tag{5}$$

This tells us that we can write the p-th Bochner Laplacian as the difference of two cross-order Laplacians multiplied element-wise with a "residual" matrix $\mathbf{R}_p$ which includes the effects of orientations:

$$\boldsymbol{\mathcal{L}}_p^B = (\mathbf{L}^\times_{(p,p-1)} - \mathbf{L}^\times_{(p,p+1)}) \odot \mathbf{R} = \mathbf{L}_p^\parallel \odot \mathbf{R}_p, \tag{6}$$

where $(\mathbf{R}_p)_{ij} = -1$ if $\sigma_i \parallel \sigma_j$ and $\sigma_i \sim \sigma_j$ and 1 otherwise. Moreover, it is possible to interpret $\mathbf{L}^{\times}_{(p,p-1)} - \mathbf{L}^{\times}_{(p,p+1)}$ as the Laplacian of a new adjacency graph, which we name *parallel adjacency graph*, where nodes represent p-simplices and edges connect parallel neighbors, hence the notation $\mathbf{L}^{\parallel}_p$.

Finally, putting together Equation (2) with Equation (6), we find

$$\boldsymbol{\mathcal{L}}_p = \mathbf{L}^{\parallel}_p \odot \mathbf{R}_p + \mathbf{F}_p, \tag{7}$$

which suggests that we may interpret diffusion with the Hodge Laplacian, neglecting the contribution of orientation, as a reaction-diffusion process where diffusion takes place among parallel p-simplices and reaction, responsible for the production and destruction of information, is given by the Forman curvature.

II. SCALE-INVARIANCE AND SPECTRAL DIMENSION

To better understand the meaning of the definition of scale-invariance employed in this work, we can look from two different angles. First, we notice,

$$C(\tau) = C^* \iff \frac{\mathrm{d}C(\tau)}{\mathrm{d}\log\tau} = 0 \iff \frac{\mathrm{d}^2 S(\tau)}{\mathrm{d}(\log\tau)^2} = 0 \quad \forall \tau \in I$$

which means that scale-invariance is associated to a range of times where the rate of change of the entropy (in logarithmic scale) is constant, i.e. $S(\tau)$ does not accelerate nor decelerate. From another point of view, it is interesting to look at the relation between the entropic susceptibility and the *spectral dimension* of the adjacency graphs of the simplicial complex. The spectral dimension [5–7] intuitively measures the dimensionality "perceived" by a diffusion process taking place on a manifold or, in our case, a graph. We define the spectral dimension $D_s(\tau)$ as the derivative w.r.t. the logarithmic diffusion time of the logarithm of the return probability $Z(\tau)$

$$D_s(\tau) = -2\frac{\mathrm{d}\log Z(\tau)}{\mathrm{d}\log\tau}. \tag{8}$$

Here τ has to be thought of as a *scale* instead of a time parameter, meaning that $D_s(\tau)$ measures dimensionality at scale τ. On d-dimensional flat manifolds $D_s(\tau) = d$ for all τ, while on d-dimensional (periodic) lattices $D_s(\tau)$ shows a large plateau whose value corresponds exactly to d [7]. In general, when $Z(\tau) \propto \tau^{-2d}$ in a scale interval $I = [\tau_{\min}, \tau_{\max}]$, $D_s(\tau)$ has a plateau equal to d. This implies that, at these specific temporal resolutions, the spectral dimension observed through the diffusion process remains consistent. Such behavior suggests the existence of an inherent dimensionality within the space at these particular scales.

One can see that there is a strong relation between the entropic susceptibility

$$C(\tau) = -\frac{\mathrm{d}S}{\mathrm{d}\log\tau}$$

and the spectral dimension, i.e.

$$C(\tau) = -\frac{1}{2}\frac{\mathrm{d}D_s(\tau)}{\mathrm{d}\log\tau} + \frac{1}{2}D_s(\tau). \tag{9}$$

From Equation (9), we thus find that scale-invariance is equivalent to

$$C(\tau) = C^* \iff D_s(\tau) = a\tau + 2C^* \tag{10}$$

i.e. the spectral dimension varies linearly w.r.t. the scale. This result tells us that informational scale-invariance corresponds not only to spaces with a well-defined intrinsic dimensionality ($a = 0$), but also the case in which the dimension varies linearly with the scale ($a \neq 0$). This last situation, as we can see by taking the definition of spectral dimension (8) and integrating Equation (10), corresponds to

$$Z(\tau) \propto e^{-\frac{a}{2}\tau}\tau^{-C^*}, \tag{11}$$

i.e. changing the scale results in the space becoming increasingly "larger" so that the return probability decreases exponentially fast.

III. IMPLEMENTATION DETAILS AND COMPUTATIONAL COMPLEXITY OF THE RENORMALIZATION METHOD

We report here some further details about the implementation of the algorithm described above.

1. Both the computation of entropic susceptibilities and the renormalization first require the construction of the adjacency matrix $A_{(k,m)}$. This is one of the most computationally expensive part of the algorithm as it requires checking in Δ, for each pair of k-hyperedges σ, η, if there exist m-hyperedges such that they are contained in $\sigma \cap \eta$, if $m < k$, or they contain $\sigma \cup \eta$, if $m > k$. The resulting complexity is thus $O(n_k^2 n_m (k+1+m+1)$, where $(k+1+m+1)$ is the cost of checking the inclusion of a set of $m+1$ elements inside one of $k+1$.

2. After $A_{(k,m)}$ is obtained, we easily get $L_{(k,m)} = D_{(k,m)} - A_{(k,m)}$. If we want to compute the entropic susceptibility, we just need the eigenvalues of $L_{(k,m)}$ while for renormalization we also need the matrix of eigenvectors U. In both cases, we can exactly perform the diagonalization $L_{(k,m)} = U \Lambda U^\top$ at a cost $O(n_k^3)$ or leverage the fact that $L_{(k,m)}$ is symmetric and sparse to employ faster approximate numerical schemes.

3. In the case of renormalization, we first need to pick the hyperparameter $\tau > 0$ corresponding to the scale at which we want to perform the renormalization. Choosing this value requires some tuning, as too small values will result in no nodes being merged, while values too high will contract every vertex which is shared by k-hyperedges to a point. In general, in accordance with [8], we empirically see that picking τ just before the first peak in the entropic susceptibility gives the most stable results.

4. Having picked τ, we compute the density matrix $\boldsymbol{\rho}_{(k,m)}(\tau) = U e^{-\tau \Lambda} U^\top$, from which we build the auxiliary matrix ζ by iterating over every entry in the upper triangular part of $\boldsymbol{\rho}_{(k,m)}$ of at a quadratic cost in the number of k-hyperedges $O(\frac{n_k(n_k-1)}{2})$.

5. We then extract clusters of hyperedges by computing the connected components of the graph with adjacency matrix ζ (which has n_k nodes and $|E|$ edges) at a cost of $O(n_k + |E|)$.

6. Finally, we perform the actual coarse graining. We associate to each node the labels of the k-hyperedges it belongs to by iterating over each k-hyperedge and looking at its nodes (cost $O((k+1)n_k)$). Then, having obtained the map f from the vertices of Δ and the vertices of the new network Δ', we build the hyperedges of Δ' by iterating over every hyperedge in Δ (cost $O(\sum_{i=1}^{k_{\max}} n_i)$) and mapping each one of them through f.

7. The procedure can then be iterated for a chosen number of steps to renormalize Δ' with the same τ or a different one.

Notice that the overall complexity of the algorithm $O(n_k^2 n_m (k+m) + n_k^3 + \sum_{i^1}^{k_{\max}} n_i)$ is controlled by the number of k-hyperedges and m-hyperedges and depends on the actual values of k and m only linearly for the set inclusion check.

The actual values of n_i depend on the specific hypergraph considered. In Table II, we display these values for the family of real-world hypergraphs used in the analysis of Figure 4 of the main text.

IV. SIMPLICIAL PARTITION AS A FUNCTION OF TIME

In Figure S1, we visually display the behavior of steps 2,3 of the renormalization method described in the Section III, applied on a 3-dimensional NGF simplicial complex Δ of 220 vertices with $\beta = 0.1$. In the top part of the figure, we show the entropic susceptibility $C_{(2,3)}$ associated to the diffusion process where information is situated on triangles (2-simplices) and flows through tetrahedra (3-simplices). As discussed in the main text, this simplicial complex is $(2,3)$-scale-invariant, as we see in the large plateau which spans multiple orders of magnitude of τ. We pick 4 different diffusion times $\tau^* \in \{0.6, 5, 100, 6000\}$, where the first one corresponds to the first peak in $C_{(2,3)}$ and the last one falls beyond the plateau. In the second row, we see the simplicial complex Δ with its triangles (2-simplices) colored according to the partition obtained at different times τ^*. In the third row, we see the associated $(2,3)$-adjacency graph with its nodes again colored according to the partition at each time.

As we can see, the $(2,3)$-adjacency graph of a 3-dimensional NGF is composed of cliques of 4 nodes (corresponding to the 4 faces in a single tetrahedron) connected together through single nodes in a tree-like fashion. The first peak in the entropic susceptibility indeed corresponds to the micro-scale associated to the integration of information in these fundamental cliques. In fact, when $\tau^* = 0.6$, the partition of the nodes of $G_{(2,3)}$ tends to assign the same label to nodes in the same clique. As the diffusion time increases, the number of sets in the partition decreases and each set identifies bigger and bigger branches. When τ^* reaches 6000, we see in the rightmost panel, we reach full information integration and all nodes (2-simplices) belong to a single set.

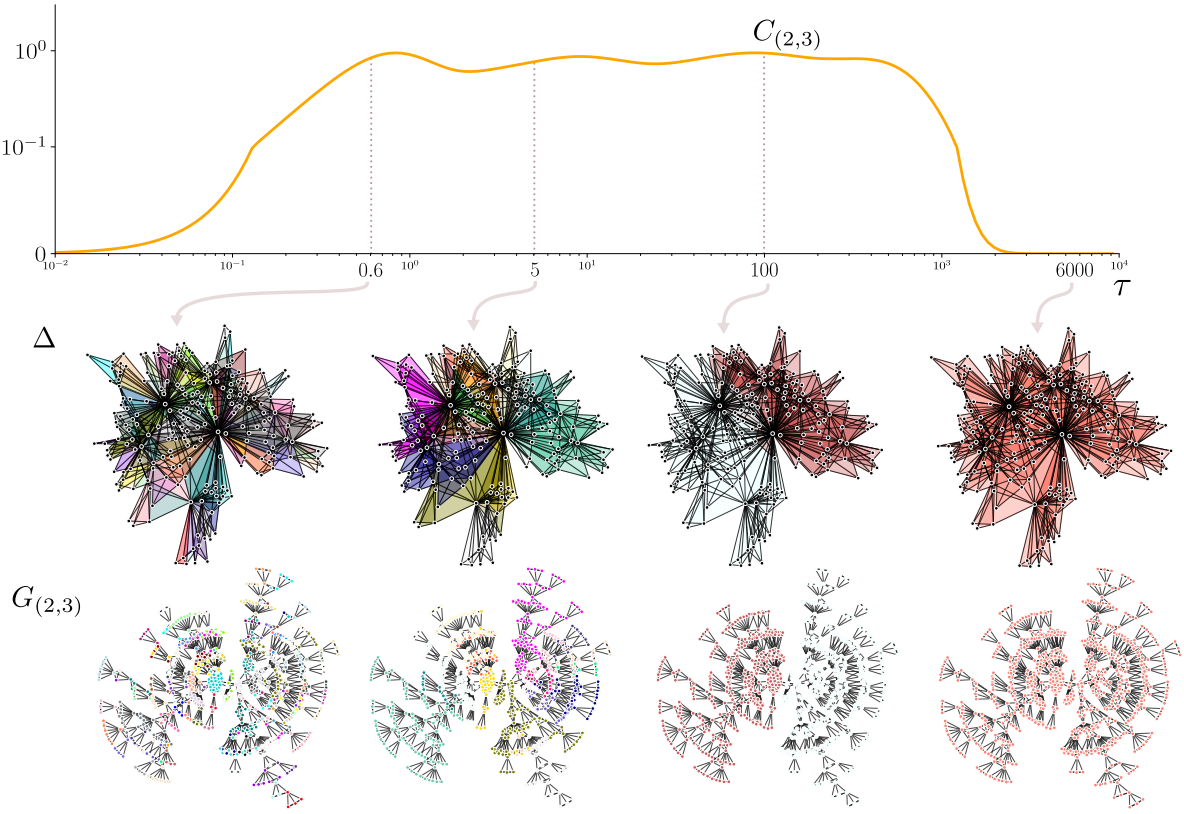

FIG. S1. Simplices partition as a function of diffusion time.

V. RENORMALIZATION OF THE PSEUDOFRACTAL SIMPLICIAL COMPLEX

As discussed in the main text, in some simplicial complexes the constraints imposed by a higher-order coarse graining scheme are essential to have a meaningful reduction of their structure. In Figures S2 and S3, we see the 2-dimensional pseudofractal simplicial complexes after one step of renormalization with all cross-order Laplacians $\mathbf{L}^{\times}_{(k,m)}$ $(k, m \in \{0, 1, 2\}\,,\ k \neq m)$, each with 10 different diffusion times τ. The renormalization with $\mathbf{L}^{\times}_{(1,2)}$ is the only one capable of disentangling the characteristic scales and transform the pseudofractal in a pseudofractal of the same type ($\tau = 2.22$ and $\tau \geq 5.56$ in the figure).

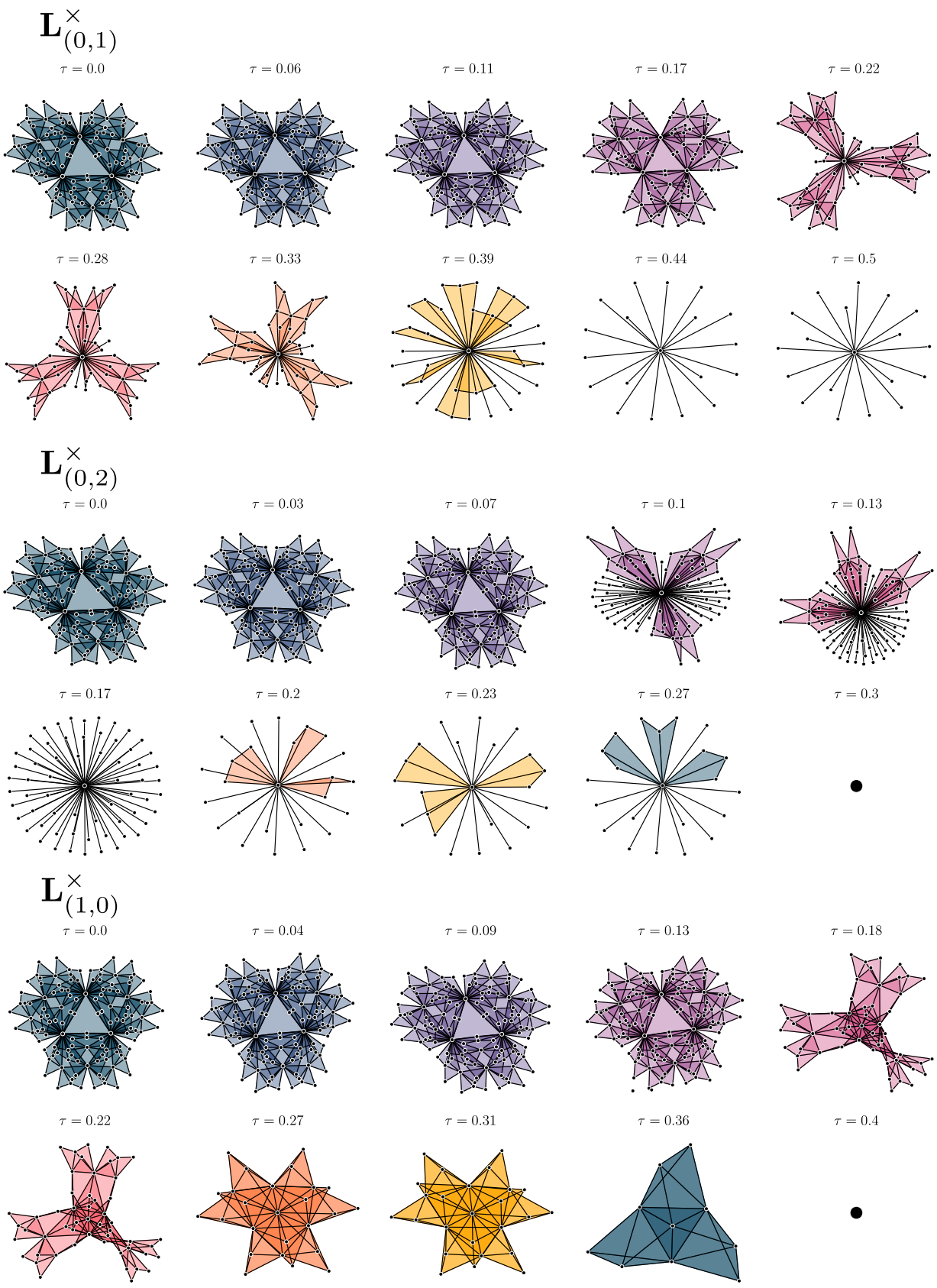

FIG. S2. Renormalization of the 2-dimensional pseudofractal simplicial complex using cross-order Laplacians $\mathbf{L}^{\times}_{(0,1)}$, $\mathbf{L}^{\times}_{(0,2)}$ and $\mathbf{L}^{\times}_{(1,0)}$.

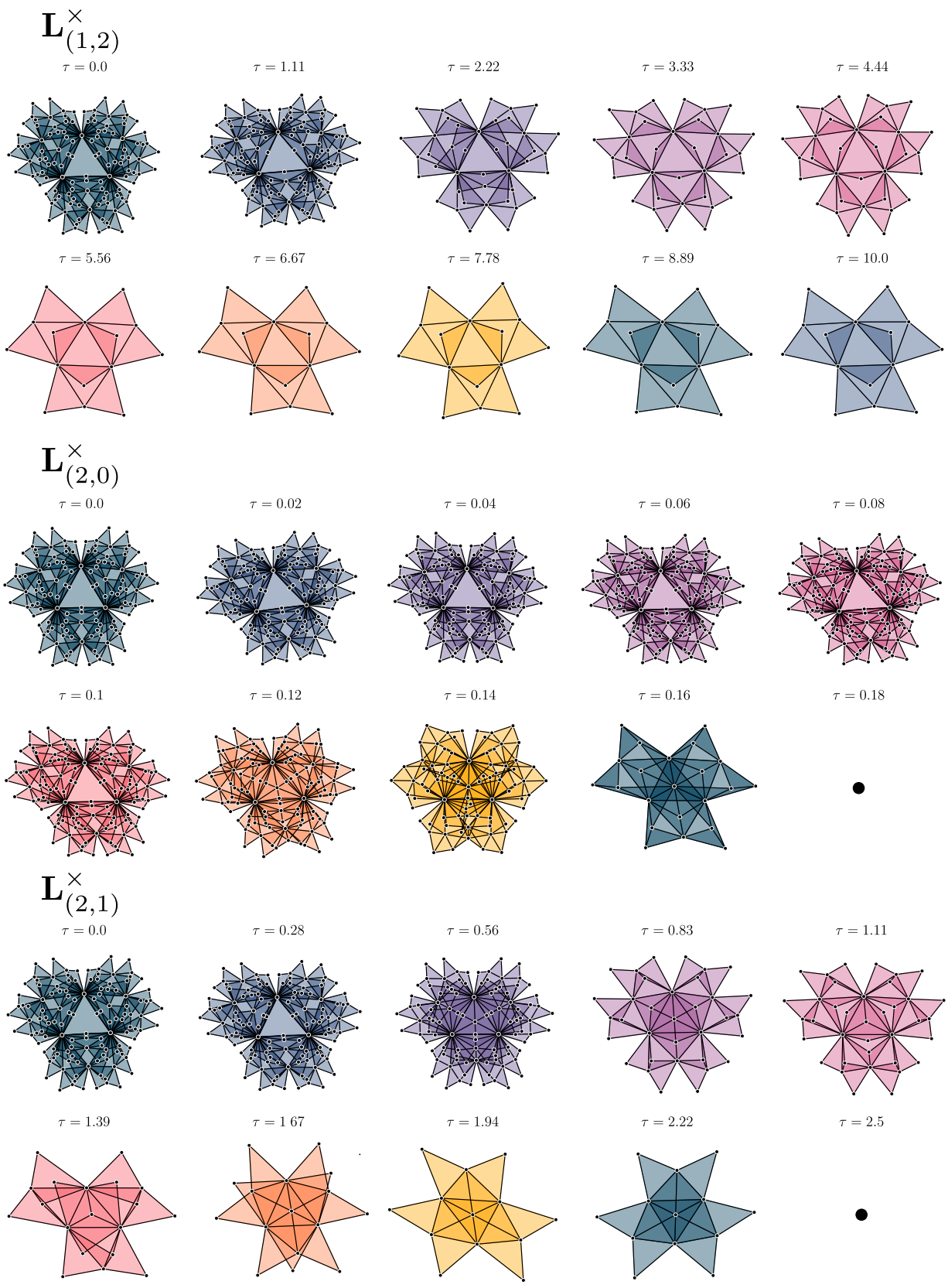

FIG. S3. Renormalization of the 2-dimensional pseudofractal simplicial complex using cross-order Laplacians $\mathbf{L}^{\times}_{(1,2)}$, $\mathbf{L}^{\times}_{(2,0)}$ and $\mathbf{L}^{\times}_{(2,1)}$.

VI. RENORMALIZATION OF NETWORK GEOMETRY WITH FLAVOR

A. Model

The NGF model [9] in dimension d is able to generate both hyperbolic manifolds and *scale-free* growing simplicial complexes, by progressively attaching d-simplices to $(d-1)$-simplices in a stochastic manner biased by the flavor s. The three possible ways in which the growth process is realized, named *flavors*, are the following:

- $s = -1$, which allows for at most 2 d-simplices attached to any $(d-1)$-simplex, resulting in d-dimensional simplicial manifolds;
- $s = 1$, which glues d-simplices to $(d-1)$-simplices by preferential attachment;
- $s = 0$, which presents intermediate properties between the two.

Let us briefly describe the growth process of the model.

At time $t = 1$ the NGF simplicial complex $\Delta^{(1)}$ is made by a single d-simplex. At each time step $t > 1$ a new d-simplex is created and attached to one of the $(d-1)$-simplices present in $\Delta^{(t)}$, chosen with probability

$$\Pi_\sigma^{[s]} = \frac{1}{Z^{[s]}(t)} e^{-\beta\epsilon_\sigma}(1 + s n_\sigma), \tag{12}$$

for each $(d-1)$-simplex σ, where

- β is the inverse temperature, which controls the amount of randomness in the process;
- ϵ_σ is the *energy* of the simplex σ, defined as the sum of the energies of its vertices, which in turn are sampled from a distribution $g(\epsilon)$;
- n_σ is the number of d-simplices which contain σ minus one;
- $Z^{[s]}(t)$ is the normalization constant.

From this it is easy to see that, when $s = -1$, all the $(d-1)$-simplices which are contained in exactly 2 d-simplices will have probability 0 to have another simplex attached to them. This results in $\Delta^{(t)}$ being a simplicial manifold (with boundary) for every $t \geq 1$.

B. Higher-order degrees

In Ref. [9], it is proven that in a d-dimensional NGF simplicial complex with flavor $s = 1$, the generalized higher-order degrees associated to m-simplices are power-law distributed (when $m \leq d-1$). In that work, however, the definition of higher-order degree differs from the one we employ i.e.

$$\deg_{(k,m)}(\sigma) = \sum_{\eta \in \Delta_k} a_{(k,m)}(\sigma, \eta). \tag{13}$$

In particular, the higher-order degree $\mathbf{k}_{m,d}(\sigma)$ of the m-simplex $\sigma \in$ is defined as the number of d-simplices which contain σ, i.e.

$$\mathbf{k}_{m,d}(\sigma) = |\{\eta \in \Delta_d : \ \sigma \subset \eta\}|. \tag{14}$$

In this section, we prove that the degrees defined in these two ways are proportional, with a proportionality constant dependent only on m and d. This means that when $\mathbf{k}_{m,d}$ is power-law distributed, then the same holds for $\deg_{(m,d)}$.

Proposition 1 *If $m < d$ and $\sigma \in \Delta_m$, then*

$$\deg_{(m,d)}(\sigma) = \left[\binom{d+1}{m+1} - 1\right] \mathbf{k}_{m,d}(\sigma). \tag{15}$$

Proof 1 *It holds that*

$$\deg_{(m,d)}(\sigma) = \sum_{\tau \in \Delta_m} a_{(m,d)}(\tau, \sigma) = \sum_{\tau \in \Delta_m, \tau \neq \sigma} |\{\eta \in \Delta_d : \sigma \cup \tau \subseteq \eta\}| = \sum_{\tau \in \Delta_m, \tau \neq \sigma} \sum_{\eta \in \Delta_d} b(\sigma; \eta) b(\tau; \eta)$$

where we define $b(\sigma; \eta) = 1$ *if* $\sigma \subset \eta$ *and* 0 *otherwise. It follows that*

$$\deg_{(m,d)}(\sigma) = \sum_{\tau \in \Delta_m, \tau \neq \sigma} \sum_{\eta \in \Delta_d} b(\sigma; \eta) b(\tau; \eta) = \sum_{\eta \in \Delta_d} b(\sigma; \eta) \sum_{\tau \in \Delta_m, \tau \neq \sigma} b(\tau; \eta) = \sum_{\eta \in \Delta_d, \sigma \subset \eta} |\{\tau \in \Delta_m : \tau \neq \sigma,\ \tau \subset \eta\}|.$$

Given that Δ *is a simplicial complex, every face of a* d*-simplex is a simplex in the simplicial complex. In particular, a* d*-simplex has exactly* $\binom{d+1}{m+1}$ *faces of order* m *and thus*

$$|\{\tau \in \Delta_m : \tau \neq \sigma,\ \tau \subset \eta\}| = \binom{d+1}{m+1} - 1. \tag{16}$$

We can thus conclude the proof by noticing that

$$\deg_{(m,d)}(\sigma) = \left[\binom{d+1}{m+1} - 1\right] \sum_{\eta \in \Delta_d, \sigma \subset \eta} 1 = \left[\binom{d+1}{m+1} - 1\right] \mathbf{k}_{m,d}(\sigma).$$

VII. RENORMALIZATION OF REAL SIMPLICIAL COMPLEXES

A. Entropic susceptibilities after renormalization

As discussed in the main text, we first considered 6 second-order clique complexes obtained from real world dataset taken from the KONECT project [10]. This means that we take each clique of three nodes in the network and consider it as a 2-hyperedge, while also keeping the three edges which formed it. We renormalize each one of them using the cross-order Laplacian associated to the highest scale-invariance parameter together with $\mathbf{L}^{\times}_{(0,1)}$ as a reference, choosing the smallest time τ^* such that the number of nodes is reduced by at least 40%. In Figure S4 the entropic susceptibility associated to the highest SIP is shown, before and after renormalization. We highlight that the metabolic network of the C. Elegans is the only scale-invariant one in which we can observe that the renormalization with $\mathbf{L}^{\times}_{(0,1)}$ better preserves the plateau than $\mathbf{L}^{\times}_{(2,0)}$, despite the SIP being lower. In the E-Road and Co-authorship networks, we instead observe that $L_{(0,1)}$ and the cross-order Laplacian are both able to preserve the plateau. This is a reflection of the fact that both network are scale-invariant in the $(0,1)$ relation too, as shown in the SIPs of Figure 4a of the main text.

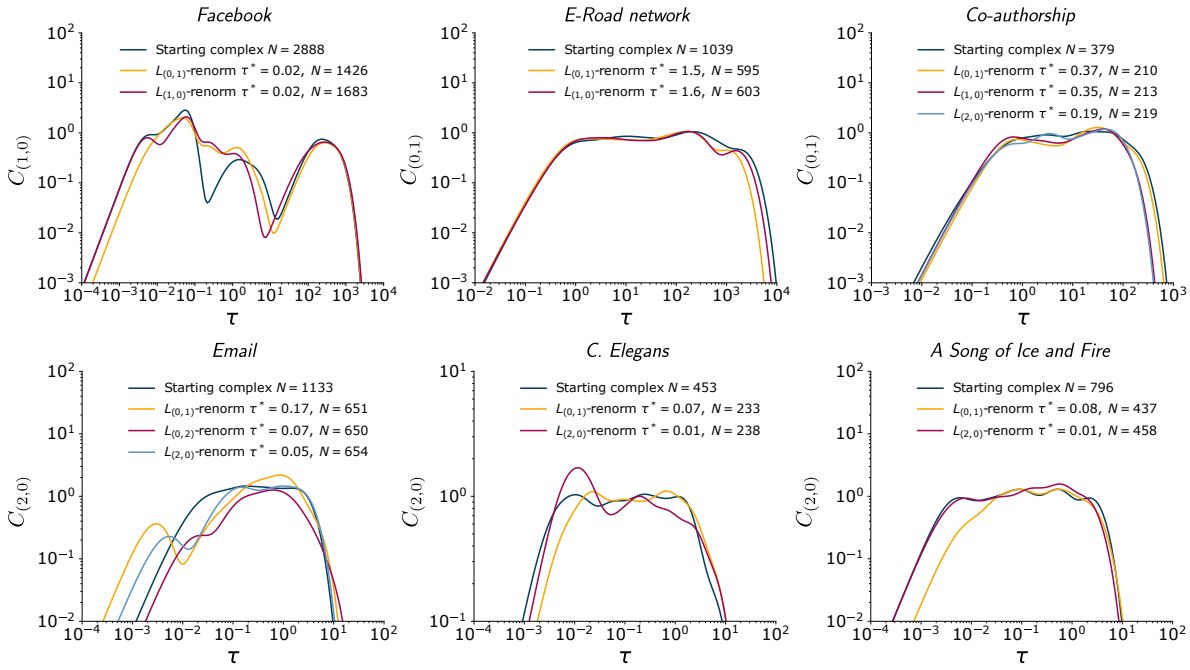

FIG. S4. Entropic susceptibility curves – associated to the highest SIP – of the real data clique-complexes before and after renormalization with different cross-order Laplacians.

B. Comparison with the Bayesian hypergraph reconstruction method

Up to now we built simplicial complexes from networks by considering their associated clique complexes Δ_c, i.e. by filling all of their cliques. One may argue, however, that not all cliques of nodes correspond to actual higher-order interactions. Our SIP measures then effectively refer to relations between cliques and not higher-order interactions.

To check the dependence of our results with respect to this assumption, we considered simplicial complexes Δ_R obtained from the networks with the Bayesian hypergraph reconstruction method of Ref. [11]. Note that with this probabilistic method, we do not fill all the cliques of the network, but only the ones for which there is sufficient statistical. We thus obtain a subcomplex of the clique complex $\Delta_R \subseteq \Delta_C$. In Figure S5a we show the scale-invariance parameters of Δ_R for each dataset (in color), comparing them with the ones obtained on the clique complexes Δ_C (in gray). Using this more refined reconstruction method, the parameters seem to have smaller values than with the clique complex. In particular, we see that the high values of higher-order SIP in Δ_C drop to negligible values in Δ_R, meaning that the scale-invariance of the structural organization of cliques is lost.

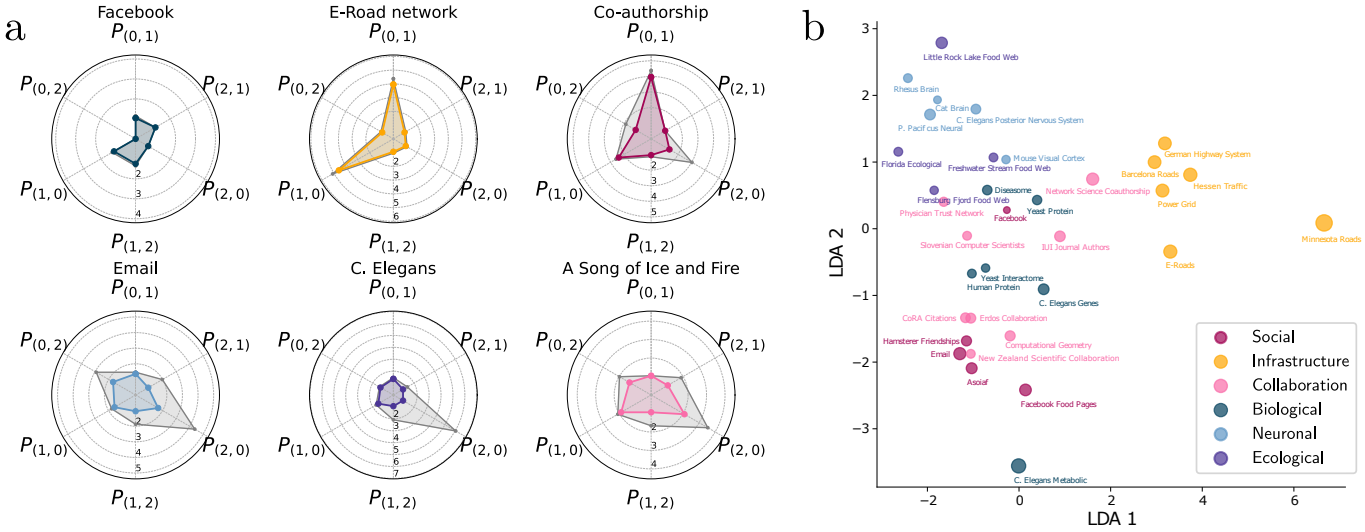

FIG. S5. **a.** Scale-invariance parameters for the 2nd order simplicial complexes obtained with the hypergraph reconstruction method (in color) and the clique complexes (in gray). **b.** A more detailed version of the Figure 4 in the main text. The projections of the 34 datasets onto the first two PCA components is shown, together with their name. The color of each point represents the type of the associated network, while the size is proportional to the value of its highest SIP.

C. Scale-invariance parameters of real datasets

As we mentioned in the main text, we then considered a larger set of pairwise networks taken from different network archives: *KONECT* [10], *ICON* [12] and *The Network Data Repository* [13]. The networks, whose details are found in Table I, were chosen to belong to the following classes:

- *Infrastructure* — road networks and power grids;
- *Collaboration* — scientific collaboration networks;
- *Social* — email networks, real and fictional social networks;
- *Biological* — gene interaction networks and protein-protein interaction networks;
- *Connectome* — brain networks;
- *Ecological* — species food webs in different ecological environments.

Each one of the networks is lifted to a simplicial complex by computing the associated clique complex. This means taking all cliques up to a chosen order and seeing them as simplices. We then compute the scale-invariance parameters up to the second order, that is $P_{(0,1)}$, $P_{(0,2)}$, $P_{(1,0)}$, $P_{(1,2)}$, $P_{(2,0)}$ and $P_{(2,1)}$ and use the resulting values as coordinates to embed it as a point in $\mathbb{R}^6$. The resulting point cloud is then projected onto its first two principal components (81.2% of total variance explained) allowing us to get Figure S5b. As we can see, networks of similar type tend to correspond to closer points. Infrastructure networks in yellow are all on the right, the collaboration and social networks, which are necessarily similar in nature, occupy the left side. In the bottom side, we see biological networks, connectomes and ecological food webs, which, while they appear more mixed, they are all of biological origin.

Name	**Type**	**Repository**	**Source**	n_0	n_1	n_2	n_3
Facebook	Social	KONECT	[14]	2890	2982	91	4
E-Road	Infrastructure	KONECT	[15]	1176	1418	32	0
Network Science Coauthorship	Collaboration	KONECT	[16]	1461	2742	3764	7159
Email	Social	KONECT	[17]	1133	5451	5343	3419
C. Elegans Metabolic	Biological	KONECT	[18–20]	453	2025	3284	2967
A Song of Ice and Fire	Social	KONECT	[21]	796	2823	5655	9331
Human Protein	Biological	KONECT	[22]	3133	6726	1047	142
Power Grid	Infrastructure	KONECT	[23]	4943	6595	651	90
Yeast Protein	Biological	KONECT	[24–27]	1872	2278	222	41
Hamsterer Friendships	Social	KONECT	[10]	1858	12534	16750	10015
German Highway System	Infrastructure	ICON	[28]	1168	1243	2	0
Cat Brain	Connectome	ICON	[29]	65	1139	3613	10125
Rhesus Brain	Connectome	ICON	[30]	91	582	1902	4048
Mouse Visual Cortex	Connectome	ICON	[31]	195	214	4	0
C. Elegans Posterior Nervous System	Connectome	ICON	[32]	272	4451	10005	16839
IUI Journal Authors	Collaboration	ICON	[33]	2288	3377	2411	1904
CoRA Citations	Collaboration	ICON	[34]	3585	11258	8631	3940
Flensburg Fjord Food Web	Ecological	ICON	[35]	180	1569	3440	4175
Hessen Traffic	Infrastructure	ICON	[36]	4660	6026	192	0
Minnesota Roads	Infrastructure	Net. Data. Rep.	[37]	2641	3302	53	0
Yeast Interactome	Biological	ICON	[38]	1278	1808	110	6
Barcelona Roads	Infrastructure	ICON	[39]	930	1798	164	1
Freshwater Stream Food Web	Ecological	ICON	[40]	113	832	124	0
Little Rock Lake Food Web	Ecological	ICON	[41]	183	2452	11292	35978
P. Pacificus Neural	Connectome	ICON	[42]	54	141	117	32
Slovenian Computer Scientists	Collaboration	ICON	[43]	262	596	488	224
New Zealand Scientific Collaboration	Collaboration	ICON	[44]	1511	4273	9337	21986
Computational Geometry	Collaboration	ICON	[45]	6158	11897	13587	20994
Physician Trust Network	Collaboration	ICON	[46]	243	924	627	220
Erdos Collaboration	Collaboration	ICON	[47]	5094	7514	1607	450
Facebook Food Pages	Social	Net. Data. Rep.	[48]	942	2101	0	0
Diseasome	Biological	Net. Data. Rep.	[49]	517	1189	1360	1391
C. Elegans Genes	Biological	Net. Data. Rep.	[50]	2989	4659	118	2
Florida Ecological	Ecological	Net. Data. Rep.	[51, 52]	128	2074	8436	14126

TABLE I. Datasets used for Figure S5b

D. Scale-invariance parameters of real hypergraphs

The analysis of the scale-invariance parameters we performed for clique complexes can be similarly applied to genuine higher-order networks. We pick a heterogeneous family of 20 real hypergraphs from the repository hosted by the XGI package [53] and compute their SIPs up to the 4-th order. These hypergraph are of different types, from contact networks to gene-disease networks, and, as we did for clique-complexes, we divide them into three coarse families: "social", "collaboration" and "biological". The values of the SIPs are shown in Figure S6 where the red polar plots are associated to "social" hypergraphs, pink to "collaboration" and blue to "biological".

A first look shows us that there are groups of networks of similar nature, whose SIPs representations appear to have very similar shapes. For instance, we see how the school and hospital contact networks display a *flower-like* scale-invariance profile, email networks show an almost constant SIP value across the different orders (except for $(4, 3)$), while bills networks show a *fan-like* shape on the left with three rays on the right.

To show the discriminatory power of the SIP measure, we once again use the values as coordinates to embed the hypergraphs into a high-dimensional space. Guided by the three macro-classes, we perform a linear discriminant (LDA) supervised dimensionality reduction, which allows us to project the datasets into the plane which maximally separates the classes.

We see in Figure S7a how the LDA projection is able to clearly separate the two classes, meaning that the information contained in the SIP profiles is able to distinguish their nature. In Figure S7, the colors show that the same projection is able to distinguish some of the networks at a finer level.

FIG. S6. Scale-invariance parameters of 20 real-world hypergraphs up to the 4-th order.

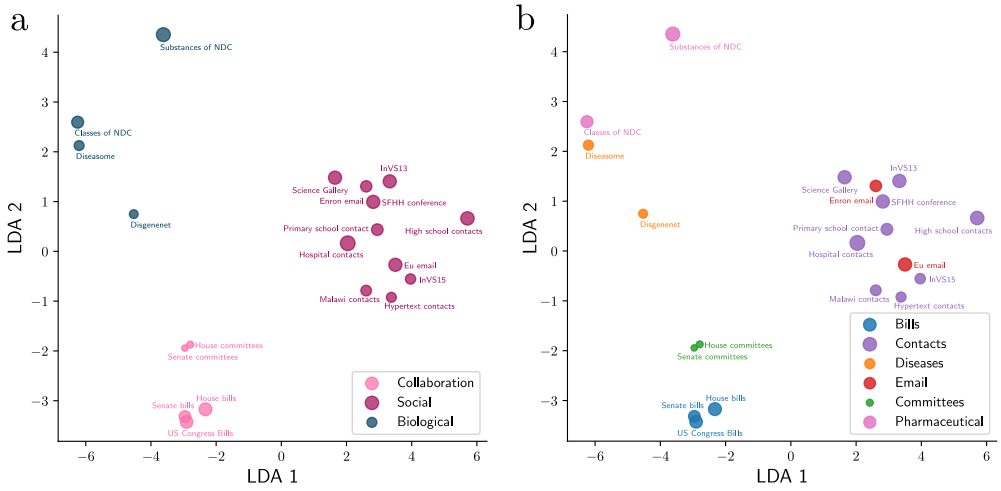

FIG. S7. LDA projection of the SIPs of the real-world hypergraphs, colored according to (**a**) the macro-class and (**b**) a finer classification.

Name	**Type**	**Finer Type**	**Source**	n_0	n_1	n_2	n_3	n_4
High School Contacts	Social	Contacts	[54]	327	5498	2091	222	7
Primary School Contacts	Social	Contacts	[55–57]	242	7748	4600	347	9
Enron Email	Social	Email	[58, 59]	143	809	317	138	63
Eu Email	Social	Email	[58, 60, 61]	986	12753	4938	2294	1359
Lyon Hospital Contacts	Social	Contacts	[57, 62]	75	1107	657	58	2
Hypertext Contacts	Social	Contacts	[57, 63]	113	2103	302	18	7
InVS13	Social	Contacts	[64]	92	741	44	2	0
InVS15	Social	Contacts	[65]	217	4142	755	12	0
Malawi Village Contacts	Social	Contacts	[66]	84	341	86	4	0
Science Gallery	Social	Contacts	[57, 63]	410	2491	808	46	5
SFHH Conference	Social	Contacts	[55, 65, 67]	403	8268	1861	258	63
US Congress Bills	Collaboration	Bills	[58, 68, 69]	1718	13871	10156	7764	5780
House Bills	Collaboration	Bills	[53]	1494	7656	4369	3350	2607
House Committees	Collaboration	Committees	[53]	1290	1	2	11	25
Senate Bills	Collaboration	Bills	[53]	294	3280	3013	2394	1685
Senate Committees	Collaboration	Committees	[53]	282	0	0	9	22
Diseasome	Biological	Disease	[49]	516	153	92	26	25
DisGeNet	Biological	Disease	[70]	12368	157	139	93	66
Classes of NDC	Biological	Pharmaceutical	[58]	628	135	92	100	73
Substances of NDC	Biological	Pharmaceutical	[58]	3414	1027	717	516	498

TABLE II. Datasets from XGI package [53] used for Figure S6 and Figure S7

[1] L.-H. Lim, Siam Review **62**, 685 (2020).
[2] B. Eckmann, Commentarii Mathematici Helvetici **17**, 240 (1944).
[3] J. R. Munkres, *Elements of Algebraic Topology* (CRC press, 2018).
[4] Forman, Discrete & Computational Geometry **29**, 323 (2003).
[5] R. Rammal and G. Toulouse, Journal de Physique Lettres **44**, 13 (1983).
[6] J. Ambjørn, J. Jurkiewicz, and R. Loll, Physical Review Letters **95**, 171301 (2005).
[7] G. Calcagni, D. Oriti, and J. Thürigen, Classical and Quantum Gravity **31**, 135014 (2014).
[8] P. Villegas, T. Gili, G. Caldarelli, and A. Gabrielli, Nature Physics **19**, 445 (2023).
[9] G. Bianconi and C. Rahmede, Physical Review E **93**, 032315 (2016).
[10] J. Kunegis, in *Proc. Int. Conf. on World Wide Web Companion* (2013) pp. 1343–1350.
[11] J.-G. Young, G. Petri, and T. P. Peixoto, Communications On Physics **4**, 1 (2021).
[12] A. Clauset, E. Tucker, and M. Sainz (2016).
[13] R. A. Rossi and N. K. Ahmed, in *AAAI* (2015).
[14] J. Leskovec and J. Mcauley, Advances in neural information processing systems **25** (2012).
[15] L. Šubelj and M. Bajec, The European Physical Journal B **81**, 353 (2011).
[16] M. E. Newman, Physical Review E **74**, 036104 (2006).
[17] R. Guimera, L. Danon, A. Diaz-Guilera, F. Giralt, and A. Arenas, Physical Review E **68**, 065103 (2003).
[18] J. Duch and A. Arenas, Physical Review E **72**, 027104 (2005).
[19] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási, Nature **407**, 651 (2000).
[20] R. Overbeek, N. Larsen, G. D. Pusch, M. D'Souza, E. S. Jr, N. Kyrpides, M. Fonstein, N. Maltsev, and E. Selkov, Nucleic acids research **28**, 123 (2000).
[21] A. Beveridge and M. Hunger, Asoiaf (2013).
[22] J.-F. Rual, K. Venkatesan, T. Hao, T. Hirozane-Kishikawa, A. Dricot, N. Li, G. F. Berriz, F. D. Gibbons, M. Dreze, N. Ayivi-Guedehoussou, *et al.*, Nature **437**, 1173 (2005).
[23] D. J. Watts and S. H. Strogatz, nature **393**, 440 (1998).
[24] H. Jeong, S. P. Mason, A.-L. Barabási, and Z. N. Oltvai, Nature **411**, 41 (2001).
[25] S. Coulomb, M. Bauer, D. Bernard, and M.-C. Marsolier-Kergoat, Proc. R. Soc. B: Biol. Sci. **272**, 1721 (2005).
[26] J.-D. J. Han, D. Dupuy, N. Bertin, M. E. Cusick, and M. Vidal, Nature Biotechnology **23**, 839 (2005).
[27] M. P. Stumpf, C. Wiuf, and R. M. May, Proceedings of the National Academy of Sciences of the United States of America **102**, 4221 (2005).
[28] M. Kaiser and C. C. Hilgetag, Physical Review E **69**, 036103 (2004).
[29] M. A. de Reus and M. P. van den Heuvel, Journal of Neuroscience **33**, 12929 (2013).
[30] N. T. Markov, J. Vezoli, P. Chameau, A. Falchier, R. Quilodran, C. Huissoud, C. Lamy, P. Misery, P. Giroud, S. Ullman, *et al.*, Journal of Comparative Neurology **522**, 225 (2014).
[31] D. D. Bock, W.-C. A. Lee, A. M. Kerlin, M. L. Andermann, G. Hood, A. W. Wetzel, S. Yurgenson, E. R. Soucy, H. S. Kim, and R. C. Reid, Nature **471**, 177 (2011).
[32] T. A. Jarrell, Y. Wang, A. E. Bloniarz, C. A. Brittin, M. Xu, J. N. Thomson, D. G. Albertson, D. H. Hall, and S. W. Emmons, Science (New York, N.Y.) **337**, 437 (2012).
[33] N. Blagus and M. Bajec, Uporabna Informatika **23**, 22 (2015).
[34] A. K. McCallum, K. Nigam, J. Rennie, and K. Seymore, Information Retrieval **3**, 127 (2000).
[35] C. D. Zander, N. Josten, K. C. Detloff, R. Poulin, J. P. McLaughlin, and D. W. Thieltges, Ecology **92**, 2007 (2011).
[36] E. C. Sancho, The hessen asymmetric network (2016).
[37] D. Gleich, Minnesota road network (2010).
[38] D. Bu, Y. Zhao, L. Cai, H. Xue, X. Zhu, H. Lu, J. Zhang, S. Sun, L. Ling, N. Zhang, *et al.*, Nucleic acids research **31**, 2443 (2003).
[39] B. Stabler, The barcelona network (2016).
[40] R. M. Thompson and CR. Townsend, Ecology **84**, 145 (2003).
[41] N. D. Martinez, Ecological monographs **61**, 367 (1991).
[42] D. J. Bumbarger, M. Riebesell, C. Rödelsperger, and R. J. Sommer, Cell **152**, 109 (2013).
[43] N. Blagus, L. Šubelj, and M. Bajec, Physica A: Statistical Mechanics and its Applications **391**, 2794 (2012).
[44] S. Aref, D. Friggens, and S. Hendy, in *Proceedings of the Australasian Computer Science Week Multiconference* (2018) pp. 1–10.
[45] B. Jones, Computational geometry database (2002).
[46] J. Coleman, E. Katz, and H. Menzel, Sociometry **20**, 253 (1957).
[47] V. Batagelj and A. Mrvar, Social Networks **22**, 173 (2000).
[48] B. Rozemberczki, R. Davies, R. Sarkar, and C. Sutton, in *Proceedings of the 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining 2019* (ACM, 2019) pp. 65–72.
[49] K.-I. Goh, M. E. Cusick, D. Valle, B. Childs, M. Vidal, and A.-L. Barabási, Proceedings of the National Academy of Sciences of the United States of America **104**, 8685 (2007).
[50] A. Cho, J. Shin, S. Hwang, C. Kim, H. Shim, H. Kim, H. Kim, and I. Lee, Nucleic acids research **42**, W76 (2014).
[51] R. E. Ulanowicz and D. L. DeAngelis, FY97: The Florida Bay Ecosystem , 20688 (1998).

[52] C. J. Melián and J. Bascompte, Ecology **85**, 352 (2004).
[53] N. W. Landry, M. Lucas, I. Iacopini, G. Petri, A. Schwarze, A. Patania, and L. Torres, Journal of Open Source Software **8**, 5162 (2023).
[54] R. Mastrandrea, J. Fournet, and A. Barrat, PLoS One **10**, e0136497 (2015).
[55] J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, L. Isella, J.-F. Pinton, M. Quaggiotto, W. Van den Broeck, C. Régis, B. Lina, and P. Vanhems, PLoS One **6**, e23176 (2011).
[56] V. Gemmetto, A. Barrat, and C. Cattuto, BMC Infect Dis **14**, 1 (2014).
[57] The sociopatterns collaboration.
[58] A. R. Benson, R. Abebe, M. T. Schaub, A. Jadbabaie, and J. Kleinberg, Proc Natl Acad Sci U.S.A **115**, E11221 (2018).
[59] Enron email dataset (2024).
[60] H. Yin, A. R. Benson, J. Leskovec, and D. F. Gleich, in *ACM Conferences* (Association for Computing Machinery, New York, NY, USA, 2017) pp. 555–564.
[61] J. Leskovec, J. Kleinberg, and C. Faloutsos, ACM Trans Knowl Discov Data **1**, 2 (2007).
[62] P. Vanhems, A. Barrat, C. Cattuto, J.-F. Pinton, N. Khanafer, C. Régis, B.-a. Kim, B. Comte, and N. Voirin, PLoS One **8**, e73970 (2013).
[63] L. Isella, J. Stehlé, A. Barrat, C. Cattuto, J.-F. Pinton, and W. Van den Broeck, J Theor Biol **271**, 166 (2011).
[64] M. Génois, C. L. Vestergaard, J. Fournet, A. Panisson, I. Bonmarin, and A. Barrat, Network Sci **3**, 326 (2015).
[65] M. Génois and A. Barrat, EPJ Data Sci **7**, 1 (2018).
[66] L. Ozella, D. Paolotti, G. Lichand, J. P. Rodríguez, S. Haenni, J. Phuka, O. B. Leal-Neto, and C. Cattuto, EPJ Data Sci **10**, 1 (2021).
[67] C. Cattuto, W. Van den Broeck, A. Barrat, V. Colizza, J.-F. Pinton, and A. Vespignani, PLoS One **5**, e11596 (2010).
[68] J. H. Fowler, Political Analysis **14**, 456 (2006).
[69] J. H. Fowler, Social Networks **28**, 454 (2006).
[70] J. Piñero, J. M. Ramírez-Anguita, J. Saüch-Pitarch, F. Ronzano, E. Centeno, F. Sanz, and L. I. Furlong, Nucleic Acids Res **48**, D845 (2020).