

# Structural Reducibility of Hypergraphs

Alec Kirkley<sup>1,2,3,\*</sup>, Helcio Felipe<sup>4</sup>, and Federico Battiston<sup>4,†</sup>

<sup>1</sup>*Institute of Data Science, University of Hong Kong, Hong Kong SAR, China*

<sup>2</sup>*Department of Urban Planning and Design, University of Hong Kong, Hong Kong SAR, China*

<sup>3</sup>*Urban Systems Institute, University of Hong Kong, Hong Kong SAR, China*

<sup>4</sup>*Department of Network and Data Science, Central European University, 1100 Vienna, Austria*



(Received 19 March 2025; accepted 3 November 2025; published 12 December 2025)

Higher-order interactions provide a nuanced understanding of the relational structure of complex systems beyond traditional pairwise interactions. However, higher-order network analyses also incur more cumbersome interpretations and greater computational demands than their pairwise counterparts. Here, we present an information-theoretic framework for determining the extent to which a hypergraph representation of a networked system is structurally redundant and for identifying its most critical higher orders of interaction that allow us to remove these redundancies while preserving essential higher-order structure.

DOI: [10.1103/xrn7-cz8v](https://doi.org/10.1103/xrn7-cz8v)

A wide variety of complex systems and relational data are characterized by higher-order, nondyadic interactions [1–5]. Such systems can be conveniently represented as hypergraphs, collections of nodes representing fundamental units of a system that are connected by hyperedges encoding interactions among an arbitrary number of nodes [6]. To investigate the higher-order architecture of networked systems, new mathematical and computational frameworks have been proposed [7–10], revealing previously unknown organizational principles and new emergent behaviors in collective phenomena ranging from contagions [11–13] and diffusion [14] to synchronization [15–19] and evolutionary dynamics [20–22]. Nevertheless, due to the high dimensionality of many real-world hypergraphs, higher-order network analyses are typically more computationally demanding and complex than pairwise network analyses. Hence, it is important to identify and exploit redundancies—which have been observed in real-world systems [23–26]—to construct more compressed representations that retain the key structural heterogeneity present in a system's original higher-order structure.

Inspired by related work in the context of multilayer networks [27–29], here we provide a simple and principled information-theoretic solution to identify the *structural reducibility* of a hypergraph—the extent to which a hypergraph provides redundant information about a system's relational structure—and remove these redundancies to create a reduced representation that retains its critical higher-order structure. Our method is interpretable and computationally efficient and can be generalized to capture the reducibility of hypergraphs when viewed at different

scales. We test our framework on a variety of synthetic network models, showcasing its wide applicability and robustness to different sources of statistical noise. Finally, we apply the framework to a corpus of real-world higher-order systems from various application domains, finding that many of these systems can be substantially structurally reduced.

**Hypergraph reducibility**—Let  $G = \{G^{(\ell)}\}_{\ell \in \mathcal{L}}$  be a hypergraph with  $L$  unique (but not necessarily consecutive) layers  $G^{(\ell)}$  indexed by  $\ell$ , each layer  $G^{(\ell)}$  containing all hyperedges of size  $\ell$  from  $G$ . Let  $\mathcal{L} = \{\ell_1, \dots, \ell_L\}$  denote the set of  $L$  unique layer indices. For example, a hypergraph  $G$  with only layers  $\ell = 2$  and  $\ell = 5$  would have  $G = \{G^{(2)}, G^{(5)}\}$  and  $\mathcal{L} = \{2, 5\}$ . We consider  $G^{(\ell)}$  as a set of undirected, sorted tuples of size  $\ell$  with no repeated entries, and let  $E^{(\ell)} = |G^{(\ell)}|$  be the number of hyperedges in  $G^{(\ell)}$  (in other words,  $G$  is a simple, undirected hypergraph). There are  $\binom{N}{\ell}$  possible undirected, sorted tuples of size  $\ell$  so  $E^{(\ell)} \leq \binom{N}{\ell}$ . We also let  $G^{(k \rightarrow \ell)}$  be the projection of layer  $k > \ell$  onto the order of  $\ell$ , which extracts all unique  $\ell$ -tuples nested within the  $k$ -tuples in  $G^{(k)}$ . For example, if  $G^{(3)} = \{(0, 1, 2), (0, 2, 4)\}$ , we would have  $G^{(3 \rightarrow 2)} = \{(0, 1), (0, 2), (1, 2), (0, 4), (2, 4)\}$ . We define  $E^{(k \rightarrow \ell)} = |G^{(k \rightarrow \ell)}|$ , similarly to the unprojected layers, and use the convention  $G^{(\ell \rightarrow \ell)} = G^{(\ell)}$ .

The structural reducibility of a hypergraph  $G$  can be defined based on the overlap among its pairs of layers  $(G^{(k)}, G^{(\ell)})$ ,  $k, \ell \in \mathcal{L}$ , where overlap is defined based on the projection of each layer to the lower order of the two. This convention is required because higher-order hyperedges have unique projections onto lower-order interactions (as defined above), but one cannot conversely determine higher-order structure uniquely from lower-order

\*Contact author: [alec.w.kirkley@gmail.com](mailto:alec.w.kirkley@gmail.com)

†Contact author: [battistonf@ceu.edu](mailto:battistonf@ceu.edu)

structure alone [1]. Higher overlap among the layers indicates higher structural redundancy among different orders of hyperedges, suggesting a higher structural reducibility for the hypergraph  $G$ . Formally, we can define the overlap of the layers indexed by  $k, \ell$  as  $E^{(k \cap \ell)} = |G^{[k \rightarrow \min(k, \ell)]} \cap G^{[\ell \rightarrow \min(k, \ell)]}|$ , which is the number of hyperedges the two layers share when they are projected to the lower order of the two layers.

Our proposed reducibility measure reflects the extent to which a hypergraph  $G$  can be *compressed* in an information-theoretic sense when we exploit the structural redundancy among its hyperedge layers of different orders (i.e., their layer overlaps). To formalize this concept mathematically, we consider transmitting the hypergraph  $G$  to a receiver using two different schemes. In the first (naïve) scheme, we transmit each of  $G$ 's layers  $G^{(\ell)}$  individually. In the second scheme, we first transmit a set of  $R \leq L$  “representative” layers  $\mathcal{R} \subseteq \mathcal{L}$  which capture most of the heterogeneity in the hyperedge structure in  $G$ , and then we transmit each remaining layer  $\ell \in \mathcal{L} \setminus \mathcal{R}$  as a noisy copy of a representative layer  $r(\ell) \in \mathcal{R}$ . A similar concept has been employed to compress multilayer network structures [30] and sets of network partitions [31] using intermediate representative structures.

We assume that the receiver knows the orders  $\mathcal{L}$  of the layers and the number of hyperedges  $E^{(\ell)}$  for each  $\ell \in \mathcal{L}$ —specifying these counts incurs a comparatively negligible information cost anyway. Since each layer  $G^{(\ell)}$  has  $\binom{N}{E^{(\ell)}}$  possible configurations of its hyperedges when  $E^{(\ell)}$  is known, we need to send a bit string of length approximately equal to  $\log_2 \binom{N}{E^{(\ell)}}$  bits to fully specify which configuration corresponds to  $G^{(\ell)}$ . The naïve transmission of  $G$  as individual layers, therefore, requires an information content of

$$H_0 = \sum_{\ell \in \mathcal{L}} \log \binom{N}{E^{(\ell)}} \quad (1)$$

bits, using the convention  $\log \equiv \log_2$  for brevity.

A better way to transmit  $G$  is to exploit the overlaps among layers of hyperedges of different sizes to save information. To do this, we first transmit a representative subset of  $R \leq L$  layers indexed by  $\mathcal{R} \subseteq \mathcal{L}$ , which incurs a cost of  $\sum_{r \in \mathcal{R}} \log \binom{N}{E^{(r)}}$  bits. Then we transmit each remaining layer  $\ell \in \mathcal{L} \setminus \mathcal{R}$  by (a) transmitting a representative layer  $r(\ell) \in \mathcal{R}$  of a higher order, costing us  $\log R$  bits as there are  $R$  layers to choose from; (b) transmitting the overlap  $E^{[r(\ell) \cap \ell]}$  among the layer  $\ell$  and its representative  $r(\ell)$ , costing us  $\log(E^{[r(\ell) \cap \ell]} + 1)$  bits as  $E^{[r(\ell) \cap \ell]} \in [0, E^{[r(\ell) \rightarrow \ell]}]$ ; and (c) transmitting layer  $\ell$  given

the constraints imposed by the overlap value  $E^{[r(\ell) \cap \ell]}$ . If  $E^{[r(\ell) \cap \ell]}$  of the possible  $E^{[r(\ell) \rightarrow \ell]}$  edges in  $G^{[r(\ell) \rightarrow \ell]}$  are present in  $G^{(\ell)}$ , and  $E^{(\ell)} - E^{[r(\ell) \cap \ell]}$  of the  $\binom{N}{\ell} - E^{[r(\ell) \rightarrow \ell]}$  possible edges absent from  $G^{[r(\ell) \rightarrow \ell]}$  are present in  $G^{(\ell)}$ , then we require  $\log \binom{E^{[r(\ell) \rightarrow \ell]}}{E^{[r(\ell) \cap \ell]}} \binom{\binom{N}{\ell} - E^{[r(\ell) \rightarrow \ell]}}{E^{(\ell)} - E^{[r(\ell) \cap \ell]}}$  bits to specify  $G^{(\ell)}$  given knowledge of  $G^{[r(\ell) \rightarrow \ell]}$  and  $E^{[r(\ell) \cap \ell]}$ . Steps (a) and (b) incur negligible information costs compared to step (c) and can be ignored. The total information content of this scheme is then

$$H_G(\mathcal{R}) = \sum_{r \in \mathcal{R}} \log \binom{N}{E^{(r)}} + \sum_{\ell \in \mathcal{L} \setminus \mathcal{R}} \log \binom{E^{[r(\ell) \rightarrow \ell]}}{E^{[r(\ell) \cap \ell]}} \binom{\binom{N}{\ell} - E^{[r(\ell) \rightarrow \ell]}}{E^{(\ell)} - E^{[r(\ell) \cap \ell]}} \quad (2)$$

bits. For a minimal information cost, the representative layer  $r(\ell) \in \mathcal{R}$  for each layer  $\ell \notin \mathcal{R}$  can be assigned as

$$r(\ell) = \operatorname{argmin}_{r \in \mathcal{R}, r > \ell} \left\{ \log \binom{E^{(r \rightarrow \ell)}}{E^{(r \cap \ell)}} \binom{\binom{N}{\ell} - E^{(r \rightarrow \ell)}}{E^{(\ell)} - E^{(r \cap \ell)}} \right\}. \quad (3)$$

The information cost  $H_G(\mathcal{R})$  of this transmission scheme depends on which layers  $\mathcal{R} \subseteq \mathcal{L}$  are selected as representatives—the better the layers  $\mathcal{R}$  capture the heterogeneity in the hyperedge structure in  $G$ , the lower the information cost  $H_G(\mathcal{R})$ . Thus, to maximize compression from layer overlap, we must find the optimal set of representative layers  $\mathcal{R}^*$  according to

$$\mathcal{R}^* = \operatorname{argmin}_{\mathcal{R} \subseteq \mathcal{L}} \{H_G(\mathcal{R})\}. \quad (4)$$

We will describe shortly how to solve this optimization problem.

The optimal information cost  $H_G(\mathcal{R}^*)$  of this transmission scheme is bounded in the interval  $H_{\ell_{\max}} \leq H_G(\mathcal{R}^*) \leq H_0$ , where

$$H_{\ell_{\max}} = \log \binom{N}{E^{(\ell_{\max})}}. \quad (5)$$

The lower bound follows from always minimally needing to transmit the top layer of  $G$  as a representative layer, at a cost  $H_{\ell_{\max}}$ , and the upper bound follows from  $H_0 = H_G(\mathcal{R} = \mathcal{L})$  being in the solution space over which we minimize  $H_G$  to find  $\mathcal{R}^*$ . Therefore, exploiting layer overlap always provides compression relative to the naïve transmission of layers independently. The reducibility of the hypergraph  $G$  can then be computed based on the extent to which  $G$  can be compressed relative to the baseline cost of  $H_0$  bits.

Using these bounds, we can construct a properly normalized *structural reducibility* measure  $\eta$  for a hypergraph  $G$  as

$$\eta = \frac{H_0 - H_G(\mathcal{R}^*)}{H_0 - H_{\ell_{\max}}}, \quad (6)$$

which satisfies  $\eta \in [0, 1]$ . If  $G$  is maximally compressible—i.e., is a nested hypergraph where all layers  $G^{(\ell)}$  with  $\ell < \ell_{\max}$  are given by  $G^{(\ell_{\max} \rightarrow \ell)}$ —then we have  $\mathcal{R}^* = \{\ell_{\max}\}$  and  $H_G(\mathcal{R}^*) = H_{\ell_{\max}}$  and, thus,  $\eta = 1$ . On the other hand, we have an information cost  $H_G(\mathcal{R}^*) \approx H_0$  when  $G$  is highly incompressible (i.e., has little to no structural overlap among its layers), as there is little shared information that can be exploited to improve on the naïve information cost of  $H_0$ ; thus,  $\eta \approx 0$ .

In Supplemental Material [35], we discuss extending our reducibility concept to understand the structural redundancy of multiscale coarse-grainings of hypergraphs (Sec. I), as well as individual hypergraph layers (Sec. V) and individual hyperedges (Sec. VI).

By solving Eq. (4) to maximize compression of  $G$ , we can also obtain a compressed hypergraph representation for  $G$  given by  $G_{\text{red}} = \{G^{(r)}\}_{r \in \mathcal{R}^*}$ , which captures the critical higher-order structure of  $G$  while removing structurally redundant layers. In Fig. 1(a), we show an example hypergraph with layers  $\mathcal{L} = \{2, 3, 4, 5\}$  and its

corresponding reduced representation of layers  $\mathcal{R}^* = \{3, 5\}$ , giving a reducibility value of  $\eta = 0.34$ .

*Optimizing reducibility*—To identify the optimal representative layers  $\mathcal{R}^*$ , when  $L \lesssim 30$ —a value satisfied by most real hypergraph datasets [32]—we can use a simple brute-force search. Letting  $\ell_{\max} = \max(\mathcal{L})$  be the largest hyperedge size in  $G$ , we must have  $\ell_{\max} \in \mathcal{R}$ , since there is no higher layer from which to transmit  $G^{(\ell_{\max})}$ . Therefore, we have  $2^{L-1}$  possible subsets of representatives  $\mathcal{R}$  that we must search through. In our method we must first compute  $E^{(k \rightarrow \ell)}$ ,  $E^{(k \cap \ell)}$ , and  $E^{(\ell)}$  for all pairs of layers  $k, \ell$  with  $k > \ell$ . For layer pairs with small values  $k, \ell \lesssim 10$ , we can directly compute  $E^{(k \rightarrow \ell)}$  and  $E^{(k \cap \ell)}$  using the projection  $G^{(k \rightarrow \ell)}$  and the lower layer  $G^{(\ell)}$ . For larger  $k, \ell$  values, however, we cannot compute the projection  $G^{(k \rightarrow \ell)}$  directly. In this case, we can compute  $E^{(k \cap \ell)}$  by iterating through the edges  $e_i \in G^{(k)}$  in a fixed order, for each edge  $e_i$  checking its overlaps  $o(e_i) = \{e_i \cap e_\tau : \tau < i\}$  with all previously checked edges. Then, we can compute the number of new projected tuples that  $e_i$  contributes to  $E^{(k \rightarrow \ell)}$  as  $\binom{k}{\ell} - E^{[o(e_i) \rightarrow \ell]}$ , where  $E^{[o(e_i) \rightarrow \ell]}$  is the number of unique subtuples of size  $\ell$  within the set of overlapping tuples  $o(e_i)$ , which can be computed recursively using the same approach. When computing the projection  $G^{(k \rightarrow \ell)}$  is infeasible, we also must compute  $E^{(k \cap \ell)}$  in a more efficient way. We do this by iterating over hyperedges  $e_k \in G^{(k)}$  and incrementing  $E^{(k \cap \ell)}$  for each edge  $e_\ell \in G^{(\ell)}$  that fully overlaps with  $e_k$ , removing  $e_i$  from  $G^{(\ell)}$  afterward.

We can then compute a matrix  $M$  such that

$$M[k, \ell] = \log \left( \frac{E^{(k \rightarrow \ell)}}{E^{(k \cap \ell)}} \right) \left( \binom{N}{\ell} - E^{(k \rightarrow \ell)} \right) \left( E^{(\ell)} - E^{(k \cap \ell)} \right) \quad (7)$$

for layer pairs  $(k, \ell), k > \ell$ . The computation of certain conditional entropies  $M[k, \ell]$  is often the computational bottleneck in practice and takes roughly  $O[(E^{(k)})^2]$  run-time for large  $k, \ell$  where projection is intractable and  $O[\binom{k}{\ell} E^{(k)}]$  run-time when using direct projection. We also compute a vector  $Q$  with entries

$$Q[\ell] = \log \left( \frac{\binom{N}{\ell}}{E^{(\ell)}} \right), \quad (8)$$

storing the individual layer information costs for all layers  $\ell$ . Then, for each valid subset  $\mathcal{R} \subseteq \mathcal{L}$ , we can compute

$$H_G(\mathcal{R}) = \sum_{r \in \mathcal{R}} Q[r] + \sum_{\ell \in \mathcal{L} \setminus \mathcal{R}} M[r(\ell), \ell], \quad (9)$$

where  $r(\ell) = \arg\min_{r \in \mathcal{R}} \{M[r, \ell]\}$ . We then select  $\mathcal{R}^*$  as the representative subset that achieves the minimum value of  $H_G(\mathcal{R})$ . There are  $2^{L-1}$  subsets  $\mathcal{R}$  to check, and each takes in the worst case  $O(L)$  operations to compute  $r(\ell)$  for

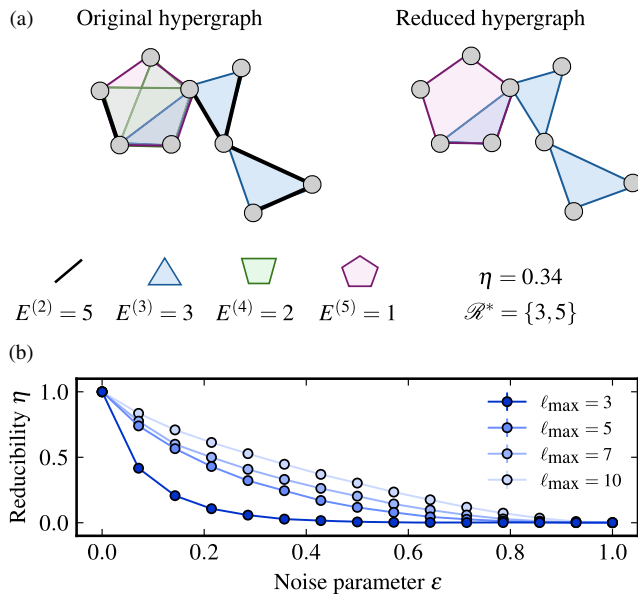


FIG. 1. Structural reducibility. (a) Hypergraph containing layers  $\mathcal{L} = \{2, 3, 4, 5\}$  of size  $E^{(\ell)}$ , which is reducible to an optimal representative layer set  $\mathcal{R}^* = \{3, 5\}$  with reducibility  $\eta = 0.34$  [Eq. (6)]. (b) Reducibility of a noisy nested hypergraph, with noise parameter  $\varepsilon$  determining the fraction of randomized hyperedges, for various hypergraph dimensions  $\ell_{\max}$ .

each layer  $\ell \in \mathcal{L} \setminus \mathcal{R}$ . The iteration over  $\mathcal{R}$  thus has a time complexity of roughly  $O(L^2 2^{L-1})$ , which in practice is tractable for most real-world datasets with  $L \lesssim 30$  (see Table I).

For systems with many layers  $\mathcal{L}$ , optimizing over subsets  $\mathcal{R} \subseteq \mathcal{L}$  through direct enumeration is unfeasible. In such cases, we use an approximate greedy method for identifying  $\mathcal{R}^*$  to compute the structural reducibility. Starting with  $\mathcal{R} = \{\ell_{\max}\}$ , we can iteratively add the best layer  $\ell$  to  $\mathcal{R}$  following the rule  $\ell = \operatorname{argmin}_{\ell \in \mathcal{L} \setminus \mathcal{R}} \{H_G(\{\ell\} \cup \mathcal{R})\}$  until  $\mathcal{R} = \mathcal{L}$ . Then, we choose among the explored solution candidates to find the representative layer set giving the lowest information cost  $H_G$ . In practice, this approximation always obtains accurate results (see Sec. IV in Supplemental Material [35]), significantly speeding up the computation and making our method available to otherwise intractable datasets.

**Reducibility of synthetic hypergraphs**—To validate our approach, we investigate our method on synthetic hypergraphs with tunable structure. First, we consider nested hypergraphs, where all lower-order interactions are fully encapsulated into those of higher order. Noisy nested hypergraphs are nested hypergraph where a noise parameter  $\varepsilon$  determines the fraction of its hyperedges to be rewired, replacing each selected hyperedge with a hyperedge of the same order drawn uniformly at random. In Fig. 1(b), we plot the results of this test, averaged over ten realizations of the randomness for each value of  $\varepsilon$ . Standard errors (vanishingly small) are shown as error bars. We can observe that  $\eta = 1$  indicates complete reducibility when  $\varepsilon = 0$  for each fully nested hypergraph and that  $\eta$  decreases smoothly as the hypergraph becomes noisier, eventually bottoming out at  $\eta = 0$  for purely random hypergraphs ( $\varepsilon = 1$ ). As the system gets larger ( $\ell_{\max}$  increases), we observe greater reducibility values, since a larger fraction of the layers are structurally redundant due to being nested within the top layer  $\ell_{\max}$ .

In a followup experiment, we examine the reducibility of more general synthetic hypergraphs with nested structure. We start by generating three planted representative layers  $G^{(3)}$ ,  $G^{(5)}$ , and  $G^{(7)}$  on  $N = 100$  nodes with  $E^{(3)}$ ,  $E^{(5)}$ ,  $E^{(7)} = 1740, 1050$ , and 50 hyperedges, respectively, drawn uniformly at random without replacement. We then generate the layers  $G^{(2)}$ ,  $G^{(4)}$ , and  $G^{(6)}$  as noisy versions of the projected layers  $G^{(3 \rightarrow 2)}$ ,  $G^{(5 \rightarrow 4)}$ , and  $G^{(7 \rightarrow 6)}$ , respectively, by selecting a fraction  $\varepsilon$  of the hyperedges in each projected layer randomly and replacing these hyperedges with those of the same size drawn uniformly at random. In Fig. 2, we plot layer-layer similarity matrices showing the network normalized mutual information [33] (over arbitrary tuple sizes) for pairs of layers in the generated hypergraphs, illustrating the effect of  $\varepsilon$  on the nested structure. As  $\varepsilon$  increases, we see a smooth decrease in the structural

reducibility  $\eta$ , with our method able to infer the planted set of representative layers  $\mathcal{R}^*$ . We compare our results with dynamical reducibility [32], which reduces structure based on the collective behavior supported by the hypergraph rather than topological overlap and which remains close to zero without detecting any change in this simple but nuanced hypergraph structure. All results are averaged over 20 realizations of the randomness at each  $\varepsilon$ . Section II in Supplemental Material [35] further investigates the reducibility of hypergraphs with tunable nested interactions.

We also examine our proposed multiscale reducibility measure in a similar experimental setting. For an  $N = 10^4$  node system, we synthesize each random hypergraph using a planted community partition  $\mathbf{b}$  by generating each of the  $E^{(\ell)}$  hyperedges in layer  $\ell = 2, \dots, 10$  as follows: (i) choose a random node  $i$  to start a hyperedge  $e$ ; (ii) add  $\ell - 1$  nodes to  $e$ , drawing each node (without replacement) from the same community as  $i$  with probability  $1 - p$  and from a different community with probability  $p$ . The result is a hypergraph that is more tightly clustered under the planted node partition  $\mathbf{b}$  as  $p \rightarrow 0$  and uncorrelated with  $\mathbf{b}$  as  $p \rightarrow 1$ . All results are averaged over ten realizations of the randomness at each  $p$ , and separate experiments are run for  $B = \{50, 200, 1000\}$  equally sized communities in  $\mathbf{b}$ . We observe that the standard reducibility of Eq. (6) cannot

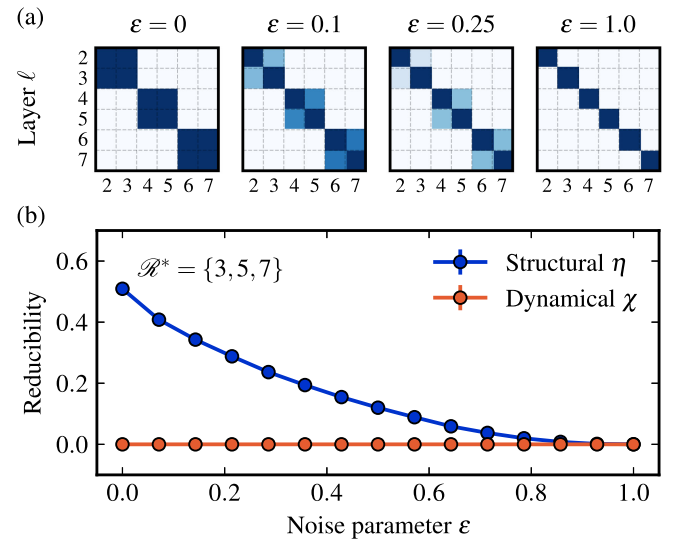


FIG. 2. Comparison of structural and dynamical  $\chi$  [32] reducibility measures. (a) Pairwise layer similarity matrices of block-nested hypergraphs at increasing levels of noise  $\varepsilon$ . Hypergraph layers  $\ell = 2, 4$ , and 6 are fully nested within  $\ell = 3, 5$ , and 7, respectively, and the similarity is lost as  $\varepsilon$  increases. (b) Structural and dynamical  $\chi$  reducibility measures against all  $\varepsilon$  values. The dynamical  $\chi$  reducibility does not detect any compressibility between layers. The structural reducibility uncovers both the structural redundancies and the planted, optimal representative layers  $\mathcal{R}^* = \{3, 5, 7\}$ .



TABLE I. Structural reducibility of empirical datasets. Daggers denote the usage of greedy minimization for obtaining  $\eta$  and  $\mathcal{R}^*$ . The greedy scheme produced identical results to the exact scheme for all networks with  $\ell_{\max} \lesssim 30$ .

Dataset	$N$	$E$	$\ell_{\max}$	$\mathcal{R}^*$	$\eta$
Coauth-mag-geology_1980	1674	903	18	{3, 5, 9, 18}	0.03
Coauth-mag-geology_1981	1075	547	29	{4, 5, 8, 29}	0.01
Coauth-mag-geology_1982	1878	987	26	{4, 6, 26}	0.02
Coauth-mag-geology_1983	1734	883	36	{4, 36}	0.03 <sup>†</sup>
Kaggle-whats-cooking	6714	39224	65	{6,8,9,11,65}	0.04 <sup>†</sup>
Contact-high-school	327	7818	5	{3, 5}	0.13
Contact-primary-school	242	12704	5	{4, 5}	0.09
Hospital-lyon	75	1824	5	{4, 5}	0.11
Hypertext-conference	113	2434	6	{3, 5, 6}	0.06
Invs13	92	787	4	{3, 4}	0.05
Invs15	217	4909	4	{3, 4}	0.10
Science-gallery	410	3350	5	{3, 5}	0.16
Sfh-conference	403	10541	9	{4,9}	0.10
Malawi-village	84	431	4	{3, 4}	0.19
Dawn	2290	138742	16	{6,7,13,16}	0.15
Ndc-classes	628	796	39	{4, 9, 10, 13, 14,23,27,39}	0.40 <sup>†</sup>
Ndc-substances	3414	6471	187	{5, 40, 187}	0.31 <sup>†</sup>
Email-enron	143	1459	37	{4,6,11,12,37}	0.17 <sup>†</sup>
Email-eu	986	24520	40	{5,7,8,9,10,11, 12,13,27,39,40}	0.19 <sup>†</sup>
Tags-ask-ubuntu	3021	145053	5	{5}	0.17
Tags-math-sx	1627	169259	5	{5}	0.26

detect any changes in the mesoscale nested structure, while the multiscale reducibility of Eq. (S3) in Supplemental Material [35] exhibits a smooth descent as  $p$  increases. For larger  $B$ , we see that the communities in  $\mathbf{b}$  become smaller and the hypergraph becomes less reducible under the coarse-graining  $\mathbf{b}$ , approaching the standard structural reducibility value.

*Reducibility of real networks*—Finally, we apply our reducibility method to a range of real higher-order networks [34], as shown in Table I. We find a great variety in the reducibility of these systems, with many systems most parsimoniously represented by only a small subset  $\mathcal{R}^*$  of their layers. In Supplemental Material [35], we further investigate the nested organization of real-world systems in Sec. III, while in Sec. IV we compare run times of the exact and greedy optimization methods for finding  $\mathcal{R}^*$ , showing that the greedy scheme is considerably faster especially for larger hypergraphs. Finally, we explore the structural and dynamical properties of all reduced empirical hypergraphs in Sec. VII. We find that the reduced hypergraph representations consistently preserve the global, mesoscale, and local connectivity of the complete empirical hypergraphs—as quantified by the effective number of connected components, community structure, and degree ordering, respectively. We also find that these reduced systems preserve the consensus times in higher-order voter model dynamics [36]. In general, we observe that such properties are better preserved as the reducibility  $\eta$  increases, due to improved compressibility of the original hypergraph structure.

*Conclusion*—Reducing the dimensionality of higher-order systems allows for more efficient analyses, with simpler interpretations and visualization. Here, we have developed a principled, efficient, and interpretable information-theoretic framework for assessing the structural reducibility of hypergraphs and removing structural redundancies to construct compressed hypergraph representations retaining the critical higher-order structure of complex networked systems. There are a number of ways in which this framework can be extended in future work to directed, weighted, temporal, or multilayer hypergraphs. This would allow the method to be applied to representations that capture additional nuances of the relational structure in a wider variety of systems. Our Letter sheds new light on the organizational principles of higher-order networks, distinguishing the extent to which lower-order information is redundant in the presence of higher-order information.

*Acknowledgments*—A. K. acknowledges support from the National Science Foundation of China (NSFC) through Young Scientist Fund Project No. 12405044. F. B. acknowledges support from the Austrian Science Fund (FWF) through Projects No. 10.55776/PAT1052824 and No. 10.55776/PAT1652425.

*Data availability*—The data that support the findings of this Letter are openly available [34]. Code is available as part of the library Hypergraphx [42].

- [1] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, *Phys. Rep.* **874**, 1 (2020).
- [2] F. Battiston, E. Amico, A. Barrat, G. Bianconi, G. Ferraz de Arruda, B. Franceschiello, I. Iacopini, S. Kéfi, V. Latora, Y. Moreno *et al.*, *Nat. Phys.* **17**, 1093 (2021).
- [3] G. Bianconi, *Higher-Order Networks* (Cambridge University Press, Cambridge, England, 2021).
- [4] C. Bick, E. Gross, H. A. Harrington, and M. T. Schaub, *SIAM Rev.* **65**, 686 (2023).
- [5] S. Majhi, M. Perc, and D. Ghosh, *J. R. Soc. Interface* **19**, 20220043 (2022).
- [6] C. Berge, *Hypergraphs: Combinatorics of Finite Sets* (Elsevier, New York, 1984), Vol. 45.
- [7] A. R. Benson, R. Abebe, M. T. Schaub, A. Jadbabaie, and J. Kleinberg, *Proc. Natl. Acad. Sci. U.S.A.* **115**, E11221 (2018).
- [8] G. Petri and A. Barrat, *Phys. Rev. Lett.* **121**, 228301 (2018).
- [9] M. Contisciani, F. Battiston, and C. De Bacco, *Nat. Commun.* **13**, 7229 (2022).
- [10] L. Di Gaetano, F. Battiston, and M. Starnini, *Phys. Rev. Lett.* **132**, 037401 (2024).
- [11] I. Iacopini, G. Petri, A. Barrat, and V. Latora, *Nat. Commun.* **10**, 2485 (2019).
- [12] G. Burgio, S. Gómez, and A. Arenas, *Phys. Rev. Lett.* **132**, 077401 (2024).
- [13] G. Ferraz de Arruda, A. Aleta, and Y. Moreno, *Nat. Rev. Phys.* **6**, 468 (2024).
- [14] L. Di Gaetano, G. Carugno, F. Battiston, and F. Coghi, *Phys. Rev. Lett.* **133**, 107401 (2024).
- [15] P. S. Skardal and A. Arenas, *Phys. Rev. Lett.* **122**, 248301 (2019).
- [16] A. P. Millán, J. J. Torres, and G. Bianconi, *Phys. Rev. Lett.* **124**, 218301 (2020).
- [17] Y. Zhang, M. Lucas, and F. Battiston, *Nat. Commun.* **14**, 1605 (2023).
- [18] M. S. Anwar, G. K. Sar, M. Perc, and D. Ghosh, *Communications in Physics* **7**, 59 (2024).
- [19] S. Majhi, S. Ghosh, P. K. Pal, S. Pal, T. K. Pal, D. Ghosh, J. Završnik, and M. Perc, *Phys. Life Rev.* **52**, 144 (2025).
- [20] U. Alvarez-Rodriguez, F. Battiston, G. F. de Arruda, Y. Moreno, M. Perc, and V. Latora, *Nat. Hum. Behav.* **5**, 586 (2021).
- [21] A. Civilini, O. Sadekar, F. Battiston, J. Gómez-Gardeñes, and V. Latora, *Phys. Rev. Lett.* **132**, 167401 (2024).
- [22] A. Kumar, S. Chowdhary, V. Capraro, and M. Perc, *Phys. Rev. E* **104**, 054308 (2021).
- [23] Q. F. Lotito, F. Musciotto, A. Montresor, and F. Battiston, *Commun. Phys.* **5**, 79 (2022).
- [24] T. LaRock and R. Lambiotte, *J. Phys. Complex.* **4**, 045007 (2023).
- [25] N. W. Landry, J.-G. Young, and N. Eikmeier, *EPJ Data Sci.* **13**, 17 (2024).
- [26] L. Gallo, L. Lacasa, V. Latora, and F. Battiston, *Nat. Commun.* **15**, 4754 (2024).
- [27] M. De Domenico, V. Nicosia, A. Arenas, and V. Latora, *Nat. Commun.* **6**, 6864 (2015).
- [28] M. De Domenico and J. Biamonte, *Phys. Rev. X* **6**, 041062 (2016).
- [29] A. Santoro and V. Nicosia, *Phys. Rev. X* **10**, 021069 (2020).
- [30] A. Kirkley, A. Rojas, M. Rosvall, and J.-G. Young, *Commun. Phys.* **6**, 148 (2023).
- [31] A. Kirkley and M. Newman, *Commun. Phys.* **5**, 40 (2022).
- [32] M. Lucas, L. Gallo, A. Ghavasieh, F. Battiston, and M. De Domenico, *arXiv:2404.08547*.
- [33] H. Felipe, F. Battiston, and A. Kirkley, *Commun. Phys.* **7**, 335 (2024).
- [34] N. W. Landry, M. Lucas, I. Iacopini, G. Petri, A. Schwarze, A. Patania, and L. Torres, *J. Open Source Software* **8**, 5162 (2023).
- [35] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/xrn7-cz8v> including Refs. [36–41].
- [36] J. Kim, D.-S. Lee, B. Min, M. A. Porter, M. San Miguel, and K.-I. Goh, *Phys. Rev. E* **111**, L052301 (2025).
- [37] M. A. Riolo, G. T. Cantwell, G. Reinert, and M. E. J. Newman, *Phys. Rev. E* **96**, 032310 (2017).
- [38] S. Morel-Balbi and A. Kirkley, *Phys. Rev. Res.* **6**, 033307 (2024).
- [39] N. X. Vinh, J. Epps, and J. Bailey, *Proceedings of the 26th Annual International Conference on Machine Learning* (2009), pp. 1073–1080, [10.1145/1553374.1553511](https://doi.org/10.1145/1553374.1553511).
- [40] A. Eriksson, D. Edler, A. Rojas, M. de Domenico, and M. Rosvall, *Commun. Phys.* **4**, 133 (2021).
- [41] V. Sood, T. Antal, and S. Redner, *Phys. Rev. E* **77**, 041121 (2008).
- [42] Q. F. Lotito, M. Contisciani, C. De Bacco, L. Di Gaetano, L. Gallo, A. Montresor, F. Musciotto, N. Ruggeri, and F. Battiston, *J. Complex Networks* **11**, cnad019 (2023).