

Phase transitions and stability of dynamical processes on hypergraphs

Guilherme Ferraz de Arruda ^{1,4}, Michele Tizzani ^{1,4} & Yamir Moreno ^{1,2,3}

Hypergraphs naturally represent higher-order interactions, which persistently appear in social interactions, neural networks, and other natural systems. Although their importance is well recognized, a theoretical framework to describe general dynamical processes on hypergraphs is not available yet. In this paper, we derive expressions for the stability of dynamical systems defined on an arbitrary hypergraph. The framework allows us to reveal that, near the fixed point, the relevant structure is a weighted graph-projection of the hypergraph and that it is possible to identify the role of each structural order for a given process. We analytically solve two dynamics of general interest, namely, social contagion and diffusion processes, and show that the stability conditions can be decoupled in structural and dynamical components. Our results show that in social contagion process, only pairwise interactions play a role in the stability of the absorbing state, while for the diffusion dynamics, the order of the interactions plays a differential role. Our work provides a general framework for further exploration of dynamical processes on hypergraphs.

¹ISI Foundation, Torino, Italy. ²Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Zaragoza, Spain. ³Department of Theoretical Physics, University of Zaragoza, Zaragoza, Spain. ⁴These authors contributed equally: Guilherme Ferraz de Arruda, Michele Tizzani. email: gui.f.arruda@gmail.com

Network science has been successful in describing the structure and dynamics of complex systems in many fields of science. Nevertheless, in the vast majority of studies, the mathematical and computational descriptions of these networks are limited to pairwise interactions. In most cases, this order of interaction is a good approximation, describing real phenomena^{1–4}. Recently, it has become increasingly evident that simplifying higher-order interactions as a set of pairwise ones can be misleading in many situations^{5,6}, as it is the case, for instance, when comparing triangles and clustering⁶. An attempt to formally solve this problem is represented by the introduction of the simplicial complexes approach^{7,8}. Although very powerful, as a topological space, simplicial complexes require mutual inclusion for the interactions⁹, which is too restrictive for general purposes. To face this problem, one needs to resort to the use of hypergraphs, which by relaxing the assumption of mutual inclusion, allow the representation of a broader range of systems.

Admittedly, the attention to higher-order systems has proliferated lately^{5,6,8,10–20}, with an increasing focus on hypergraphs in fields such as mathematics^{6,12,13,15,16,21,22}, physics^{8,10,11,17–19,23–25}, and computer science^{26–34}. In many fields, the dynamical systems approach appears naturally. This is the case of many problems tackled in the physics literature mentioned above, where we are interested in macroscopic changes that emerge from local interactions. Specifically, this viewpoint was explored in social contagion models in refs. ^{8,17}, where first- and second-order transitions and hysteresis were found. Along the similar lines, the problems of evolutionary game theory¹⁸, synchronization^{23–25}, and random walks^{19,33,34} were also explored. Each field's focus is different, but it is often useful to use an abstract model that has a direct physical interpretation. Perhaps, the simplest example of this is the use of random walks in deep learning and machine learning techniques^{33,34}. Despite this interest, a general theory of dynamical processes on higher-order structures is still largely missing.

Here we address this open problem by building a mathematical framework that allows performing a linear stability analysis for general processes on arbitrary hypergraphs. This approach highlights the importance of the graph projection—an underlying weighted graph representation of a hypergraph—for the dynamics.

The proposed methodology makes it possible to decompose a hypergraph in uniform structures, hence enabling the characterization of their role in the system's dynamics. Finally, to show the usefulness of our approach, we analytically study and recover some results reported for social contagion^{8,17}, and diffusion processes, also providing new key insights into these paradigmatic dynamics. The framework discussed in this paper could be used to explore different processes when pairwise interactions are an oversimplification of the system, which has the potential to bring more insights into the understanding of higher-order dynamics of interacting systems.

Results and discussion

A hypergraph, $\mathcal{H} = \{\mathcal{V}, \mathcal{E}\}$, is defined as a set of nodes, $\mathcal{V} = \{v_i\}$, with $v_i \in \mathbb{Z}^+$, $N = |\mathcal{V}|$ the number of nodes and a set of hyperedges $\mathcal{E} = \{e_j\}$, where e_j is a subset of \mathcal{V} with arbitrary cardinality $|e_j|$. We also denote \mathcal{E}_i as the set of hyperedges that contain the node i . Note that if $\max(|e_j|) = 2$, we recover a graph, whereas one has a simplicial complex if for each hyperedge with $|e_j| > 2$, its subsets are also contained in \mathcal{E} . The adjacency matrix¹⁵ can be defined as

$$\mathbf{A}_{ik} = \sum_{\substack{e_j \in \mathcal{E} \\ i,k \in e_j}} \frac{1}{|e_j| - 1}, \quad (1)$$

which can be interpreted as a weighted projected graph. Note that $\mathbf{A}_{ii} = 0$ for all i . An example of a hypergraph and its graph projection is shown in Fig. 1. A more general definition of the adjacency matrix, including also diagonal elements, can be expressed in terms of its weighted counterpart defined as

$$\mathbf{W}_{ik} = \sum_{\substack{e_j \in \mathcal{E} \\ i,k \in e_j}} \frac{w_{ik}(e_j)}{|e_j| - 1}, \quad (2)$$

where $w_{ik}(e_j)$ is an arbitrary number. When $w_{ik}(e_j) = w(|e_j|)$, the weight is a function of the cardinality of the hyperedge, e_j , weighting differently the contribution of each hyperedge. For instance, if we consider $w(|e_j|) = |e_j| - 1$, we recover the adjacency

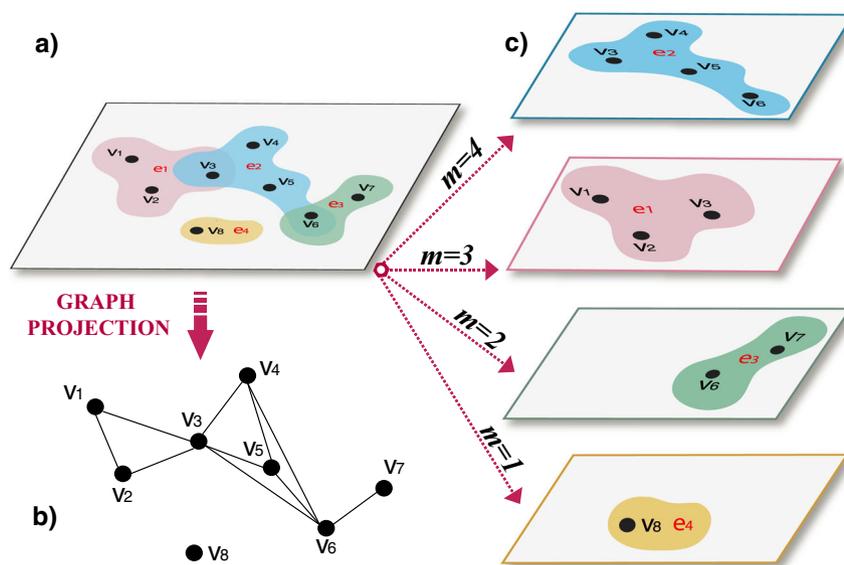


Fig. 1 Graphical representation of a hypergraph. The hypergraph in **a**, its projection in **b**, and structural decomposition in **c**. Mathematically, The set of nodes was represented by $\mathcal{V} = \{v_1, v_2, v_3, v_4, v_5, v_6, v_7, v_8\}$, while the set of hyperedges was $\mathcal{E} = \{e_1, e_2, e_3, e_4\}$, where the hyperedges are $e_1 = \{v_1, v_2, v_3\}$, $e_2 = \{v_3, v_4, v_5, v_6\}$, $e_3 = \{v_6, v_7\}$, and $e_4 = \{v_8\}$. We show how the original hypergraph (**a**) can be represented both by **c** decomposing it into m -uniform hypergraphs (color-coded) or by **b** projection—where hyperedges are simplified as $|e_j|$ -cliques, with $j = 1, \dots, 4$.

matrix used in ref. 19. Furthermore, the Laplacian matrix is defined as $\mathbf{L} = \mathbf{D} - \mathbf{A}$, where $\mathbf{D} = \text{diag}(k_i)$ and the degree is given as $k_i = \sum_{j=1}^N \mathbf{A}_{ij}$. Note that, in order to define the Laplacian matrix on the weighted case we also need to redefine \mathbf{D} accordingly.

We can also decompose the hypergraph into m -uniform hypergraphs, as motivated in ref. 16. Formally,

$$\mathbf{A}_{ik}^m = \sum_{\substack{e_j \in \mathcal{E}, |e_j|=m \\ i, k \in e_j}} \frac{1}{|e_j| - 1}, \quad (3)$$

therefore $\mathbf{A} = \sum_m \mathbf{A}^m$ as showed in Fig. 1. Consequently, we can also define \mathbf{D}^m , \mathbf{L}^m , and \mathbf{W}^m .

Linear stability analysis (LSA). A general dynamical process on the hypergraph can be written as

$$\frac{dx_i}{dt} = f_i(x_i) + \sum_{e_j \in \mathcal{E}_i} g_j(x_{\{e_j\}}), \quad (4)$$

where x_i is the state of the node i , $f_i(x_i)$ is a $\mathbb{R} \rightarrow \mathbb{R}$ function that depends only on the state of node i and $g_j(x_{\{e_j\}})$ is a $\mathbb{R}^{|e_j|} \rightarrow \mathbb{R}$ function that takes all the states of the nodes on the hyperedge e_j , here denoted as $x_{\{e_j\}}$, and compute its contribution to x_i . While $f_i(x_i)$ represents the internal dynamics of the node, $g_j(x_{\{e_j\}})$ is the external interaction, expressed as the hyperedge contribution.

Linear stability analysis for higher-order systems has been explored both in simplicial complexes^{24,25}, where the mutual inclusion constrain imposes a limit on the dynamics, and for particular dynamics with a specific Laplacian for hypergraphs²³. Despite these efforts, a general theory of dynamical processes on higher-order structures is still largely missing. Similarly to ref. 2, we can perform a linear stability analysis for Eq. (4) around a known fixed point, $x_i = x_i^*$. The linearized equations are expressed as

$$\left. \frac{de_i}{dt} = \frac{df_i(x_i)}{dx_i} \right|_{x_i=x_i^*} e_i + \sum_{e_j \in \mathcal{E}_i} \sum_{k \in e_j} \left. \partial_{x_k} g_j(x_{\{e_j\}}) \right|_{\mathbf{x}=\mathbf{x}^*} e_k, \quad (5)$$

where \mathbf{x} is a vector whose components are x_i . We can express Eq. (5) in its matrix form decomposing a general hypergraph into uniform hypergraphs as

$$\frac{d\boldsymbol{\epsilon}}{dt} = (\mathbf{F}(\mathbf{x}^*) + \mathbf{G}(\mathbf{x}^*))\boldsymbol{\epsilon} = \mathbf{M}\boldsymbol{\epsilon}, \quad (6)$$

where

$$\mathbf{F}_{ii}(\mathbf{x}^*) = \begin{cases} \left. \frac{df_i(x_i)}{dx_i} \right|_{x_i=x_i^*} & i = j \\ 0 & \text{otherwise} \end{cases} \quad (7)$$

$$\mathbf{G}_{ik}(\mathbf{x}^*) = \sum_{e_j \in \mathcal{E}_i} \left. \partial_{x_k} g_j(x_{\{e_j\}}) \right|_{\mathbf{x}=\mathbf{x}^*} \quad (8)$$

$$= \sum_{e_j \in \mathcal{E}_i} \mathbb{1}_{\{k \in e_j\}} \left. \partial_{x_k} g_j(x_{\{e_j\}}) \right|_{\mathbf{x}=\mathbf{x}^*},$$

$$\mathbf{M}(\mathbf{x}^*) = \mathbf{F}(\mathbf{x}^*) + \mathbf{G}(\mathbf{x}^*). \quad (9)$$

The fixed point \mathbf{x}^* is stable if

$$\Lambda_i(\mathbf{M}) < 0 \quad \forall i \in \{1, 2, \dots, N\}, \quad (10)$$

where $\Lambda_i(\mathbf{M})$ is the i th eigenvalue of the Jacobian \mathbf{M} . Note that the indicator function $\mathbb{1}_{\{k \in e_j\}}$, which is one if the node k is inside the hyperedge e_j , and zero otherwise, in this case is redundant,

considering that g_j depends only on the states of the nodes inside e_j . The general form of the adjacency matrix in Eq. (2) can be expressed as

$$\begin{aligned} \mathbf{W}_{ik} &= \sum_{\substack{e_j \in \mathcal{E} \\ k \in e_j}} w_{ik}(e_j) \\ &= \sum_{e_j \in \mathcal{E}_i} \mathbb{1}_{\{k \in e_j\}} w_{ik}(e_j), \end{aligned} \quad (11)$$

when $w_{ik}(e_j) = (|e_j| - 1) \partial_{x_k} g_j(x_{\{e_j\}})$ calculated at the fixed point \mathbf{x}^* , $\mathbf{G} = \mathbf{W}$. This shows that, although we have a high-order structure ($N^{\max(|e_j|)}$ in the tensorial representation¹³), near the fixed point, the dynamics is described only by $\mathbf{G} \in \mathbb{R}^{N \times N}$, which is a weighted graph-projection of the hypergraph, the weights being determined by the dynamics. In special cases, as shown next, this projection appears in the form of the Laplacian or the adjacency matrix \mathbf{A} in Eq. (1).

Moreover the matrix \mathbf{G} can be decomposed in m -order contributions \mathbf{G}^m that correspond to the m th uniform hypergraph in the structural decomposition, such that

$$\mathbf{G}_{ik}^m(\mathbf{x}^*) = \sum_{\substack{e_j \in \mathcal{E}_i \\ |e_j|=m}} \left. \partial_{x_k} g_j(x_{\{e_j\}}) \right|_{\mathbf{x}=\mathbf{x}^*}, \quad (12)$$

$$\mathbf{G}(\mathbf{x}^*) = \sum_{m=1}^{\max(|e_j|)} \mathbf{G}^m(\mathbf{x}^*). \quad (13)$$

The structures of cardinality m do not contribute to the stability of the fixed point \mathbf{x}^* if $\mathbf{G}_{ik}^m(\mathbf{x}^*) = 0 \quad \forall i, k$. A similar approach can be tackled by the bipartite representation of the hypergraph, as shown in the next section.

Linear stability analysis: comparison with the bipartite representation. In order to define the bipartite representation of a hypergraph, we have to introduce the incidence matrix. The common definition is given by

$$\mathcal{I}_{ij} = \begin{cases} 1 & \text{if } i \in e_j \\ 0 & \text{if } i \notin e_j \end{cases}, \quad (14)$$

where $\mathcal{I} \in \mathbb{R}^{N \times M}$ as shown in refs. 19,20,23. In this case, we have two sets of nodes, one representing the hypergraph nodes, and the other the hyperedges. The relative projection is encoded in the weighted adjacency matrix¹⁹

$$\tilde{\mathbf{A}} = \mathcal{I}\mathcal{I}^T - \mathbf{D}, \quad (15)$$

where $\mathbf{D} = \text{diag}(k_i)$ is a diagonal matrix whose diagonal elements are the degrees. The element $\tilde{\mathbf{A}}_{ij}$ can be interpreted as the number of hyperedges shared by nodes i and j . Notice that $\sum_j \tilde{\mathbf{A}}_{ij}$ is the number of neighbors of i , which can be different from the degree. On the other hand, the projection in Eq. (1), can be defined as

$$\mathbf{A} = \mathcal{B}\mathcal{B}^T - \hat{\mathbf{D}}, \quad (16)$$

$$\hat{\mathbf{D}}_{ii} = \sum_{e_j \cap \{i\} \neq \emptyset} \frac{1}{|e_j| - 1}, \quad (17)$$

where the weighted incidence matrix is

$$\mathcal{B}_{ij} = \begin{cases} (|e_j| - 1)^{-\frac{1}{2}} & \text{if } i \in e_j \\ 0 & \text{if } i \notin e_j \end{cases} \quad (18)$$

and $\hat{\mathbf{D}}$ is a diagonal matrix. In this case, $\sum_j \mathbf{A}_{ij}$ correctly represents the degree, like in standard graph theory. Moreover, the

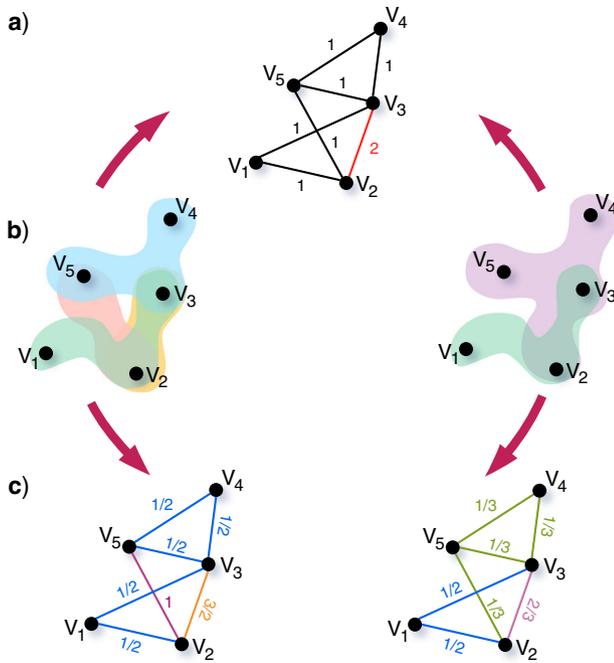


Fig. 2 Different hypergraph projections. Projection of two different hypergraphs built on the same set of nodes, $\mathcal{V} = \{v_1, v_2, v_3, v_4, v_5\}$. In **a** projection following the incidence matrix in Eq. (15), **b** original hypergraphs, and **c** the weighted projection for the model presented in Eq. (16). In **a** and **c**, the edges' colors represent their weights as indicated next to each edge.

projection in Eq. (15) can represent multiple hypergraphs, as shown in Fig. 2a, b, while, in the case of Eq. (16) we could not find the same projection for two different hypergraphs (Fig. 2b, c). To the best of our knowledge, the proof of the one to one mapping is an open problem.

From a dynamical point of view, Eq. (4) can be written in terms of the bipartite representation as

$$\frac{dx_i}{dt} = f_i(x_i) + \sum_{j=1}^M \mathcal{B}_{ij} g_j(x_{\{e_j\}}). \quad (19)$$

In this case, the linear stability is expressed by

$$\frac{d\epsilon_i}{dt} = \left. \frac{df_i(x_i)}{dx_i} \right|_{x_i=x_i^*} \epsilon_i + \sum_{j=1}^M \sum_{k=1}^M \mathcal{B}_{ij} \mathcal{B}_{kj} \partial_{x_k} g_j(x_{\{e_j\}}) \Big|_{\mathbf{x}=\mathbf{x}^*} \epsilon_k, \quad (20)$$

recovering Eq. (5).

Linear stability analysis of separable interacting functions (LSA-SIF). We now focus on the special case of a diffusion-like process on hypergraphs. In this case, $g_j(x_{\{e_j\}})$ is separable and we can express the interaction function as

$$g_j(x_{\{e_j\}}) = \frac{1}{|e_j| - 1} \sum_{\substack{k \in e_j \\ k \neq i}} (g(x_i) - g(x_k)), \quad (21)$$

where the weight $(|e_j| - 1)^{-1}$ is necessary due to conservation purposes. Assuming that $f_i(x_i)$ is the same for all nodes, the fixed point is symmetrical, i.e., $x_i^* = x_j^*$ for any pair i, j . In this case, Eqs. (7) and (8) reduce to

$$\mathbf{F} = \alpha \mathbf{I} \text{ and } \mathbf{G} = \beta \mathbf{L}, \quad (22)$$

where $\mathbf{I} \in \mathbb{R}^{N \times N}$ is the identity matrix, and

$$\alpha = \left. \frac{df(x)}{dx} \right|_{\mathbf{x}=\mathbf{x}^*} \text{ and } \beta = \left. \frac{dg(x)}{dx} \right|_{\mathbf{x}=\mathbf{x}^*}. \quad (23)$$

Therefore, Eq. (5) is expressed as

$$\frac{d\epsilon_i}{dt} = \left(\alpha + \sum_{e_j \in \mathcal{E}_i} \sum_{k \in e_j} \frac{\beta}{|e_j| - 1} \right) \epsilon_i - \sum_{e_j \in \mathcal{E}_i} \sum_{k \in e_j} \frac{\beta \epsilon_k}{|e_j| - 1}, \quad (24)$$

which in matrixial form reads

$$\frac{d\boldsymbol{\epsilon}}{dt} = (\alpha \mathbf{I} + \beta \mathbf{L}) \boldsymbol{\epsilon}, \quad (25)$$

where $\boldsymbol{\epsilon}$ is the vector whose components are ϵ_i . If μ_i are the eigenvalues of \mathbf{L} , the stability condition for the fixed point \mathbf{x}^* can be written as $\alpha + \beta \mu_i < 0$, for all i . Next, considering that the Laplacian matrix is semi-positive definite, the former condition in terms of the largest eigenvalue, μ_n , is

$$\frac{1}{\mu_n} > - \left[\frac{dg(x)}{dx} / \frac{df(x)}{dx} \right]_{x_i=x_i^*}. \quad (26)$$

Linear stability analysis of interacting functions with symmetrical derivatives (LSA-SD). Now, let us consider a second scenario for $g_j(x_{\{e_j\}})$ such that

$$g_j^*(x_{\{e_j\}}) = \frac{1}{(|e_j| - 1)} g_j(x_{\{e_j \setminus \{i\}\}}), \quad (27)$$

which depends, for a node i , only on x_k , where $k \in e_j$ and $k \neq i$, and the weight is the same as for Eq. (21). Assuming that $f_i(x_i)$ is the same for all nodes, and also all the derivatives of g_j^* are all the same in the fixed point, Eq. (7) reduces to

$$\frac{d\epsilon_i}{dt} = \alpha \epsilon_i + \sum_{e_j \in \mathcal{E}_i} \sum_{k \in e_j} \frac{\beta \epsilon_k}{|e_j| - 1}, \quad (28)$$

or, equivalently,

$$\frac{d\boldsymbol{\epsilon}}{dt} = (\alpha \mathbf{I} + \beta \mathbf{A}) \boldsymbol{\epsilon}. \quad (29)$$

Note that, the former restriction covers the wide class of normalized symmetric functions²³. Denoting by λ_i , the eigenvalues of \mathbf{A} , the stability condition reads

$$\frac{1}{\lambda_n} < - \left[\frac{dg(x)}{dx} / \frac{df(x)}{dx} \right]_{x_i=x_i^*} < \frac{1}{\lambda_1}. \quad (30)$$

In these two general cases, the projected adjacency matrix, Eq. (1), arises naturally from the linear stability analysis around the fixed point, which means that only the projected structure is relevant for the dynamics. Moreover, as it can be seen from Eqs. (26) and (30), we can decouple the structural and dynamical contributions to the stability conditions.

Linear stability analysis of uniform hypergraphs (LSA-UH). So far we have focused on arbitrary structures. However, it is worth emphasizing that structural constraints might allow for a simplification of the linear stability condition. A special class of hypergraphs that allows for that are uniform hypergraphs, for which all hyperedges have the same cardinality. Note that although the cardinality of the hyperedges is the same, the degree distribution could be arbitrary. Moreover, even if constrained, this class of hypergraphs was used to study epidemic spreading³⁵ and evolutionary games¹⁸. As all the hyperedges have the same cardinality, the interacting functions will be the same, i.e., $g_j(x_{\{e_j\}}) = g(x_{\{e_j\}})$. Assuming symmetric derivatives and denoting them by β as in Eq. (21), and, given that we are restricted to undirected hypergraphs, \mathbf{G} will comply with one of the three following cases: (i) separable functions, thus $\mathbf{G} = \beta \mathbf{L}$, (ii) normalized functions of its entries²³, where $g(x_{\{e_j\}})$ does not distinguish between x_i and the other arguments, in this case $\mathbf{G} = \beta$

(**A + D**), which has the same form of sing-less Laplacian in graph theory, (iii) independent of x_i , i.e., $g(x_{\{e_j\}}) = g_j(x_{\{e_j \setminus \{i\}\}})$, thus $\mathbf{G} = \beta \mathbf{A}$.

Note that although the former derivations are specific for these cases, Eq. (10) holds for general dynamics in an arbitrary hypergraph. The framework outlined up to now is general and can be readily applied to many dynamical processes in which one is forced to go beyond pairwise interactions. In what follows, we discuss two of such applications, namely, social contagion^{8,17} and diffusion processes. For the social contagion case, the stability of the process is related to the adjacency matrix, following a similar form as Eq. (30), where we were able to employ the structural decomposition to evaluate the role of pairwise and higher-order structures. On the other hand, for the diffusion, the stability depends on the Laplacian matrix, yielding to a condition of the form of Eq. (26). In this case, we show that hypergraphs' structural organization may introduce a nontrivial behavior, and we remark that the Laplacian operator is connected to many other dynamical processes, such as random walks and synchronization. With these two processes, we exemplify possible applications of the linear stability analysis and the structural decomposition, providing insights about these dynamics in hypergraphs and suggesting further implications in other processes.

Application to a process of social contagion. In this dynamical process, the state x_i represents the probability of an individual to be active. The deactivation mechanism is given by $f_i(x_i) = -\delta x_i$ and the interaction functions $g_j(x_{\{e_j\}})$ are given by the product of the probability that i is inactive and the other nodes in the hyperedge e_j are active times a contact rate, $\beta_{|e_j|}$. Hence,

$$g_j(x_{\{e_j\}}) = \frac{\beta_{|e_j|}}{|e_j| - 1} (1 - x_i) \prod_{\substack{k \in e_j \\ k \neq i}} x_k. \quad (31)$$

Separating the pairwise and higher-order contributions to the state of the node, we have

$$\begin{aligned} \frac{dx_i}{dt} = & -\delta x_i + \beta_2 \sum_{\substack{e_j \in \mathcal{E}_i \\ |e_j|=2 \\ k \in e_j, k \neq i}} (1 - x_i) x_k + \\ & + \sum_{\substack{e_j \in \mathcal{E}_i \\ |e_j|>2}} \frac{\beta_{|e_j|}}{|e_j| - 1} (1 - x_i) \prod_{\substack{k \in e_j \\ k \neq i}} x_k. \end{aligned} \quad (32)$$

Assuming a symmetric fixed point, i.e., $x_i = x^*$,

$$\begin{aligned} \frac{d\epsilon_i}{dt} = & \left(-\delta - \sum_{e_j \in \mathcal{E}_i} \frac{\beta_{|e_j|} (x^*)^{(|e_j|-1)}}{|e_j| - 1} \right) \epsilon_i \\ & + \sum_{e_j \in \mathcal{E}_i} \sum_{\substack{k \in e_j \\ k \neq i}} \frac{\beta_{|e_j|} (1 - x^*) (x^*)^{(|e_j|-2)}}{|e_j| - 1} \epsilon_k, \end{aligned} \quad (33)$$

where we can apply Eq. (10) to evaluate the stability of the system's dynamics. The mean-field form of Eq. (32) is given by

$$\frac{dx}{dt} = -\delta x + P^\nu(x) = -\delta x + \sum_{k=1}^{\nu} \beta_k c(k) (1 - x) x^{k-1}, \quad (34)$$

where $\nu = \max_j \{|e_j|\}$ is the maximum cardinality and $c(k)$ is the ratio between the number of hyperedges with cardinality k and the number of pairwise interactions, which characterizes the structure of the hypergraph. In the steady state this is a

polynomial equation whose solutions are the fixed points of the process and their stability can be evaluated using Eq. (33).

Social contagion: general behavior. A fundamental observation from Eq. (33) is that only pairwise interactions, given by the off-diagonal terms, are responsible for the stability of the absorbing state, the fixed point $x_i = 0$. Specifically, the stability condition for this state depends on the second-order adjacency matrix from Eq. (3) and is given by

$$\frac{\beta}{\delta} < (\Lambda_{\max}(\mathbf{A}^2))^{-1}. \quad (35)$$

Interestingly enough, in this scenario, the dynamics behaves like an SIS process on complex networks, and is characterized by a continuous phase transition³⁶. Therefore, we claim that in a social contagion model, pairwise interactions are a necessary condition for a continuous phase transition. We can support this claim using an argument by contradiction as follows. Near the absorbing state the probability for a node to be active is very small, hence, assuming an arbitrary small number $\eta \approx 0$, we can consider $x_i \in O(\eta)$. Therefore, from Eq. (32), the coefficients $\beta_{|e_j|}$ are multiplying terms of order $O(\eta^{(|e_j|-1)})$. At the steady state, $\delta x_i \sim \beta_{|e_j|} x_i$, assuming $\delta \in O(1)$ without loss of generality. From this point, we can differentiate three cases depending on $\beta_{|e_j|}$:

- (i) If $\beta_{|e_j|} \in O(1)$ for all $|e_j| \in \{1, 2, \dots, \max(|e_j|)\}$, then the leading coefficient is β_2 therefore $\beta_2 x_i \in O(\eta)$ as assumed. In this case we have a continuous phase transition with respect to the control parameter β_2 ;
- (ii) If $\beta_2 \in O(1)$ and $\beta_{|e_j|} \in O(\eta^{-(|e_j|-2)})$, for $|e_j| > 2$, then the leading coefficient is $\beta_{|e_j|} x_i \in O(\eta)$, thus the transition can be either continuous or discontinuous;
- (iii) If $\beta_{|e_j|} \in O(\eta^{-(|e_j|-1)})$ for all $|e_j| \in \{1, 2, \dots, \max(|e_j|)\}$, then the leading coefficient is $\beta_{|e_j|} x_i \in O(1)$, which contradicts the initial assumption of $x_i \in O(\eta)$. This suggests a discontinuity if $\beta_2 \in O(1)$ and at least one of the higher-order coefficients is $O(\eta^{-(|e_j|-1)})$.

Social contagion: mean-field approach. To further extend the analysis in ref. 8, in order to compute $c(k)$, we consider higher-order contact patterns from two real datasets. The data were collected with proximity sensors in a high school^{14,37}, and a primary school^{14,38}. Our results are summarized in Fig. 3a, b, where we consider the feasible fixed points and their stability. In this case, we assume $\beta_2 = \beta$ and $\beta_k = \beta e^k$, in order to emphasize the role of lower and higher-order structures. Notice that, given the same parameters, we can have either continuous or discontinuous phase transitions with hysteresis, depending only on the structure of the hypergraph. Considering the graph case, i.e., $\max_j \{|e_j|\} = 2$, the social contagion model is an SIS epidemic spreading. Thus, the behavior near the fixed point $x^* = 0$ is

$$\frac{d\epsilon_i}{dt} = -\delta \epsilon_i + \beta_2 \sum_{e_j \in \mathcal{E}_i} \epsilon_k. \quad (36)$$

Note that the stability of the absorbing state is given by $\lambda_{\max} < \frac{\delta}{\beta_2}$, as for the quenched mean-field³⁹.

Next, we analyze the simplicial complex case, where the structure is homogeneous and consists only of pairwise interactions and triangles, hence $n = 3$. In this case, Eq. (34) has the

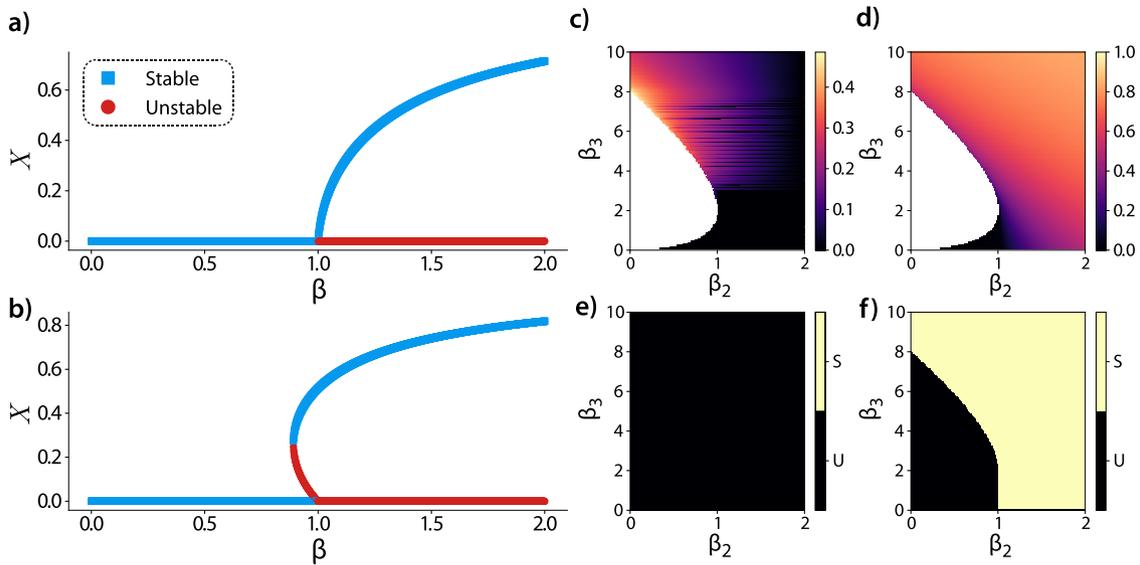


Fig. 3 Stability analysis of a social contagion dynamics on real-world, and synthetic hypergraphs. In **a**, the substrate is a high school contact pattern, while in **b**, a primary school contact pattern. The order parameter, X , is the fraction of active nodes, while β is the spreading rate. **c-f** Results for the stability analysis of a simplicial complex in the space of the spreading rates for pairwise interactions, β_2 , and triangles, β_3 . **c, e** Correspond, respectively, to the solution x_2^* (Eq. (37)) and its stability (Eq. (38)), whereas **d, f** represent the solution x_2^+ (Eq. (37)) and its stability (Eq. (38)), respectively. Here, x_2^* and x_2^+ are solutions of the mean-field equation that describes the social contagion model. In **e, f**, U and S stand for unstable and stable, respectively.

three solutions, $x_1^* = 0$, and

$$x_2^\pm = \frac{(\frac{\beta_2}{2} - \beta_2) \pm \sqrt{(\beta_2 - \frac{\beta_2}{2})^2 - 2\beta_3(\delta - \beta_2)}}{\beta_3}. \quad (37)$$

From Eq. (33) we derive the following condition

$$\left(-\delta - \beta_2 x^* - \frac{\beta_3 x^{*2}}{2} + \beta_2(1 - x^*) + \beta_3(1 - x^*)x^*\right) < 0, \quad (38)$$

thus, $x_1^* = 0$ is stable for $\beta_2 < \delta$. The feasibility of the solution x_2^\pm was shown in ref. 8. Complementarily, Fig. 3c-f shows the results of the corresponding stability analysis, illustrating how our approach could provide new insights. As noted before, we stress that Eq. (33) is not limited to the case of simplicial complexes, but it can be used to further explore higher-order structures. Our findings are in alignment with ref. 40.

Social contagion: uniform hypergraphs. Restricting our analysis to uniform hypergraphs, Eq. (32) reduces to

$$\frac{dx_i}{dt} = -\delta x_i + \frac{\beta_{|e_j|}}{|e_j| - 1} \sum_{e_j \in \mathcal{E}_i} (1 - x_i) \prod_{\substack{k \in e_j \\ k \neq i}} x_k, \quad (39)$$

whose solutions are

$$\mathbf{x} = \beta^* \left(\mathbf{A} \mathbf{x}^{[|e_j|-1]} - \mathbf{A} \mathbf{x}^{[|e_j|]} \right), \quad (40)$$

where the $\beta^* = \frac{\beta_{|e_j|}}{\delta(|e_j|-1)}$ is the scaled spreading rate and $\mathbf{x}^{[k]}$ is the element-wise k th power vector, i.e., $\mathbf{x}^{[k]} = [x_i^k]^T$. We now focus on homogeneous degree distributions where all the nodes have the same degree. In this case, $x = x_i, \forall i \in \mathcal{V}$ and the solutions, x^* (β^*), of Eq. (40) are the roots

$$\beta^* \langle k \rangle (x^{|e_j|-1} - x^{|e_j|}) - x = 0, \quad (41)$$

which are stable if

$$-\delta - \beta^* \langle k \rangle \delta (|e_j - 1|) (1 - x^*) (x^*)^{|e_j|-2} < 0. \quad (42)$$

We observe numerically that Eq. (41) has three feasible solutions, of which two are stable, including the absorbing state $x^* = 0$, and one unstable. The first non-zero positive value of the stable solution is denoted by $Q_1(\beta_c^*)$, where β_c^* is the tipping point, using the same notation as ref. 17. As shown in Fig. 4a, b, both $Q_1(\beta_c^*)$ and β_c^* increase with the cardinality, in particular the last one follows a linear trend as shown in Fig. 4c. Notice that in this annealed approach we can consider $\langle k \rangle$ just as a scaling factor of β^* .

From Eq. (41), considering that β^* is a free parameter, $x = 0$ is always a solution. Moreover, for finite β^* and in the limit of infinite cardinality, in the range of $0 \leq x \leq 1$, $x = 0$ is the only solution. Finally, in the case of infinite β^* , the only feasible solution is $x = 1$. This argument aligns with the numerical results of Fig. 4c. For $|e_j| = 2$, trivially we have an SIS dynamics and a continuous phase transition. On the other hand, in the limit of infinite cardinality, the system displays a discontinuous phase transition and $\beta_c^* \rightarrow \infty$ and $Q_1(\beta_c^*) = 1$. For finite cardinalities the tipping point is also finite.

Application to diffusion processes. Another immediate application is the diffusion process². In this case, $g_j(x_{\{e_j\}})$ is separable and $g(x_i) = x_i$. The dynamics of the process is given by

$$\frac{dx}{dt} = -D \mathbf{L} \mathbf{x}. \quad (43)$$

where D is the diffusion constant and the minus sign is a consequence of the stability condition in Eq. (26). We can arbitrarily decompose the Laplacian in m -uniform hypergraphs as done for the adjacency matrix in Eq. (3). In particular, we want to consider if the lower-order contributions are enough to describe the dynamics. To this end, we can separate the Laplacian in two components $\mathbf{L} = \mathbf{L}_O(\theta) + \mathbf{L}_R(\theta)$ given by

$$\mathbf{L}_O(\theta) = \sum_{m=1}^{\theta} \mathbf{L}^m \quad \text{and} \quad \mathbf{L}_R(\theta) = \sum_{m=\theta+1}^{\max(|e_j|)} \mathbf{L}^m, \quad (44)$$

where θ differentiates between higher and lower-order Laplacian, $\mathbf{L}_R(\theta)$ and $\mathbf{L}_O(\theta)$, respectively. In order to underline the important

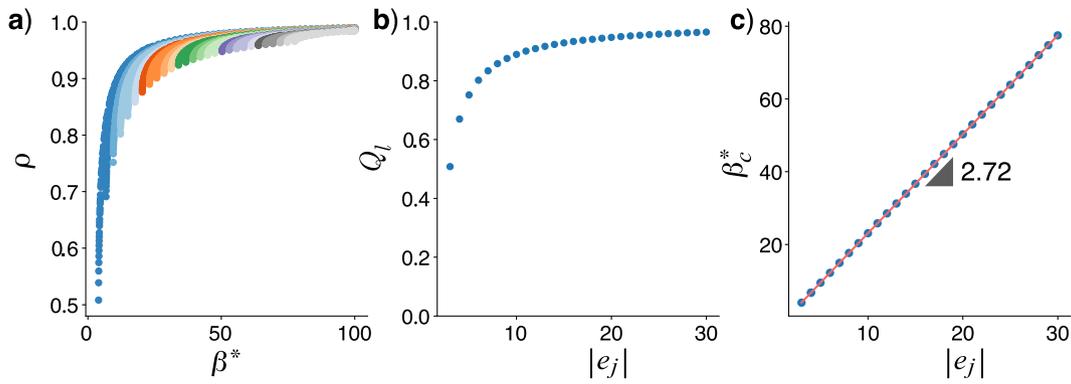


Fig. 4 Numerical results for the social contagion dynamics on uniform hypergraphs. **a** The feasible stable solution as a function of the scaled spreading rate β^* for different values of cardinality (color-coded), in the range from $|e_j| = 3$ in blue, to $|e_j| = 30$ in light gray. **b** The first non-zero positive value of the stable solution, Q_l , as a function of the cardinality. **c** Depicts the behavior of β^* at the tipping point, β_c^* , as a function of the cardinality, revealing a linear behavior.

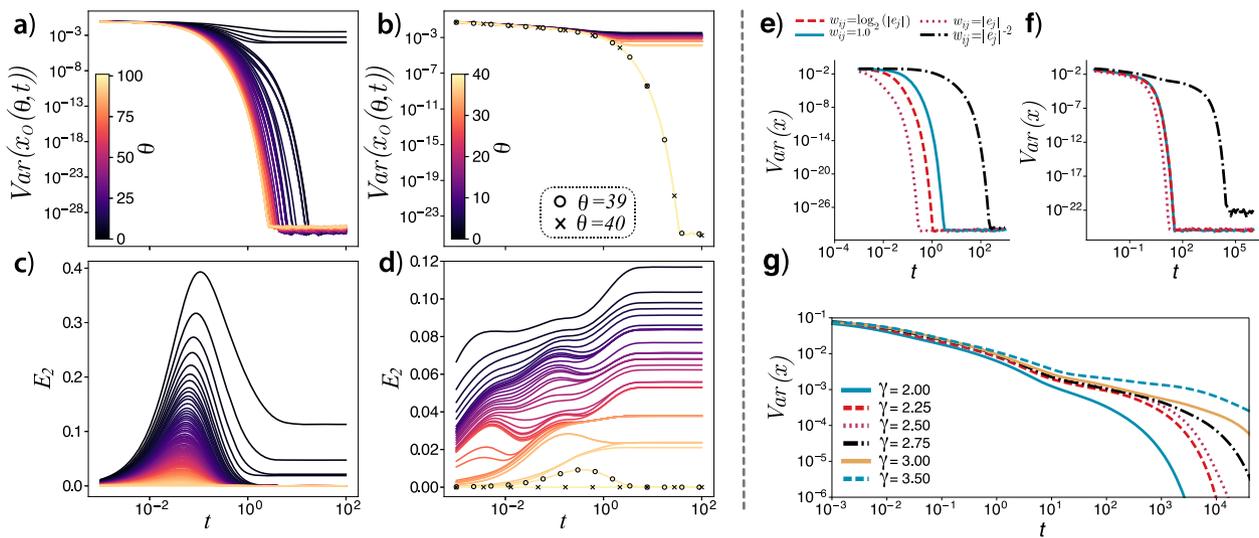


Fig. 5 Characterization of diffusion on real and synthetic hypergraphs. **a-c** Represents real hypergraphs, while **e-g** synthetic ones. **a, b** The variance of states, $x_o(\theta, t)$, up to order θ , in the diffusion, while **c** and **d** display the error E_2 , both as a function of time and structures considered, θ (color-coded). To represent the hypergraphs, we used in **a** and **c**, a power-law distribution of cardinalities, whereas in **b** and **d**, the hypergraph is given by the real email data (see text). The three panels to the right (**e-g**) show results obtained for weighted diffusion on hypergraphs, in terms of the variance of the state of the complete system x , defined by Eq. (43). **e, f** The effect of the different weights as a function of the cardinality, Eq. (2), see the legend. **e** Corresponds to a synthetic hypergraph with a power-law distribution of cardinalities, $P(|e_j|) \sim |e_j|^{-2.05}$, and **f** to the email hypergraph. **g** Explores the dependency with the exponent of the power-law weighting function.

role played by the structure on the dynamics, we next analyze a diffusion process on both a synthetic hypergraph generated according to a power-law cardinality distribution, $P(|e_j|) \sim |e_j|^{-2.05}$, and a real hypergraph using data from email exchange at a European research institution^{14,41,42}. In Fig. 5a, b, we show the variance of $x_o(\theta, t) = e^{-DL_o(\theta)t}x(0)$, which is the solution of the diffusion process on the lower-order Laplacian L_o . For a small enough θ , we have disconnected components and the process does not reach the final state of the complete Laplacian. On the other hand, when the final state is reached, different θ 's correspond to different timescales, and in particular, convergence is faster when increasing θ . In fact, for the synthetic case we need less information (structures) to represent the whole process, in the email dataset the convergence is only reached for $\theta = 1$, as shown in Fig. 5b. The relative error can be estimated as

$$E_2 = \frac{\|x(t) - x_o(\theta, t)\|_2}{\|x(t)\|_2}. \quad (45)$$

For the synthetic case, as shown in Fig. 5c, the error curves show

peaks that are higher for lower values of θ . For long enough times, different plateau is reached depending on how many components are considered. The same is true for the email case, Fig. 5d, however we have a multiple peaked pattern, suggesting that we have various structures, each one with a different time scale. As expected, only for $\theta = 39$, the error goes to zero, highlighting that real structures might have nontrivial behavior.

Another important aspect is the role on the dynamics of considering different weights for the hyperedges^{8,17,18}. In particular, by redefining the Laplacian with the weighted adjacency matrix, Eq. (2), different choices of $w(|e_j|)$ can be used, as shown in Fig. 5e-g. As expected, increasing the weights reduces the convergence time. Aside from the different timescales in the two examples, we show that real data might also have a qualitatively different behavior, as shown in Fig. 5f, g. In particular, we focus on the power-law weighting functions, $w(|e_j|) = |e_j|^{-\gamma}$, in Fig. 5g. In this case, since higher-order structures are heavily penalized, we observe a greater time modulation by increasing γ . We also note that a well-established approach to this

problem could be the analysis of some of the spectral modes of the Laplacian. In this case, it is not possible to decouple the contributions of different orders from the spectrum. On the other hand, the structural decomposition explicitly takes into account each order, allowing a straightforward analysis of the dynamics.

Conclusions

In this paper, we have developed a framework that allows to describe the relationship between the structural organization of hypergraphs and a highly general class of dynamical processes, beyond pairwise interactions. Within this framework, the graph projection appears naturally from a linear stability analysis around a given fixed point. Therefore, the dimensionality of the hypergraph, $N^{\max(|e_j|)}$ in its tensor representation, around the fixed point is reduced to a N^2 space. Moreover, we showed that using a structural decomposition in m -uniform hypergraphs, given by Eq. (3), enables the individual analysis of each m th order component. Altogether, the methodology provides the stability condition for a general class of dynamical process and also allows to identify the structural components that play a role in the stability of any known fixed point. Of particular interest are the special cases LSA-SIF, LSA-SD and LSA-UH as one is able to separate the structural and dynamical contributions to stability. We also applied the proposed framework to two relevant dynamical processes, namely, social contagion and diffusion dynamics. For the former, we not only recovered known results for graph-based and simplicial complexes, but we obtained a necessary condition for the stability of the absorbing state, showing that only pairwise interactions play a role for such a state. In addition, our analysis revealed the conditions for the continuity of the phase transition also for real systems, where different kinds of bifurcations might appear. Furthermore, in the special case of uniform hypergraphs, we have shown that both the tipping point and $Q_I(\beta_c^*)$ increase with the cardinality. In particular, for finite cardinalities the tipping point is also finite, whereas in the infinite limit the two possible solutions are $x=0$ and $x=1$, being the latter only possible for $\beta^* \rightarrow \infty$. Finally, with regard to diffusion processes, we characterized the role played by the structural decomposition of the hypergraph and different weighting functions for the hyperedges on the dynamics of the system. In conclusion, we are confident that this investigation will motivate further research on the spectral properties of hypergraphs, their structural characterization, and in providing relevant insights in different areas where higher-order interactions can not be neglected, also beyond the community of dynamical processes on complex networks.

Data availability

The dataset used is available at: <https://www.cs.cornell.edu/~arb/data/>.

Code availability

Custom code that supports the findings of this study is available from the corresponding author upon request.

Received: 1 December 2020; Accepted: 11 January 2021;

Published online: 12 February 2021

References

- Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D.-U. Complex networks: structure and dynamics. *Phys. Rep.* **424**, 175–308 (2006).
- Newman, M. *Networks: an introduction* (Oxford University Press, Inc., 2010).
- Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925–979 (2015).
- Cimini, G. et al. The statistical physics of real-world networks. *Nat. Rev. Phys.* **1**, 58–71 (2019).
- Lambiotte, R., Rosvall, M. & Scholtes, I. From networks to optimal higher-order models of complex systems. *Nat. Phys.* **15**, 313–320 (2019).
- Chodrow, P. S. Configuration models of random hypergraphs. *J. Complex Netw.* **8**, cnaa018, <https://doi.org/10.1093/comnet/cnaa018> (2020).
- Petri, G. & Barrat, A. Simplicial activity driven model. *Phys. Rev. Lett.* **121**, 228301 (2018).
- Iacopini, I., Petri, G., Barrat, A. & Latora, V. Simplicial models of social contagion. *Nat. Commun.* **10**, 1–9 (2019).
- Armstrong, M. *Basic topology, Undergraduate texts in mathematics* (Springer, 2013).
- Estrada, E. & Rodríguez-Velázquez, J. A. Subgraph centrality and clustering in complex hyper-networks. *Phys. A: Stat. Mech. Appl.* **364**, 581–594 (2006).
- Ghoshal, G., Zlatić, V., Caldarelli, G. & Newman, M. E. J. Random hypergraphs and their applications. *Phys. Rev. E* **79**, 066118 (2009).
- Bodó, Á., Katona, G. Y. & Simon, P. L. SIS epidemic propagation on hypergraphs. *Bull. Math. Biol.* **78**, 713–735 (2016).
- Banerjee, A., Char, A. & Mondal, B. Spectra of general hypergraphs. *Linear Algebra Appl.* **518**, 14–30 (2017).
- Benson, A. R., Abebe, R., Schaub, M. T., Jadbabaie, A. & Kleinberg, J. Simplicial closure and higher-order link prediction. *Proc. Natl. Acad. Sci.* **115**, E11221–E11230 (2018).
- Banerjee, A. On the spectrum of hypergraphs. *Linear Algebra Appl.* <https://doi.org/10.1016/j.laa.2020.01.012> (2020). (In press).
- Ouvrard, X., Goff, J.-M. L. & Marchand-Maillet, S. Adjacency and tensor representation in general hypergraphs part 1: e-adjacency tensor uniformisation using homogeneous polynomials. <https://arxiv.org/abs/1712.08189> (2017).
- de Arruda, G. F., Petri, G. & Moreno, Y. Social contagion models on hypergraphs. *Phys. Rev. Res.* **2**, 023032 (2020).
- Alvarez-Rodriguez, U. et al. Evolutionary dynamics of higher-order interactions in social networks. *Nat. Hum. Behav.* <https://doi.org/10.1038/s41562-020-01024-1> (2021).
- Carletti, T., Battiston, F., Cencetti, G. & Fanelli, D. Random walks on hypergraphs. *Phys. Rev. E* **101**, 022308 (2020).
- Battiston, F. et al. Networks beyond pairwise interactions: structure and dynamics. *Phys. Rep.* **874**, 1–92 (2020).
- Qi, L. & Luo, Z. *Tensor analysis: spectral theory and special tensors. Other titles in applied mathematics* (Society for Industrial and Applied Mathematics, 2017).
- Ouvrard, X. Hypergraphs: an introduction and review. <https://arxiv.org/abs/2002.05014> (2020).
- Mulas, R., Kuehn, C. & Jost, J. Coupled dynamics on hypergraphs: Master stability of steady states and synchronization. *Phys. Rev. E* **101**, 062313 (2020).
- Lucas, M., Cencetti, G. & Battiston, F. Multiorder laplacian for synchronization in higher-order networks. *Phys. Rev. Res.* **2**, 033410 (2020).
- Gambuzza, L. V. et al. The master stability function for synchronization in simplicial complexes. <https://arxiv.org/abs/2004.03913> (2020).
- Karypis, G., Aggarwal, R., Kumar, V. & Shekhar, S. Multilevel hypergraph partitioning: applications in VLSI domain. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **7**, 69–79 (1999).
- Tian, Z., Hwang, T. & Kuang, R. A hypergraph-based learning algorithm for classifying gene expression and arrayCGH data with prior knowledge. *Bioinformatics* **25**, 2831–2838 (2009).
- Voloshin, V. I. *Introduction to graph and hypergraph theory* (Nova Science Publishers, 2009).
- Bretto, A. *Hypergraph theory: an introduction* (Springer Publishing Company, Incorporated, 2013).
- Valdivia, P., Buono, P., Plaisant, C., Dufournaud, N. & Fekete, J. Analyzing dynamic hypergraphs with parallel aggregated ordered hypergraph visualization. *IEEE Trans. Vis. Comput. Graph.* **27**, 1–13 (2021).
- Payne, J. Deep hyperedges: a framework for transductive and inductive learning on hypergraphs. <https://arxiv.org/abs/1910.02633> (2019).
- Jiang, J., Wei, Y., Feng, Y., Cao, J. & Gao, Y. Dynamic hypergraph neural networks. In *Proc. Twenty-Eighth International Joint Conference on Artificial Intelligence*, 2635–2641 (International Joint Conferences on Artificial Intelligence Organization, 2019).
- Chitra, U. & Raphael, B. J. Random walks on hypergraphs with edge-dependent vertex weights. In *Proc. 36th International Conference on Machine Learning* (eds. Chaudhuri, K. & Salakhutdinov, R.), Vol. 97, 1172–1181 (PMLR, Long Beach, California, USA, 2019).
- Hayashi, K., Aksoy, S. G., Park, C. H. & Park, H. Hypergraph random walks, laplacians, and clustering. In *Proc. 29th ACM International Conference on Information & Knowledge Management*, 495–504 (Association for Computing Machinery, New York, NY, USA, 2020).
- Jhun, B., Jo, M. & Kahng, B. Simplicial SIS model in scale-free uniform hypergraph. *J. Stat. Mech. Theory Exp.* **2019**, 123207 (2019).
- Mieghem, P. V. Epidemic phase transition of the SIS type in networks. *EPL (Europhys. Lett.)* **97**, 48004 (2012).

37. Mastrandrea, R., Fournet, J. & Barrat, A. Contact patterns in a high school: a comparison between data collected using wearable sensors, contact diaries and friendship surveys. *PLoS ONE* **10**, e0136497 (2015).
38. Stehlé, J. et al. High-resolution measurements of face-to-face contact patterns in a primary school. *PLoS ONE* **6**, e23176 (2011).
39. Miegheem, P. V., Omic, J. & Kooij, R. Virus spread in networks. *IEEE/ACM Trans. Netw.* **17**, 1–14 (2009).
40. Kuehn, C. & Bick, C. A universal route to explosive phenomena. <https://arxiv.org/abs/2002.10714> (2020).
41. Leskovec, J., Kleinberg, J. & Faloutsos, C. Graph evolution: Densification and shrinking diameters. *ACM Trans. Knowl. Discov. Data* **1**, <https://doi.org/10.1145/1217299.1217301> (2007).
42. Yin, H., Benson, A. R., Leskovec, J. & Gleich, D. F. Local higher-order graph clustering. In *Proc. 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (ACM Press, 2017).

Acknowledgements

We acknowledge support from Intesa Sanpaolo Innovation Center. Y.M. acknowledges partial support from the Government of Aragón and FEDER funds, Spain through Grant ER36-20R to FENOL, and by MINECO and FEDER funds (Grant FIS2017-87519-P). The funders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Author contributions

G.F.A., M.T., and Y.M. conceived the study, performed the calculations, analysed and discussed the results, and wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42005-021-00525-3>.

Correspondence and requests for materials should be addressed to G.F.d.A.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021