Emerging Models and Paradigms in Network Science

Part #4: Higher-Order Interactions

Ph.D. Course at University of Calabria
27–30 May 2024

Francesco Cauteruccio

DIEM, University of Salerno

fcauteruccio@unisa.it

francescocauteruccio.info

Introduction

- Over the past decades, a variety of complex systems has been successfully described as networks whose interacting pairs of nodes are connected by links.
- Yet, complex systems exhibit interactions that often occur in groups of three or more nodes:
  - These cannot be described simply in terms of dyads, e.g., pairwise connections.
- Such interactions make the *higher-order architecture* of complex systems.
- Taking the higher-order structure of these systems into account can:
  - **Enhance our modeling capacities**, and
  - **Help us understand and predict dynamical behavior**.
- As for network science, a novel context is emerging: *networks beyond pairwise interactions* [4].

Complex Systems

- If we take an ecosystem and break it into pieces, chances are that our understanding of population dynamics will be slim at best:
  - Explain epileptic seizures starting from the individual neurons of the human brain;
  - Viral rumors spreading across societies from individual human psychology;
- What is missing?
  - **The rich pattern of nonlinear interactions between the system components.**
- Reductionism: the collective behaviors of a complex system can be simply understood and predicted by considering the units of the system in isolation [2].
  - Nowadays, this idea has been completely abandoned.
- Networks emerged as a reference modeling tool for complex systems.

Network Science is Not Enough

- Networks = collection of nodes (the elementary units of the systems) + edges (existence of interactions between pairs of such units).
- Applications to real-world systems are only possible if we describe interactions in detailed and varied ways:
  - Message networks require a *direction*,
  - Flow networks require a *weight*,
  - Contact networks require a *time*.
- However, a reasonable question arises:
  - **Are networks themselves enough to provide a complete description of complex systems?**

… Probably Not!

- An important limitation of networks is that **they capture pairwise interactions exclusively!**
- Yet, many empirical systems seem to display group interactions:
    - Social systems [6], neuroscience [11], ecology [8], biology [12], etc.
    - Neuronal dynamics display mesoscopic behaviors that require interactions among multiple neurons to be predicted [7].
    - Three or more species routinely compete for food and territory [9].
    - Social mechanisms, such as peer-pressure or collaborations, inherently go beyond the idea of dyadic connections [5].
- Different approaches leveraged the language of pairwise networks to describe interactions of *higher-order*:
    - Bipartite graphs [10]
- **Is there a mathematical framework that can explicitly and naturally describe group interactions?**
    - **Simplicial complexes** and **hypergraphs**

Low- and High-Order Representations

- An *interaction set* is a set $I = \{p_0, p_1, \dots, p_{k-1}\}$ containing an arbitrary number k of basic elements of the system under study.
  - We indicate them as *nodes* or *vertices*.
- Interactions can be classified by the number of entities they involve:
- We denote the *order* (or *dimension*) of an interaction involving k nodes to be $k - 1$:
  - A node interaction with itself only is a 0-order interaction.
  - An interaction between two nodes is a 1-order interaction.
  - An interaction between three nodes is a 2-order interaction, and so on.
- We consider *higher-order interactions* to be k-interactions with $k \geqslant 2$.
  - Low-order interactions are those with $k \leqslant 1$.
- *Higher-order systems* are those systems displaying interactions in groups of more than two elements.

Interacting System

- We define an *interacting system* $(V, \mathcal{I})$ as the family of interactions $\mathcal{I} = \{I_0, \ldots, I_n\}$ taking place on a node set V.
- Example:
  - $V = \{a, b, c, d, e\}, \mathcal{I} = \{[a, b, c], [a, d], [d, c], [c, e]\}$.
  - $\mathcal{I}$ contains three 1-interactions and one 2-interaction.
- While the complete information about the systems is given in $\mathcal{I}$, the study of most interesting properties of the system requires the choice of a **representation**.

Graph-based Representation I

- A simple graph $G = (V, E)$.
- The most natural choice to represent $\mathcal{I}$ is to unfold each higher-order interaction in $\mathcal{I}$ in terms of 1-interactions built from pairs of nodes in I.
  - $\mathcal{I}_G = \{[a, b], [b, c], [c, a], [a, d], [d, c], [c, e]\}$.
- Despite the power of graph representations, it is impossible to explicitly describe group interactions.
  - For instance, it is impossible to tell from $\mathcal{I}_G$ whether the original interaction set contained $[a, c, d]$ or not.

Graph-based Representation II

- A bipartite graph is defined by two node sets (U, W) and an edge set E containing only edges (u, w) such that $u \in U$ and $w \in W$.
  - A bipartite graph is generally used to represent higher-order interactions in the following way:
    - $U = V$, $W = \mathcal{I}$, and the links in E connect a node (in V) to an interaction (of arbitrary order in W) in which it is involved.
- While this representation is useful, the nodes of the original system do not interact directly with each other anymore.
  - Their relation is always mediated by the interaction layer (W), which is of a different nature from the node layer itself.
- One could argue that *cliques* are a convenient way to analyze higher-order interactions:
  - A clique of size k is defined as a fully connected subgraph of k nodes.
  - In our example, both sets a, b, c and a, d, c form 3-cliques.
  - Conversely, in $\mathcal{I}_G$ we only had a true 2-interaction in $[a, b, c]$, while the fictitious one $[a, d, c]$ is emerging as a byproduct of $[a, d] \cup [d, c]$ with the $[a, c]$ edge induced by $[a, b, c]$.
  - Thus, if we consider all cliques present at the graph level, we would find a 2-interaction that was not included in the original interaction set!

Higher-order Representations I

- Why not encode interactions exactly as they are?
- *Simplices*
  - A k-simplex σ is a set of $k+1$ nodes $\sigma = \{p_0, p_1, \ldots, p_k\}$.
  - Dimension of a simplex = order of an interaction.
- *Simplicial Complexes*
  - A simplicial complex is a collection of n simplices $K = \{\sigma_0, \sigma_1, \ldots, \sigma_n\}$ such that for every k-simplex $\sigma = \{p_0, p_1, \ldots, p_k\} \in K$, all its subfaces of any dimension belong to K too.
  - For instance, if $\{a, b, c\} \in K$, then $\{a\}, \{b\}, \{c\}, \{a, b\}, \{a, c\}, \{b, c\} \in K$ must hold.

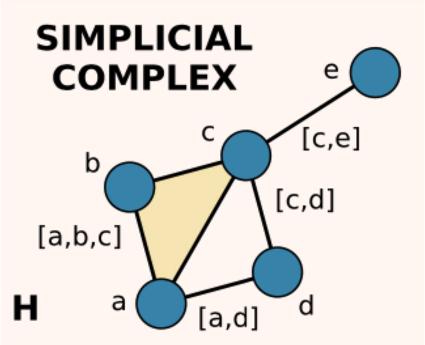

Figure: A simplicial complex. Source: [4]

Higher-order Representations I

- Simplicial complexes overcome some of the problems encountered with lower dimensional representations.
- However, they are limited by the requirement on the existence of all subfaces, which in some cases poses a too restrictive constraint:
  - Collaborations in scientific papers: a paper by three authors and none by the corresponding pairs of authors.
  - Gene pathways: exactly four genes are needed to perform a function, but the subgroups are not responsible for any function on their own.

Higher-order Representations II

- *Hypergraphs*
  - A hypergraph is defined by a node set V and a set of hyperedges H that specify which nodes participate in which way within an interaction.
    - Each hyperedge $h \in H$ is a non-empty subset of V, that is $h \subseteq V$ and $h \neq \emptyset$.
- Hypergraphs provide the most general and unconstrained description of higher-order interactions.
  - Note that a hypergraph can include a 2-interaction without any requirements on the existence of 1-interactions related to it.
  - It is also possible to define hyperedges that include other hyperedges.
  - Such flexibility comes with additional complexity in treating them.

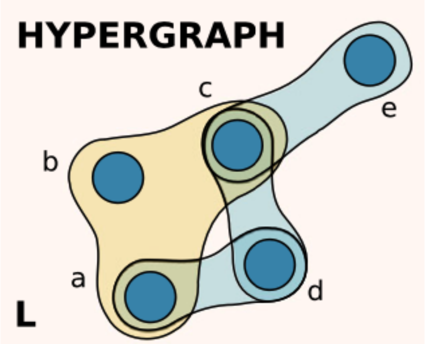

Figure: A hypergraph. Source: [4]

Measures and Structures

- In what follows, we mainly focus on hypergraphs. See [4] for a discussion on simplicial complexes.
  - Nevertheless, almost all concepts hold for simplicial complexes also.
- We'll take a look at different measures on hypergraphs:
  - Matrix representations and properties of higher-order systems,
  - Duality and line graphs,
  - Walks, paths, and centrality measures,
  - Clustering coefficient.

Matrix Representations: The Incidence Matrix

- The incidence matrix of a hypergraph $\mathcal{H} = (V, H)$ is a $n \times m$ matrix $I = \{I_{i\alpha}\}$ where $n = |V|$ and $m = |H|$.
  - The entry $I_{i\alpha}$ in row i and column α is 1 if node i is involved in hyperedge α, 0 otherwise.
  - In hypergraphs allowing for a node to be represented more than once in each hyperedge, it can be useful to weight the entries of the incidence matrix I.
- I can be useful to characterize various properties of a hypergraph, thus of a higher-order system.
- The *degree* of a node i in $\mathcal{H}$ is the sum of the elements of the ith-row of I, and we denote it as $d(i) = \sum_{\alpha \in H} I_{i\alpha}$.
  - $d(i)$ indicates the number of hyperedges containing the node i.
- The *size* of a hyperedge α in $\mathcal{H}$ is the number of nodes involved in it, that is $size(\alpha) = \sum_{i \in V} I_{i\alpha}$.

Matrix Representation: The Adjacency Matrix

- From I, we can construct another matrix that fully encodes the connectivity of the hypergraph, that is, the *adjacency matrix* A.
  - We define $A = II^T - D$, where D is the diagonal matrix whose diagonal entries are the number of hyperedges a node belongs to, that is its degree.
  - A is a $n \times n$ matrix whose elements a_{ij} for $i \neq j$ are the number of hyperedges that contain both i and j.
  - In the case of weighted hyperedges, we can write $A = IWI^T - D$, where I is the incidence matrix, W is the diagonal matrix with the weights of the hyperedges along the diagonal, and D is a diagonal matrix with the degrees of the nodes along the diagonal [14].
- Another useful structure is the *intersection profile* $P = I^T I$ which is a $m \times m$ matrix whose elements $P_{\alpha\beta}$ count the number of nodes in common between two hyperedges α and β, where m is the number of hyperedges.

Matrix Representation – Example

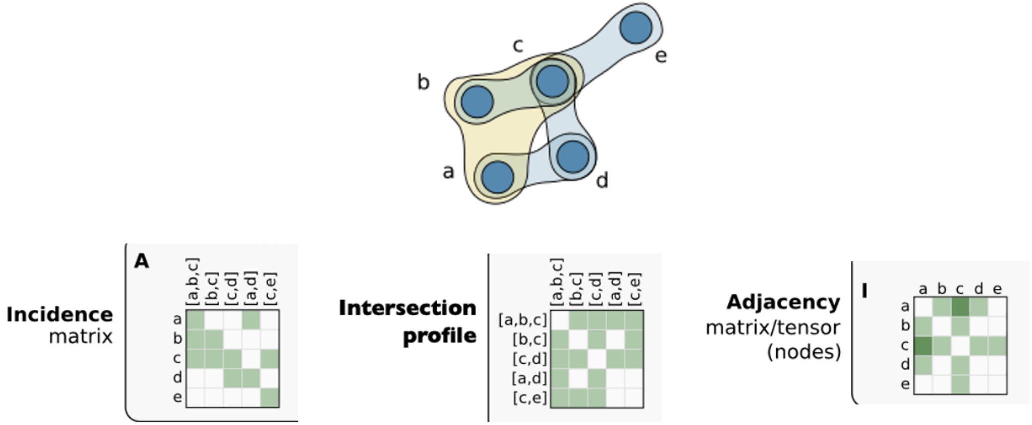

Figure: A hypergraph H and its matrix representations. Source [4]

Hypergraph Properties

- There are different ways to refer to common hypergraph properties:
- The *order* of a hypergraph is its number of nodes, i.e., $n = |V|$.
- The *size* of a hypergraph is its number of hyperedges, i.e., $m = |H|$.
- The *size* of a hyperedge $h \in H$ is its cardinality, i.e., $|h|$.
- The *degree* of a node i is the number of hyperedges containing it, i.e., $d(i) = |\{h \in H : i \in h\}|$.
    - In the literature, the degree is often called *hyperdegree*, while the former is defined as the number of neighbors of i, i.e., the number of nodes that share at least one hyperedge with i.
- A hypergraph is said to be k-uniform if all hyperedges have size k.

Dual Hypergraph

- Let $H = (V, E)$ be a hypergraph with node set $\{v_1, \ldots, v_n\}$ and family of edges $E = \{e_1, \ldots, e_m\}$.
- The *dual hypergraph* of H, denoted as $H^* = (E^*, V^*)$, has node set $E^* = \{e_1^*, \ldots, e_m^*\}$ and family of hyperedges $V^* = \{v_1^*, \ldots, v_n^*\}$, where $v_i^* = \{e_k^* : v_i \in e_k\}$.
    - Practically speaking, the dual of a hypergraph is the hypergraph constructed by swapping the roles of nodes and edges.
    - The dual of a hypergraph with incidence matrix I is the hypergraph associated with the transposed incidence matrix I^T.
    - $(H^*)^* = H$.

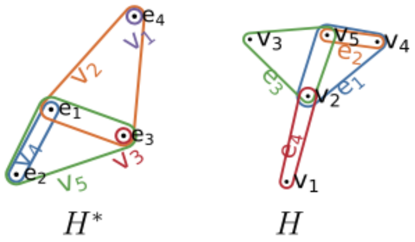

Figure: A hypergraph H and its dual H^*.
Source: [1]

Line Graph

- Let $H = (V, E)$ be a hypergraph with node set $\{v_1, \ldots, v_n\}$ and family of edges $E = \{e_1, \ldots, e_m\}$.
- The *line graph* of H, denoted as $L(H)$, is the graph on vertex set $\{e_1^*, \ldots, e_m^*\}$ and edge set $\{\{e_i^*, e_j^*\} : e_i \cap e_j \neq \emptyset \text{ for } i \neq j\}$.
  - Practically speaking, the line graph of a hypergraph is a graph having as nodes the original hyperedges, where a link between two nodes indicates that the two corresponding hyperedges share at least one node in the hypergraph.
- The literature reports different names for these objects:
  - $L(H)$ is the *line graph*, *representative graph*, or *intersection graph*,
  - $L(H)$ and $L(H^*)$ are often referred to as *top and bottom projections* of H,
  - $L(H^*)$ is the *2-section*, *clique graph*, or *clique expansion*.

Line Graph and Dual Hypergraph – Example

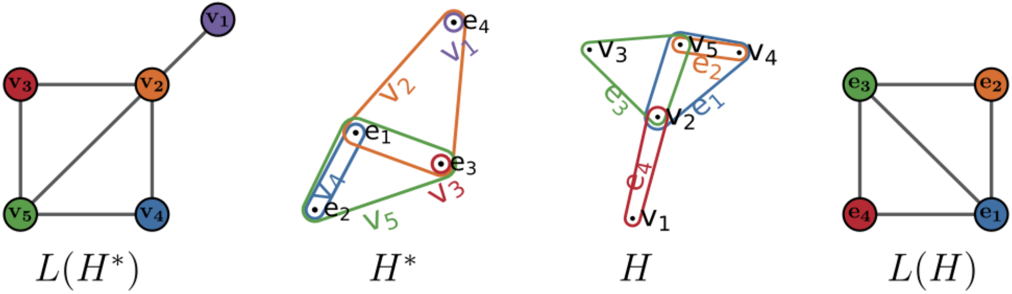

Figure: A hypergraph H, its dual H^*, the line graphs $L(H)$ and $L(H^*)$

From Graph Walks to Hypergraph Walks

- One of the most fundamental concepts in graph theory is that of a walk.
- For a graph $G = (V, E)$, a *walk of length k* is a sequence of nodes $v_0, v_1, \ldots, v_k$ such that each pair of successive vertices are adjacent.
  - Any valid graph walk can be equivalently expressed as either a sequence of adjacent vertices or a sequence of incident edges.
- Such observation no longer holds for hypergraphs.
  - Two hyperedges can intersect at any number of nodes, and two nodes can belong to any number of shared hyperedges.
- In the literature, two walk concepts have been defined:
  - Walks on the node level – consisting of successively adjacent nodes,
  - Walks on the hyperedge level – consisting of successively intersecting hyperedges.
- Both notions are captured when duality is considered:
  - A node-based walk on H is simply a hyperedge-based walk on H^*.

From Graph Walks to Hypergraph Walks

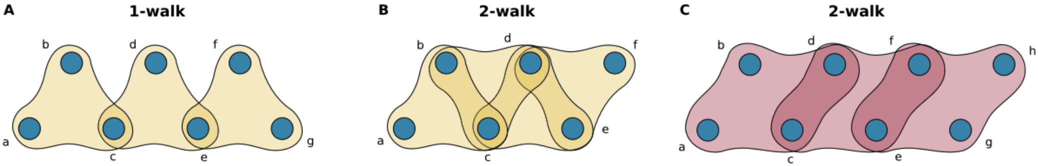

Figure: A 1-walk, and two 2-walks. Source: [4]

s-walk

- Given $s > 0$, an *s-walk* of length k between two hyperedges f and g is a sequence of hyperedges
$$f = e_{i_0}, e_{i_1}, \ldots, e_{i_k} = g$$
where for $j = 1, \ldots, k$, we have $s \leqslant |e_{i_{j-1}} \cap e_{i_j}|$ and $i_{j-1} \neq i_j$.
    - Notation from [1].
- Given this definition, a number of basic yet important properties, measures, and structures can be defined on hypergraphs:
    - s-connected components,
    - s-distance,
    - s-paths,
    - etc.
- In the literature, the framework under which these definitions apply is referred to as "the s-walk framework" [1].

Connected Components

- For a hypergraph $H = (V, E)$, a subset of hyperedges $C \subseteq E$ is called s-connected if there exists an s-walk between all $f, g \in C$, and is further called an s-connected component if there is no s-connected $J \subseteq E$ such that $C \subsetneq J$.
  - Note that the order of an s-connected component is bounded above by $|E_s|$, where $E_s = \{e \in E : |e| \geqslant s\}$.
- This concept is particularly useful in analyzing the structure and dynamics of hypergraphs.

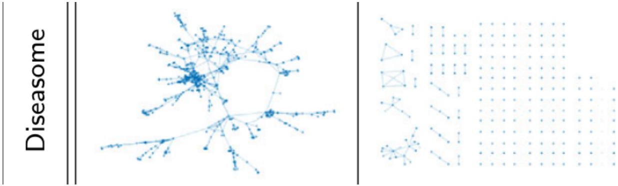

Figure: 1-components (left) and 2-components (right) within the dataset *Diseasome* in [1]

s-line Graph

- Let $H = (V, E)$ be a hypergraph with node set $V = \{v_1, \ldots, v_n\}$ and edge set $E \supset E_s$ where $E_s = \{e \in E : |e| \geqslant s\} = \{e_1, \ldots, e_k\}$, for an integer $s \geqslant 1$.
- The *s-line graph of H*, denoted by $L_s(H)$, is the graph on vertex set $\{e_1^*, e_k^*\}$ and edge set $\{\{e_i^*, e_j^*\} : |e_i \cap e_j| \geqslant s \text{ for } i \neq j\}$.
  - Each node in the s-line graph represents a hyperedge with at least s nodes in the hypergraph, and two nodes are linked in the s-line graph if their corresponding hyperedges intersect in at least s vertices in the hypergraph.

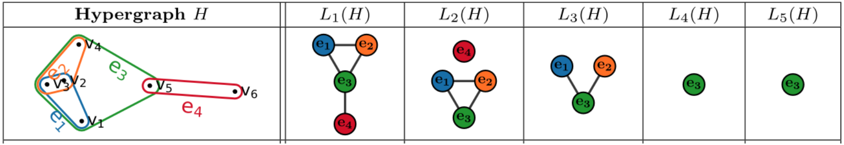

Figure: Line graphs for a hypergraph. Source: [1]

Distance and Centrality

- Let $H = (V, E)$ be a hypergraph and $E_s = \{e \in E : |e| \geqslant s\}$. We define the s-distance function $d_s : E_s \times E_s \to \mathbb{Z}_{\geqslant 0}$ by

$$d_s(f, g) = \begin{cases} \text{length of the shortest } s\text{-walk}, & \text{if an } s\text{-walk between } f, g \text{ exists;} \\ \infty, & \text{otherwise} \end{cases}$$

  - Note that (E_s, d_s) is a metric space [1].
- Measures based on distance:
  - The s-eccentricity of a hyperedge f is $\max_{g \in E_s} d_s(f, g)$.
  - The s-diameter is the maximum s-eccentricity over all edges in E_s, while the s-radius is the minimum.
  - The average s-distance of H is $\binom{|E_s|}{2}^{-1} \sum_{f, g \in E_s} d_s(f, g)$.
  - The s-closeness centrality of a hyperedge f is $\frac{|E_s| - 1}{\sum_{g \in E_s} d_s(f, g)}$.
- These measures are generally computed for the largest s-component of the hypergraph.
  - These can also be defined on a per-component basis [1].

Paths, Cycles, and Clustering Coefficients I

- For a hypergraph $H = (V, E)$, let the sequence of hyperedges $\omega = (e_{i_0}, e_{i_1}, \ldots, e_{i_k})$ be an s-walk of length k, and let $I_j = e_{i_{j-1}} \cap e_{i_j}$ be the j-th intersection.
- The s-walk can be also defined as:
  1. An s-trace if $i_x \neq i_y$ for all $x \neq y$ (all hyperedges are pairwise distinct by label);
  2. An s-meander if ω is an s-trace in which $I_x \neq I_y$ for all $x \neq y$ (all intersections are pairwise distinct);
  3. An s-path if ω is an s-meander in which $I_x \setminus I_y \neq \emptyset$ for all $x \neq y$ (no intersection is included in another).
- This categorization can help in defining high-order substructures or motifs that cannot be determined from the s-line graph alone.

Paths, Cycles, and Clustering Coefficients II

- In some cases, we need these definitions to be *closed*, i.e., $e_{i_0} = e_{i_k}$.
- For a hypergraph H, an s-triangle is a closed s-path of length 3, and an s-wedge is an s-path of length 2. For an s-wedge e_0, f, e_2, we say f is the center of the s-wedge.
    - The s-local clustering coefficient of a hyperedge $f \in E_s$ is given by:

$$s\text{-LCC}(f) = \begin{cases} \frac{\text{number of } s\text{-triangles containing } f}{\text{number of } s\text{-wedges centered at } f} & \text{if } f \text{ is the center of an } s\text{-wedge;} \\ 0 & \text{otherwise} \end{cases}$$

    - The s-global clustering coefficient of a hypergraph H is given by:

$$s\text{-GCC}(H) = \frac{3 \cdot \text{total number of } s\text{-triangles}}{\text{total number of } s\text{-wedges}}$$

Applications

- Hypernetwork Science allows us to study and investigate novel and old concepts in complex systems exhibiting higher-order interactions.
- A plethora of work related to hypergraphs and hypernetworks is starting to appear in the literature.
- We now discuss two recent works on hypernetwork science:
  - **Segregation measures on hypernetworks**
    - Failla, A., Rossetti, G. and Cauteruccio, F. (2024) Beyond Boundaries: Capturing Social Segregation on Hypernetworks. *In peer-review, available from authors.*
  - **Approaches for the hypergraph Influence Maximization problem**
    - Auletta, V., Cauteruccio, F. and Ferraioli, D. (2024) Heuristics for the Influence Maximization Problem on Hypergraphs. *In peer-review, available from authors.*

Segregation measures on hypernetworks I

- Segregation is generally defined as the extent to which system entities are separated or clustered based on certain attributes or features.
  - Understanding segregation is crucial for insights into system structure and dynamics, with implications for social equity, cohesion, resilience, and functionality.
- Current research lacks studies on segregation in hypernetworks due to higher-order interaction complexity.
- This work aims to:
  - Define a framework for extending pairwise segregation measures to hypernetworks.
  - Introduce *Random Walk HyperSegregation* (RWHS), a novel measure for hypernetwork segregation.
  - Propose two RWHS variants: *(i)* meet-wise and *(ii)* jump-wise RWHS.
- Experiments on synthetic and real-world hypernetwork topologies.

Segregation measures on hypernetworks II

- We deal with node-attributed hypergraphs:
  - Let L be a set of node labels (e.g., political group),
  - A node-attributed hypergraph $H_L = (V, E, L)$ is a hypergraph H where to each node is assigned a label (a group) $l \in L$.
- We propose a *segregation measure schema*, i.e., a general framework to extend classical segregation measures to hypernetworks, defined as

$$\mathfrak{F} = \langle H_L, f^{ie}(\cdot), \rho(\cdot) \rangle$$

where
  - H_L is a node-attributed hypergraph,
  - $f^{ie} : E \to [0, 1]$ is a generalized hyperedge type function,
  - $\rho : H_L \to [-1, 1]$ is a generalized segregation measure.

Segregation measures on hypernetworks III

- $f^{ie}(e_i)$ indicates whether the hyperedge $e_i \in E$ is *internal* or *external* w.r.t. a provided definition
  - This depends on the context, and it is based on the labels of the contained nodes.
  - For instance, we define e_i as internal only if all its nodes belong to the same group, i.e., have the same label.
- $\rho(H_L)$ indicates how segregated the hypernetwork is; we extend it with two classical measures
  - $\rho_{E-I}(H_L)$ is based on the classical E-I index, which calculates the ratio of the net difference between inter-group (external) and intra-group (internal) ties to the overall number of ties.
  - $\rho_Q(H_L)$ is based on Gupta's Q index, which was introduced to analyze the effects of mixing patterns of sexual contacts on the spread of the HIV epidemic.

Segregation measures on hypernetworks IV

- We then propose a random walker model to leverage random walks on the hypernetworks as effective proxies for information flow.
  - The aim is to analyze walkers' behaviors to infer whether segregation occurs.
- We deal with a collection of t random walks starting from a node v_i of length k, and we denote it as $W_i^{t,k}$.
- We define a segregation measure called Random Walk HyperSegregation (RWHS), which is based on the collection $W_i^{t,k}$ and provides two variants:
  - Meet-wise RWHS – the ratio of nodes in the same group as v_i that appears in the collection of random walks starting from it.
  - Jump-wise RWHS - it counts the pairs of subsequent steps (in the random walks) whose nodes belong to the same group.

Segregation measures on hypernetworks V

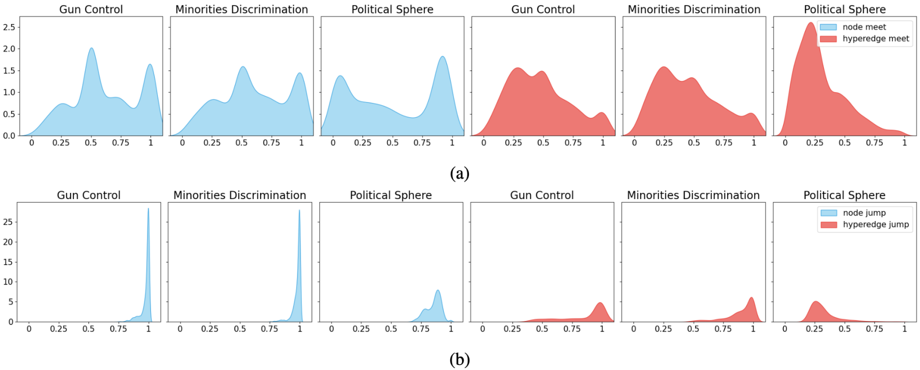

Figure: Meet-wise (a) and Jump-wise (b) RWHS score distribution on the *Reddit Politics* dataset. The *y*-axis represents the density estimate. The RWHS perfectly captures the general tendency for higher segregation than expected at random.

IM on Hypergraphs

- The Influence Maximization (IM) problem has been widely analyzed in graph topologies.
- Still very limited interest in hypergraph topologies.
- We propose two families of approaches to tackle the IM problem on hypergraphs:
  - SMARTPROPS – a general algorithm that exploits the centrality values of nodes in order to select the best seeds.
  - HC and ES – two metaheuristics based on *hill-climbing* and *evolution strategy*, respectively.

IM on Hypergraphs - The IM problem

- Let us define a hypergraph $H = (V, E)$, a value $k \in \mathbb{Z}_{>0}$, and a diffusion process model on hypergraph σ_H
- The IM problem consists in finding a subset $S^* \subseteq V, |S| = k$, such that the expected number of infected nodes is maximized.
  - Formally,
  $$S^* = argmax_{S \subset V, |S|=k} \sigma_H(S)$$
  where $\sigma_H(S)$ indicates the number of reached nodes at the end of the diffusion process starting from nodes in S.
  - S is also called the *seed node set*.
- Diffusion process models are theoretical frameworks designed to simulate how information spreads through a network.
  - Well-known ones are Independent Cascade, Linear Threshold, etc [3].

IM on Hypergraphs - The IM problem

- We use the Susceptible-Infected (SI) model with Contact Process (CP) dynamics on hypergraphs [13].
  - In SICP, a node can be either in a susceptible (S) or infected (I) state.
  - An S-state node can be infected by each of its neighbors in the I-state with a given infection rate β.
- The diffusion works as follows:
  1. Nodes in S are set to be infected (I-state), and others are set as S-state.
  2. At each time step t, we find the I-state nodes.
     - For each I-state node v_i, we find the set E_i of all hyperedges E_i containing v_i.
     - A hyperedge e_j is chosen from E_i uniformly at random, and each of the S-state nodes in e_j will be infected by v_i with probability β.
  3. The diffusion terminates after T steps, and $\sigma_H(S)$ is the number of nodes in I-state.

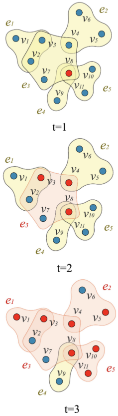

Figure: An example of diffusion in the SICP model. Source: [13]

IM on Hypergraphs - SMARTPROPS

- SMARTPROPS is a general algorithm to identify an optimal seed node set.
- The idea is to evaluate nodes w.r.t. a given property ϕ, and select the top-k ones to define the seed node set.
- We propose four variants of SMARTPROPS, each one based on a specific property:
  - SMARTDEG – ϕ is the degree of each node in H.
  - SMARTHYPERDEG – ϕ is the hyperdegree of each node in H.
  - SMARTSHAPDEG – ϕ is the Shapley Degree value of each node in H computed on the line graph $L(H)$.
  - SMARTSHAPCLOSE – ϕ is the Shapley Closeness value of each node in H computed on the line graph $L(H)$.

Algorithm 1 Pseudocode of the SMARTPROPS algorithm

function SMARTPROPS(H, k, ϕ, ρ)
    Input: hypergraph $H(V, E)$, $k \in \mathbb{Z}_{>0}$, node property function $\phi : V \to \mathbb{R}$, node threshold function $\rho : V \to \mathbb{R}$
    Output: $R^* \subseteq V$
    $snodes \leftarrow$ SORTEDDESC(V, ϕ)
    $R^* \leftarrow$ QUEUE$()$
    PUSH$(R^*,$ GET$(snodes, 1))$
    $i \leftarrow 2$
    while SIZE$(R^*) < k$ **do**
        $v_i \leftarrow$ GET$(snodes, i)$
        $th_i \leftarrow \rho(v_i)$
        $c_i \leftarrow 0$
        for $e_j \in E_i$ **do**
            $c \leftarrow |\{u_q \in R^* : u_q \in e_j\}|$
            if $c > 0$ **then** $c_i \leftarrow c_i + 1/c$
        if $c_i \geq th_i$ **then** PUSH$(R^*,$ GET$(snodes, i))$
        $i \leftarrow i + 1$
    return R^*

Figure: Pseudocode of SMARTPROPS

IM on Hypergraphs - Experiments

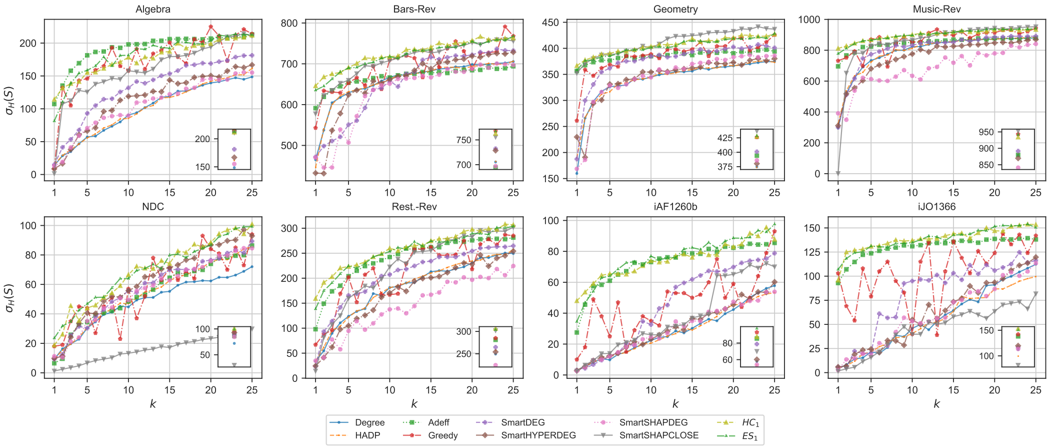

Figure: Values of $\sigma_H(S)$ averaged over 100 runs. SICP $\beta = 0.01$ and $T = 25$. Datasets available from [13]

Thanks for your attention!

Francesco Cauteruccio, Ph.D.
University of Salerno

fcauteruccio@unisa.it
https://francescocauteruccio.info

References I

Sinan G Aksoy, Cliff Joslyn, Carlos Ortiz Marrero, Brenda Praggastis, and Emilie Purvine.
Hypernetwork science via high-order hypergraph walks.
EPJ Data Science, 9(1):16, 2020.

Philip W Anderson.
More is different: Broken symmetry and the nature of the hierarchical structure of science.
Science, 177(4047):393–396, 1972.

Suman Banerjee, Mamata Jenamani, and Dilip Kumar Pratihar.
A survey on influence maximization in a social network.
Knowledge and Information Systems, 62:3417–3455, 2020.

Federico Battiston, Giulia Cencetti, Iacopo Iacopini, Vito Latora, Maxime Lucas, Alice Patania, Jean-Gabriel Young, and Giovanni Petri.
Networks beyond pairwise interactions: Structure and dynamics.
Physics Reports, 874:1–92, 2020.

References II

Austin R Benson, Rediet Abebe, Michael T Schaub, Ali Jadbabaie, and Jon Kleinberg.
Simplicial closure and higher-order link prediction.
Proceedings of the National Academy of Sciences, 115(48):E11221–E11230, 2018.

Austin R Benson, David F Gleich, and Jure Leskovec.
Higher-order organization of complex networks.
Science, 353(6295):163–166, 2016.

Elad Ganmor, Ronen Segev, and Elad Schneidman.
Sparse low-order interaction network underlies a highly correlated and learnable neural population code.
Proceedings of the National Academy of sciences, 108(23):9679–9684, 2011.

Jacopo Grilli, György Barabás, Matthew J Michalska-Smith, and Stefano Allesina.
Higher-order interactions stabilize dynamics in competitive network models.
Nature, 548(7666):210–213, 2017.

References III

- Jonathan M Levine, Jordi Bascompte, Peter B Adler, and Stefano Allesina.
  Beyond pairwise mechanisms of species coexistence in complex communities.
  Nature, 546(7656):56–64, 2017.
- Mark EJ Newman, Steven H Strogatz, and Duncan J Watts.
  Random graphs with arbitrary degree distributions and their applications.
  Physical review E, 64(2):026118, 2001.
- Giovanni Petri, Paul Expert, Federico Turkheimer, Robin Carhart-Harris, David Nutt, Peter J Hellyer, and Francesco Vaccarino.
  Homological scaffolds of brain functional networks.
  Journal of The Royal Society Interface, 11(101):20140873, 2014.
- Alicia Sanchez-Gorostiaga, Djordje Bajić, Melisa L Osborne, Juan F Poyatos, and Alvaro Sanchez.
  High-order interactions distort the functional landscape of microbial consortia.
  PLoS Biology, 17(12):e3000550, 2019.

References IV

Ming Xie, Xiu-Xiu Zhan, Chuang Liu, and Zi-Ke Zhang.
An efficient adaptive degree-based heuristic algorithm for influence maximization in hypergraphs.
Information Processing & Management, 60(2):103161, 2023.

Dengyong Zhou, Jiayuan Huang, and Bernhard Schölkopf.
Learning with hypergraphs: Clustering, classification, and embedding.
Advances in neural information processing systems, 19, 2006.