

<https://doi.org/10.1038/s42005-024-01614-9>

Enhancing predictive accuracy in social contagion dynamics via directed hypergraph structures

Check for updates

Juyi Li¹, Xiaoqun Wu^{1,2}✉, Jinhu Lü^{3,4}✉ & Ling Lei¹

Evidence from both theoretical and empirical studies suggests that higher-order networks have emerged as powerful tools for modeling social contagions, such as opinion formation. In this article, we develop a model of social contagion on directed hypergraphs by considering the heterogeneity of individuals and environments in terms of reinforcing contagion effects. By distinguishing the directedness between nodes and hyperedges, we find that the bistable interval of the discontinuous phase transition decreases as the directedness strength decreases. Additionally, directed hypergraphs tend to generate bistable intervals when nodes with a large hyperdegree are more likely to adopt a specific opinion, as evidenced by simulations of directionality assignments for three sets of real networks. These findings provide two approaches to enhance the accuracy of predicting social contagion dynamics: one is to increase the stubbornness of all individuals, and the other is to prioritize increasing the stubbornness of highly influential individuals.

The study of complex networks has fundamentally transformed our understanding of various natural systems and social structures^{1–5}, providing a comprehensive framework to analyze interactions across diverse disciplines. These networks have been instrumental in shedding light on phenomena ranging from disease spreading^{6–8} and diffusion processes^{9–11} to the consensus of multi-agent systems^{12,13} and the formation of opinions^{14–16}. For instance, in epidemiology, networks model the spread of diseases¹⁷ by categorizing nodes as healthy or infected, thus facilitating the simulation of infection dynamics across populations. Such models have extended their applicability to the realm of social sciences, where they serve to understand social contagion processes, including the spread of opinions and behaviors through populations.

Despite their broad applications, traditional network models predominantly focus on pairwise interactions, a simplification that, while useful, often fails to capture the complexity inherent in many natural and social systems. These systems, exemplified by brain networks¹⁸, social networks¹⁹, and ecological networks²⁰, feature multi-body interactions that play a crucial role in their dynamics. The limitations of traditional models in accurately representing these interactions have prompted a growing interest among scholars in higher-order network structures^{21,22}, such as simplicial complexes^{23–25} and hypergraphs^{26–28}. These advanced models are better suited to capturing the multifaceted interactions

occurring within complex systems, offering a more nuanced understanding of their dynamics.

However, a notable gap in the current body of research is the predominantly symmetric structure of these higher-order interactions, which overlooks the potential for asymmetry^{29,30} and the unique roles that different nodes may play within the same interaction group. This oversight limits the models' ability to fully grasp the complexity of interaction dynamics, particularly in social systems where individual differences significantly influence the overall process.

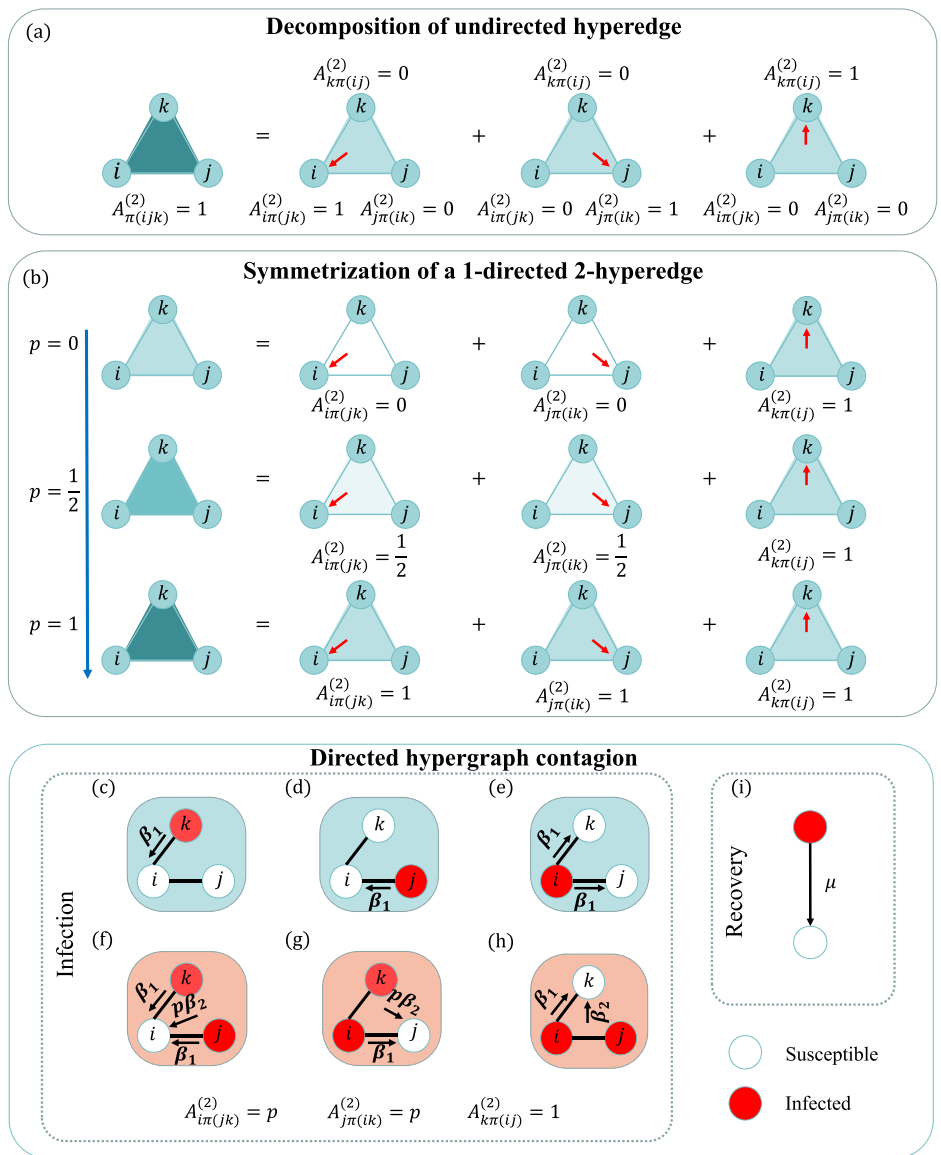
Addressing this gap, directed hypergraphs^{31,32} emerge as a powerful tool, enabling the differentiation between nodes within higher-order edges and reflecting the variability in individual roles and influences. This approach represents a significant advancement in modeling complex systems, particularly in the context of social contagion. Social contagion^{33,34}, akin to biological contagion³⁵ but involving the spread of ideas, behaviors, and norms, can be significantly affected by the directional influence of individuals within a network. Understanding these directional influences is crucial for accurately modeling and predicting the dynamics of social contagion.

In this work, we explore the impact of social contagion through the lens of M -directed hypergraphs³¹, an approach that allows for the detailed examination of directed properties assigned to both nodes and hyperedges.

¹School of Mathematics and Statistics, Wuhan University, Hubei, China. ²School of Computer Science and Software Engineering, Shenzhen University, Guangdong, China. ³School of Automation Science and Electrical Engineering, Beihang University, Beijing, China. ⁴Zhongguancun Laboratory, Beijing, China.

✉e-mail: xqwu@whu.edu.cn; jhlu@iss.ac.cn

Fig. 1 | The contagion process of M-directed hypergraphs. Panel **a** illustrates the decomposition of an undirected 2-hyperedge into three directed hyperedges. Panel **b** showcases a weighted 1-directed 2-hyperedge, which signifies a directed hyperedge when $p = 0$ and transitions back to an undirected 2-hyperedge as p increases to 1. Panels **c** through **(i)** present all contagion scenarios under a 1-directed 2-hyperedge, supplemented by two pairwise couplings. In Panels **c–h**, the squares represent higher-order interactions. The color scheme used is light blue to indicate that the higher-order interaction did not have an effect and orange to indicate that the higher-order interaction had an effect. In these panels, k is designated as the “point” node, while i and j serve as the “source” nodes. Lastly, Panel **i** indicates that all infected nodes revert to the susceptible state with a probability of μ at each time step. (Panels (a) and (b) are considered suitable adaptations of Fig. 1 and Fig. 5³¹).



By distinguishing between these two types of directedness, our model provides a richer understanding of how individual differences and the structure of interactions influence contagion processes. Through rigorous analysis, we investigate how variations in directedness affect the stability and accuracy of contagion dynamics, offering new insights into the underlying mechanisms of social spread. Our findings reveal that as the intensity of directedness decreases, the bistability interval of the discontinuous phase transition diminishes, suggesting that the accuracy of contagion processes improves with increased directedness among individuals. Furthermore, our analysis of real-world networks demonstrates that networks featuring nodes with high hyperdegree and directedness properties exhibit increased bistability intervals, highlighting the complex interplay between network structure and contagion dynamics.

Results

M-directed hypergraph

To introduce the concept of a directed hypergraph, we define a 1-directed d-hyperedge as a set of $d + 1$ nodes, where d nodes are “source” nodes collectively pointing towards the remaining one node³¹. This definition allows us to consider an undirected d-hyperedge as the union of $d + 1$ 1-directed d-hyperedges (see Fig. 1a). This decomposition framework is a natural extension of the pairwise interactions, where an undirected interaction edge

in a simple network can be decomposed into two directed interaction edges. With the directed hypergraph definition, we can classify nodes into two categories: “source” nodes and “point” nodes³¹. A 1-directed d-hyperedge, where the source nodes are denoted as j_1, j_2, \dots, j_d pointing to node i , can be represented as a tensor matrix $A^{(d)31}$:

$$A_{j_1 \dots j_d}^{(d)} = 1 \Rightarrow A_{in(j_1 \dots j_d)}^{(d)} = 1 \tag{1}$$

where $\pi(j_1 \dots j_d)$ represents any permutation of the indices j_1, \dots, j_d . Similarly, we can define m-directed d-hyperedges composed of $d + 1$ nodes, where $d + 1 - m$ source nodes point toward the remaining m nodes. Letting $s = d + 1 - m$, the tensor matrix for this type of hyperedge is given by³¹:

$$A_{\pi(j_1 \dots j_m) \pi'(j_1 \dots j_s)}^{(d)} = 1, \tag{2}$$

where $\pi(j_1 \dots j_m)$ and $\pi'(j_1 \dots j_s)$ represent any permutation of the indices j_1, \dots, j_m and j_1, \dots, j_s , respectively.

In this paper, we focus on weighted 1-directed 2-hyperedges. The weight is p , where $p \in [0, 1]$. The three nodes i, j , and k in the hypergraph network with weighted 1-directed 2-hyperedges form a hyperedge, where the source nodes are i and j . The elements in the tensor matrix A^2 are

represented as:

$$\begin{aligned} A_{ijk}^2 &= p & A_{jik}^2 &= p & A_{kij}^2 &= 1 \\ A_{ikj}^2 &= p & A_{jki}^2 &= p & A_{kji}^2 &= 1. \end{aligned} \tag{3}$$

As the value of p increases, the value of i and j pointing towards k direction will gradually increase. The hypergraph reverts to an undirected one when $p = 1$ (refer to Fig. 1b).

Contagion model in directed hypergraphs

In the original study of contagion models in higher-order networks, it was initiated with a simplicial complex¹⁵. This complex involves a d -simplex that infects a susceptible node with a probability of β_d only if the d -simplex contains a susceptible state and the remaining d nodes are in an infected state. Additionally, a d -simplex can also be affected by its lower-dimensional simplex with infectious probability $\beta_{d-1} \dots \beta_1$. For example, a 2-simplex where consider two nodes are in an infected state. The susceptible nodes within this 2-simplex have a probability of being infected equal to $2\beta_1 + \beta_2$. It takes into account both the triangular structure formed by three nodes and the contributions from the two edges.

The simplicial complex is a specific type case of hypergraphs, so our focus shifts to studying the contagion model on hypergraphs and its extension to directed hypergraphs. In this context, we introduce the weighted 1-directed 2-hyperedge consisting of three nodes i, j , and k , where i and j are the source nodes directed towards node k . The weights assigned to these hyperedges are denoted as p . The contagion process in hypergraphs is akin to that in simplicial complexes, with one crucial distinction: hypergraphs can represent subsets of the complete set within a simplicial complex. This paper explores a weighted directed hypergraph, where required weights are assigned to the probability of contagion. For example, if node i is susceptible and nodes j and k are infected, the probability of node i becoming infected in this hypergraph is $p\beta_2$. Similarly, if node j is susceptible, it is also infected with a probability of $p\beta_2$. In all other cases, the probability of infection remains β_2 . The recovery rate for all infected nodes is denoted as μ (refer to Fig. 1c–i).

Contagion model in directed uniform hypergraphs

To study the contagion process on directed uniform hypergraphs, we use the mean-field model. The directed uniform hypergraph maintains identical dimensions of the hyperedges and the same hyper degree for each node, with each node having an equal probability of being a “source” node within each hyperedge. First, we give the equations in the general 1-directed d -hyperedges of the uniformly weighted directed hypergraph. Define the set of infection probabilities as $B \equiv \{\beta_1, \dots, \beta_d\}$ and the recovery rate as μ . The hypergraph consists of N nodes, and the infection probability of node i at moment t is $x_i(t)$, where $x_i(t) \in [0, 1]$. At each moment t , the infection density is represented by the macroscopic parameter $\rho(t) = \frac{1}{N} \sum_{i=1}^N x_i(t)$. We can write the equation for the infection density $\rho(t)$ using the mean field method:

$$\rho(t + 1) = (1 - \mu)\rho(t) + \beta_d \langle k_d \rangle \rho(t)^d (1 - \rho(t)) \left(\frac{1}{d + 1} + \frac{d}{d + 1} p \right), \tag{4}$$

where $\langle k_d \rangle$ denotes the d -dimensional average hyperdegree of the node. Note that even in a more general hypergraph model, where hyperedges of each dimension are present, the assumptions made still hold. In this case, the hyperedges of different dimensions are all 1-directed i -hyperedges ($i > 1$). The equation is as follows:

$$\begin{aligned} \rho(t + 1) &= (1 - \mu)\rho(t) + \sum_{i=2}^d \beta_i \langle k_i \rangle \rho(t)^i (1 - \rho(t)) \left(\frac{1}{i + 1} + \frac{i}{i + 1} p \right) \\ &+ \beta_1 \langle k_1 \rangle \rho(t) (1 - \rho(t)), \end{aligned} \tag{5}$$

where the last term represents a pairwise interaction network, without higher order. If the higher order terms are removed, Eq. (5) degenerates to the standard Susceptible-Infectious-Susceptible (SIS) model.

We primarily focus on the case where $d = 2$. In Eq. (5), we simplify the equation to:

$$\begin{aligned} \rho(t + 1) &= (1 - \mu)\rho(t) + \beta_2 \langle k_2 \rangle \rho(t)^2 (1 - \rho(t)) \left(\frac{1}{3} + \frac{2}{3} p \right) \\ &+ \beta_1 \langle k_1 \rangle \rho(t) (1 - \rho(t)). \end{aligned} \tag{6}$$

To find the equilibrium point, we define $\lambda_1 = \frac{\beta_1 \langle k_1 \rangle}{\mu}$, $\lambda_2 = \frac{\beta_2 \langle k_2 \rangle (\frac{1}{3} + \frac{2}{3} p)}{\mu}$ and equate the right side of Eq. (6) to zero. By letting $\lim_{t \rightarrow +\infty} \rho(t) = \rho$, we obtain

$$\rho(\lambda_2 \rho^2 + (\lambda_1 - \lambda_2)\rho + (1 - \lambda_1)) = 0. \tag{7}$$

We get three solutions to this equation, which are $\rho_1 = 0$, $\rho_{2+} = \frac{-(\lambda_1 - \lambda_2) + \sqrt{(\lambda_1 - \lambda_2)^2 - 4\lambda_2(1 - \lambda_1)}}{2\lambda_2}$, and $\rho_{2-} = \frac{-(\lambda_1 - \lambda_2) - \sqrt{(\lambda_1 - \lambda_2)^2 - 4\lambda_2(1 - \lambda_1)}}{2\lambda_2}$. These solutions are meaningful only when $\rho \in [0, 1]$. Please refer to Supplementary Note 3 for using the function image analysis method, which differs from the analysis in the original article¹⁵.

To ensure that Eq. (7) has three solutions, the conditions $\lambda_2 > 1$ and $0 < \lambda_1^c < \lambda_1 < 1 < \lambda_2$ must be satisfied. The contagion process and the network parameters, such as setting $\langle k_1 \rangle$, $\langle k_2 \rangle$, β_2 , and μ are fixed as constants. When parameter conditions $p = 1$, $\lambda_2 > 1$ and $p = 0$, $0 < \lambda_2 < 1$ are met, a critical case of a discontinuous phase transition occurs. Specifically, when $p = 1$, the maximum phase transition interval is obtained, resulting in a bistable interval with two locally stable points ρ_1 and ρ_{2+} . However, as p gradually decreases, the bistable interval decreases as well. Eventually, there is a certain critical value of p where the phase transition disappears. This critical value for $\lambda_2 = 1$ is

$$p^* = \frac{3}{2} \frac{\mu}{\beta_2 \langle k_2 \rangle} - \frac{1}{2}, \tag{8}$$

where $\frac{\beta_2 \langle k_2 \rangle}{\mu} > 1$ and $\frac{\beta_2 \langle k_2 \rangle}{3\mu} < 1$ need to be satisfied. In the parameter settings of Fig. 2, it is ensured that $\lambda_2 > 1$ and $\frac{1}{3}\lambda_2 < 1$. From Fig. 2a, different initial values do not produce a phase transition interval due to $p < p^*$. However, when $p > p^*$, p increasing leads to an increase in the phase transition interval, as observed in Fig. 2b–d. In Fig. 3, the relationship between p , β_1 and ρ is plotted for different values of β_2 using Eq. (7). Fig. 3a, b correspond to $\beta_2 = 0.03$ and $\beta_2 = 0.01$, respectively. It is evident from comparing the two panels that p has little effect on the contagion threshold when β_2 is relatively small (i.e., when higher-order contagion is relatively weak). However, as β_2 gradually increases, the directedness parameter p begins to influence the contagion threshold. This corresponds to the increase in the phase transition interval observed in Fig. 2. Additionally, the stability of the solution in various cases is analyzed in Supplementary Note 3. The conclusions are further verified through simulations depicted in Supplementary Fig. 4.

We can draw an analogy between disease contagion and viewpoint contagion. From a viewpoint contagion process, we would expect the outcome to be more accurate. To achieve this, we need to have more nodes in the network that are either stubborn or have weak receptivity to information within each hyperedge. This is the only way to demonstrate the impact of directedness on the outcome of the contagion in a directed hypergraph. The accuracy of the contagion outcome is achieved by eliminating the bistable interval through the influence of directed connections.

Contagion model in k -adjacency-3-directed uniform hypergraphs

We need to discuss a crucial question of whether directedness is a characteristic of the node itself or the hyperedge. We illustrate the need to distinguish between the two kinds of directedness in Supplementary Note 1. If it is considered a property of the hyperedges, the directedness can be assigned

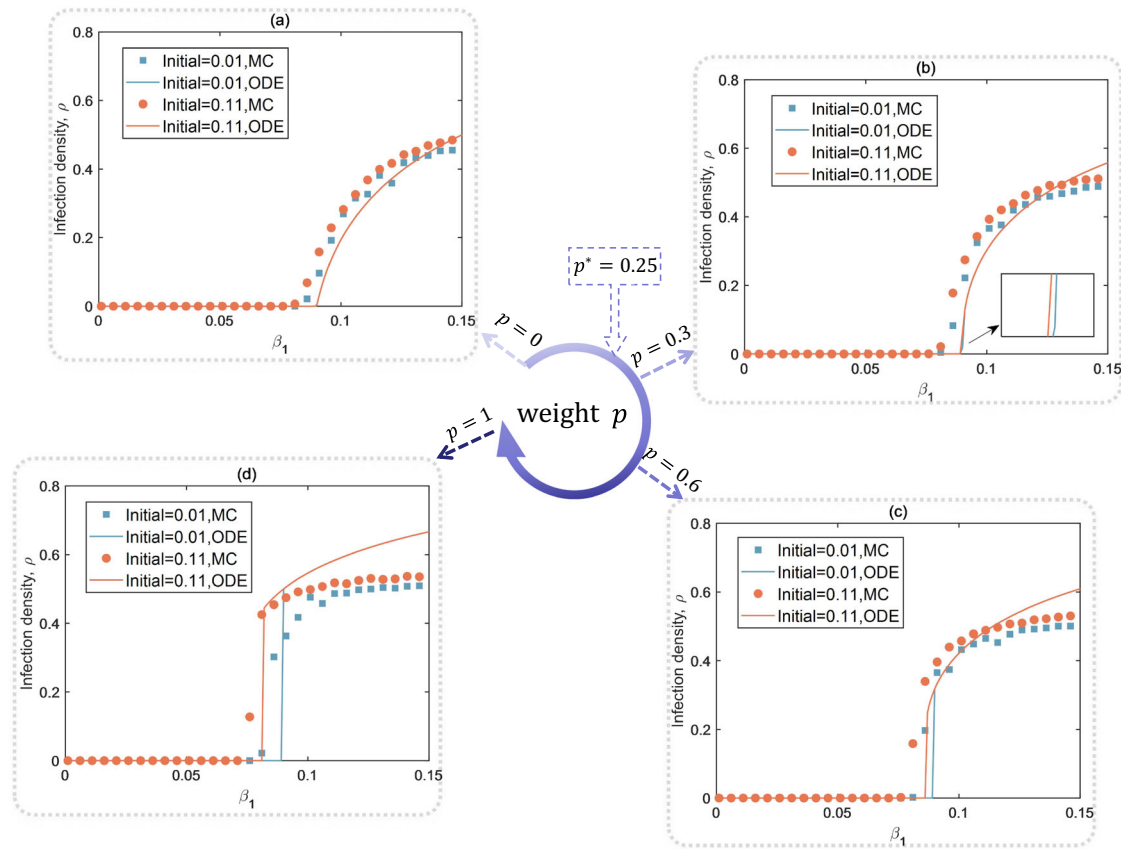


Fig. 2 | Directed hypergraph contagious mean-field model for $d = 2$ in the synthetic directed uniform hypergraph. The parameters μ , $\langle k_1 \rangle$, $\langle k_2 \rangle$, and β_2 are kept constant at values of 0.9, 10, 6, and 0.3, respectively. **a**, **b**, **c**, and **d** correspond to p -values of 0, 0.3, 0.6, and 1, respectively. Utilizing Eq. (8), we compute $p^* = 0.25$. The

solid line in the graph represents the theoretical result derived from Eq. (6), while the circles and squares illustrate the results obtained from Monte Carlo (MC) simulations. The colors blue and red are used to indicate the initial infection proportions of 0.01 and 0.11, respectively.

by probabilistic homogenization to obtain Eq. (6) when studying directed uniform hypergraphs. However, if the direction is a property of the node, the model needs to be reconstructed. Here, we will begin with the simplest case, focusing on k -adjacency-3-directed uniform hypergraphs. The properties of “source” and “point” are inherent to the nodes themselves and exhibit uniform characteristics within each hyperedge they participate in. Due to the characteristics of the k -adjacency-3-directed uniform hypergraphs, the hyperdegree of each node is k and the dimension of each hyperedge is 3 (see Fig. 4a, b). We can classify the nodes into two categories: ρ_{point} and ρ_{source} , denoting them as $\rho_1 = \rho_{\text{point}}$ and $\rho_2 = \rho_{\text{source}}$. It is necessary for each hyperedge to include one “point” and two “sources”. With this specification, we can formulate the equation based on the contagion model as follows:

$$\begin{cases} \rho_1(t+1) = (1-\mu)\rho_1(t) + k\beta_2(1-\rho_1(t))\rho_2(t)^2, \\ \rho_2(t+1) = (1-\mu)\rho_2(t) + p k \beta_2(1-\rho_2(t))\rho_1(t)\rho_2(t), \end{cases} \quad (9)$$

where k is the 3-dimensional hyperdegree of each node. To determine the equilibrium point of the equation, we make the right-hand side of Eq. (9) equal to 0:

$$\begin{cases} -\mu\rho_1 + k\beta_2(1-\rho_1)\rho_2^2 = 0, \\ \mu\rho_2 + p k \beta_2(1-\rho_2)\rho_1\rho_2 = 0. \end{cases} \quad (10)$$

Solving Eq. (10), one solution is $\rho_1 = 0$ and $\rho_2 = 0$. When neither ρ_1 nor ρ_2 is not 0, the second equation of Eq. (10) yields:

$$\rho_1 = \frac{\mu}{p k \beta_2(1-\rho_2)}. \quad (11)$$

By iteratively solving the above equations, the following conclusions have been obtained (please refer to Supplementary Note 4):

1. When $0 < p \leq \frac{\mu}{k\beta_2}$, $\rho = 0$ is the only equilibrium point of the equation in the interval $[0, 1]$.
2. When $\frac{\mu}{k\beta_2} < p < p^*$, $\rho = 0$ is the only equilibrium point of the equation in the interval $[0, 1]$.
3. When $p = p^*$, both $\rho = 0$ and $\rho = \frac{1}{3}\rho_1^* + \frac{2}{3}\rho_2^*$ are equilibrium points of the equation in the interval $[0, 1]$.
4. When $p^* < p < 1$, the equation has three solutions in the interval $[0, 1]$.

The value of p^* is obtained by solving the cubic equation for ρ^* in the interval $[0, 1]$. The critical point for the equation to undergo a phase transition is determined to be $p = p^*$. Since solving the cubic equation involving ρ^* depends on β_2 , it follows that p^* is also a function of β_2 . Fig. 5 illustrates the relationship between p , β_2 , and ρ . It is evident that when β_2 is relatively small, no matter how much p increases, transmission cannot occur even if the system reverts to an undirected hypergraph. However, as β_2 increases, the contagion threshold slowly starts to decrease. It only requires a small amount of directedness represented by p , to make the k -adjacency-3-directed hypergraph contagious.

The “point” node and the “source” node in the k -adjacency-3-directed uniform hypergraph represent the acceptant individual and the stubborn individual, respectively. A low directedness value p indicates that the stubborn individual is highly resistant to accepting the views of others. In this case, the viewpoint contagion scale is 0 for a given contagion rate β_2 . However, as the directedness p increases, the contagion scale of views gradually rises reaching a certain threshold of p^* , where stubborn individuals start to gradually accept certain information. Therefore, the presence of

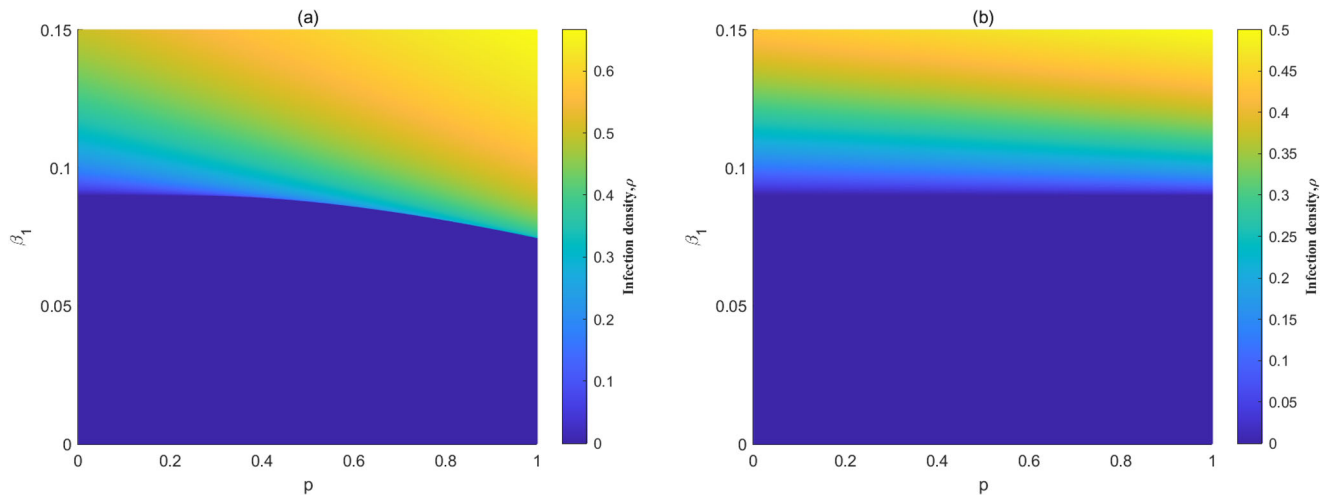
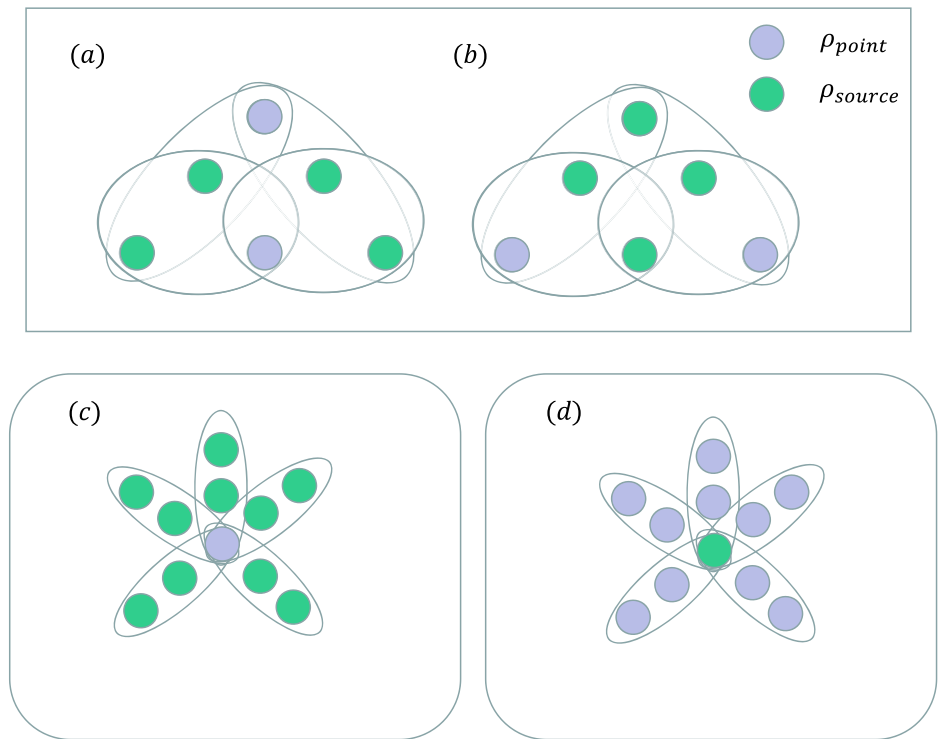


Fig. 3 | Influence of parameters on theoretical values in directed uniform hypergraphs. We discussed the correlation between infection density ρ , parameter p , and β_1 , with β_2 held constant. **a** and **b** present the results for parameters $\beta_2 = 0.3$ and

$\beta_2 = 0.1$, respectively. The value of ρ is derived from Eq. (7). Since Eq. (7) yields two results, we graphically represent the outcome for the stable ρ_{2+} .

Fig. 4 | The k -adjacency-3-directed uniform hypergraphs and directed hyperstar networks. Panels **a** and **b** depict two distinct cases of k -adjacency-3-directed uniform hypergraphs, both of which can be expressed using Eq. (9). Panel **c** designates the center node as the “point” and the leaf node as the “source”. Conversely, Panel **d** identifies the center node as the “source” and the leaf node as the “point”.



directedness in directed hypergraphs significantly influences the contagion process within the network.

The influence of directedness in the process of hypergraph contagion

To model the more general directed hypergraph network, we cannot use the mean-field approach. Instead, we introduce the general individual-based mean-field (IBMF) theory³⁶ and incorporate directedness to obtain the contagion model. Since the IBMF theory assumes independence between node states, it ignores correlations between nodes. Consequently, we can describe the directed hypergraph contagion process as follows:

$$\rho_i(t + 1) = (1 - \mu)\rho_i(t) + (1 - \rho_i(t))(1 - q_i^1(t)q_i^2(t) \dots q_i^d(t)). \quad (12)$$

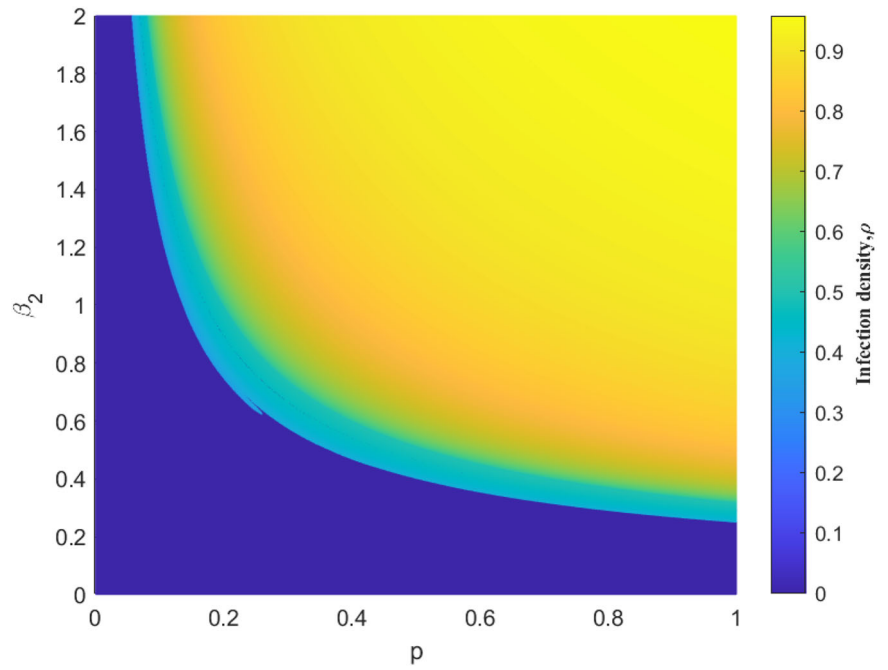
Here, $q_i^1(t)$ represents the probability that node i is not infected by any pairs of interacting neighbors:

$$q_i^1(t) = \prod_{j_1}^N (1 - \beta_1 a_{ij_1} \rho_{j_1}(t)). \quad (13)$$

Similarly, $q_i^2(t)$ represents the probability that node i is not infected by any of its 3 directed hyperedges:

$$q_i^2(t) = \prod_{\substack{j_1, j_2 \\ \{i, j_1, j_2\} \in N(t)}}^N (1 - \beta_2 a_{ij_1 j_2} \rho_{j_1}(t) \rho_{j_2}(t)), \quad (14)$$

Fig. 5 | The relationship between p , β_2 , and ρ . The value of ρ is computed using Eq. (10), which yields four solutions, one of which is zero. Among the remaining three solutions, at most two fall within the interval $[0, 1]$, as illustrated in Supplementary Note 4. The stability analysis in Supplementary Note 4 indicates that the solution with the maximum value is stable. If a numerical solution within the $[0, 1]$ interval exists, the maximum stable numerical solution is depicted.



where $N(i)$ denotes the 2-dimensional hyperedge associated with node i . Finally, $q_i^d(t)$ indicates the probability that node i is not infected by any of its $d + 1$ directed hyperedges:

$$q_i^d(t) = \prod_{\substack{j_1, j_2, \dots, j_d \\ \{j_1, j_2, \dots, j_d\} \in N(i)}} (1 - \beta_d a_{ij_1 \dots j_d}(t) \rho_{j_1}(t) \rho_{j_2}(t) \dots \rho_{j_d}(t)). \quad (15)$$

In Eqs. (13), (14), and (15), $a_{ij_1 \dots j_d}$ represents the elements in the d -dimensional tensor, while $N(i)$ denotes the d -dimensional hyperedge associated with node i . Additionally, it is necessary to order the node coordinates using concatenation numbers to avoid redundant calculations of hyperedges when the results are concatenated. If the direction is considered a node property, then for “point” node i , $a_{ij_1 \dots j_d} = \begin{cases} 1, & \{i, j_1, \dots, j_d\} \in N(i) \\ 0, & \{i, j_1, \dots, j_d\} \notin N(i) \end{cases}$ and for “source” node i , $a_{ij_1 \dots j_d} = \begin{cases} p, & \{i, j_1, \dots, j_d\} \in N(i) \\ 0, & \{i, j_1, \dots, j_d\} \notin N(i) \end{cases}$. On the other hand, if the direction is treated as a hyperedge property, the elements of the tensor matrix positions corresponding to “source” nodes and “point” nodes in a hyperedge are assigned the values p and 1, respectively. Furthermore, if the correlation between nodes is taken into account, the contagion process can be accurately simulated using the epidemic link equations method³⁷.

We applied the aforementioned method to examine the impact of directedness in a specific hypergraph called the **hyperstar network**, also known as the sunflower network³⁸. In this study, we set the dimension of each hyperedge to be 3, with each hyperedge associated with a central node. For our analysis, consider a hyperstar network consisting of r hyperedges, resulting in a total of $N = 2r + 1$ nodes. To observe the effect of directedness, we focus on two different hyperedges within the network. In one case, the central node serves as the “point” node, while the leaf nodes act as the “source” nodes. In the other case, the center is considered the “source” node, and the leaf nodes are the “point” nodes. This configuration can be visualized in Fig. 4c, d.

In the first case, where the center node is the “point” and the leaf nodes are the “sources”, we can categorize the nodes into two groups: the center

node denoted as ρ_c , and the leaf nodes denoted as ρ_l . The contagion equation in this case can be expressed as:

$$\begin{cases} \rho_c(t + 1) = (1 - \mu)\rho_c(t) + (1 - \rho_c(t))[1 - (1 - \beta_2 \rho_l^2(t))^r], \\ \rho_l(t + 1) = (1 - \mu)\rho_l(t) + (1 - \rho_l(t))[1 - (1 - \beta_2 p \rho_c(t) \rho_l(t))]. \end{cases} \quad (16)$$

Equation (16) is a high-dimensional equation, and our objective is to study the effect of directedness near the contagion threshold. To achieve this, we perform a Taylor expansion of the equation around 0 and obtain the following conclusions through iterations (refer to Supplementary Note 5):

1. When $0 < p \leq \frac{\mu}{\beta_2}$, $\rho = 0$ is the only equilibrium point of the equation in the interval $[0, 1]$.
2. When $\frac{\mu}{\beta_2} < p < p^*$, $\rho = 0$ is still the only equilibrium point of the equation in the interval $[0, 1]$.
3. When $p = p^*$, both $\rho = 0$ and $\rho = \frac{1}{2r+1} \rho_c^* + \frac{2r}{2r+1} \rho_l^*$ are equilibrium points of the equation in the interval $[0, 1]$.
4. When $p^* < p < 1$, the equation has three solutions within the interval $[0, 1]$.

In the second case, we can analyze the case where the leaf nodes are considered the “points” and the center nodes are the “sources” in the hyperstar network. Similarly, we can express the equation as follows:

$$\begin{cases} \rho_c(t + 1) = (1 - \mu)\rho_c(t) + (1 - \rho_c(t))[1 - (1 - \beta_2 p \rho_l^2(t))^r], \\ \rho_l(t + 1) = (1 - \mu)\rho_l(t) + (1 - \rho_l(t))[1 - (1 - \beta_2 \rho_c(t) \rho_l(t))]. \end{cases} \quad (17)$$

Based on our analysis (refer to Supplementary Note 5), we can draw the following conclusions:

1. When $0 < p < p^*$, $\rho = 0$ is the only equilibrium point of the equation in the interval $[0, 1]$.
2. When $p = p^*$, both $\rho = 0$ and $\rho = \frac{1}{2r+1} \rho_c^* + \frac{2r}{2r+1} \rho_l^*$ are equilibrium points of the equation in the interval $[0, 1]$.
3. When $p^* < p < 1$, the equation has three solutions within the interval $[0, 1]$.

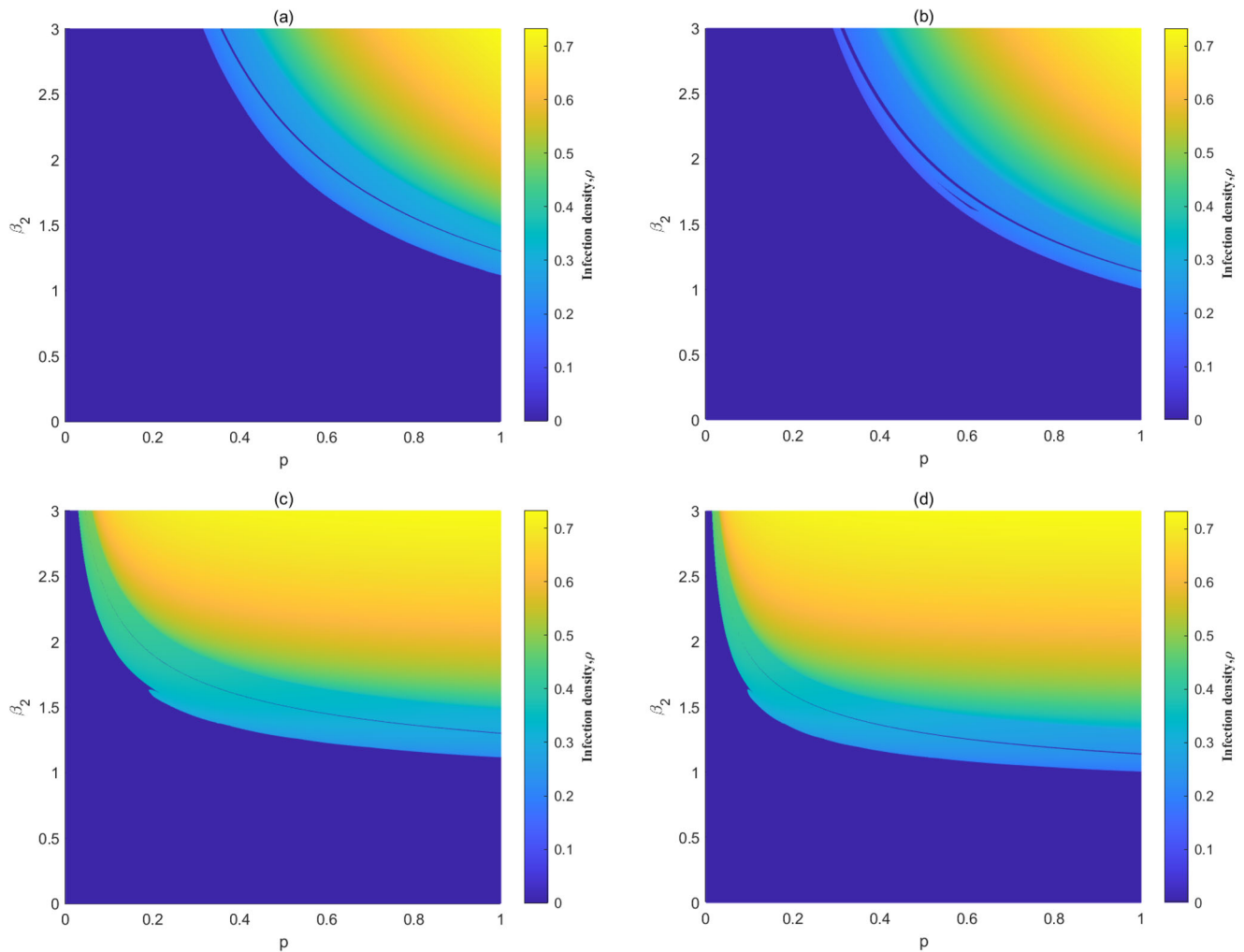


Fig. 6 | The correlation between infection density ρ , parameter p , β_2 , and r . Panels **a** and **b** represent hyperstar networks where the center node serves as the “point” node. Panels **c** and **d** represent hyperstar networks where the center node serves as

the “source” node. The number of petals r in panels **a** and **c** is set to 20, and in panels **b** and **d** is set to 40. These panels depict the stable solutions derived through numerical computation using Taylor expansion from Eqs. (16) and (17).

In Fig. 6, by comparing the two cases mentioned above and fixing the parameter β_2 , we observe that the critical threshold p_a^* for Fig. 6a is much smaller than the critical threshold p_c^* for Fig. 6c. A hyperstar network can be seen as analogous to a scientific research team in reality. The central node represents the team leader who has multiple research directions, with each direction being a hyperedge. The researchers working on those research directions are represented by the leaf nodes. In Fig. 6a, the center node is considered as the “point” node, indicating the team leader is an open-minded person who accepts different points (acceptant person). The other researchers represented by the leaf nodes, are depicted as stubborn individuals. For an opinion to spread contagiously, all stubborn researchers must have a high level of information receptivity. On the other hand, in Fig. 6c, the central node is considered the “source” node, representing a scenario where a team leader is a stubborn person, and the other researchers are open to accepting different opinions (acceptant people). In this case, only takes a small level of acceptance from the stubborn leader for a viewpoint to spread. This is demonstrated by comparing Fig. 6a, c, with the parameters $\beta_2 = 2$ and $p = 0.3$ held constant. It is evident that when the leaf nodes are stubborn, the viewpoints are not contagious, whereas when the center node is stubborn, the viewpoints become contagious. Thus, it can be inferred that in both scenarios, the stubbornness of the center node has a lesser impact on the overall size of the contagion compared to the stubbornness

of all leaf nodes. Furthermore, a comparison of Fig. 6a, b, as well as Fig. 6c, d, reveals that in both scenarios, with all parameters held constant, the infection density ρ increases as the number of petals r increases.

Contagion in realistic directed hypergraph networks

To study the contagion process in directed hypergraphs, we analyzed three different sets of social network structures. We processed the data collected from the three sets of networks. First, the pairwise coupling relationship of the networks was extended to higher-order interactions. Second, a specific classification on the assignment of directedness was made into two broad categories: node directedness and edge directedness (see data addition with directedness of the “Methods” section). The numerical simulation results were obtained by the IBMF theory algorithm.

The network structures depicted in Fig. 7 have not been extracted yet. Here, we only depict the low-order network structure of pairwise interactions, without incorporating the higher-order network structure. Please refer to the data description and processing of the “Methods” section for details on how to integrate higher-order network structures. Nevertheless, we can still observe certain distinctions among the three networks. Both the Vote and Infect-dublin networks exhibit significant clustering effects, whereas the Netscience network lacks such clustering effects. The observation of the clustering effect causes a notable

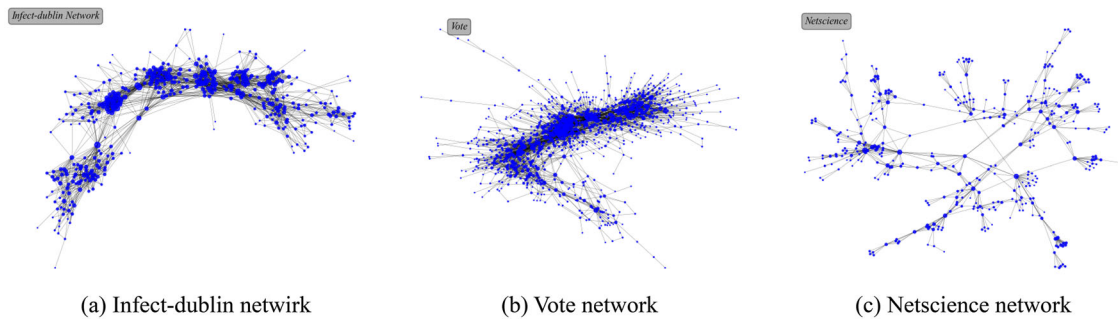


Fig. 7 | Three realistic network structures. The Infect-dublin network in panel **a** demonstrates a noticeable clustering effect. Similarly, the Vote network in panel **b** exhibits clustering, highlighting the significance of network structure in understanding contagion dynamics. Conversely, the Netscience network in panel **c** does not show a significant clustering effect. These networks represent unexplored structures in terms of higher-order variations. The specific methods used in this

study are described in the data description and processing of the “Methods” section. However, the visual analysis of these network structures reveals that Infect-dublin and Vote networks exhibit noticeable clustering effects, whereas Netscience does not. This distinction becomes more significant when considering higher-order contagion effects in subsequent analyses.

difference between the IBMF theoretical results and the actual MC results in terms of theoretical simulation. In addition, we observe that the simulation is not very near the threshold value effective. This might be caused due to the fact that IBMF theory ignores the correlation between nodes. Nonetheless, for contagion scale simulations with relatively large parameters of β_1 , the theory remains relatively accurate.

In the following analysis, we focus on comparing the effects of different choices of directedness on the contagion threshold and the phase transition interval. We examine this comparison using the data from Figs. 8, 9 and 10. First, vertically, we find that minimizing the phase transition interval is possible by increasing the probability of the nodes with a small hyperdegree becoming a “point”. Conversely, maximizing the phase transition interval can be achieved by increasing the probability of the nodes with large hyperdegree becoming a “point”. In realistic viewpoint propagation, if we aim for accuracy, it is necessary to make nodes with high hyperdegree (nodes with high influence) slightly more stubborn. Simply reducing their receptiveness to information will ensure accuracy in the contagion scale. Secondly, observing the simulation horizontally across various real networks, we note that there is an increase in the phase transition interval with higher directedness intensity p . Therefore, to ensure accuracy in the contagion process, it is recommended to decrease the directedness intensity p .

Conclusion

In summary, this paper expands on the original hypergraph contagion model and introduces the directed hypergraph contagion model. Firstly, the uniform directed hypergraph contagion model is studied, and the critical situation of phase transition is analyzed using the image method. It is concluded that the size of the discontinuous phase transition interval gradually increases as the value of directedness is enhanced. Secondly, directedness is classified into two specific types: node properties and hyperedge properties. They are discussed on two different types of networks. It is found that directedness belonging to different property cases has a certain degree of influence on both the contagion process and the criticality of phase transition. Finally, we compare three different real network data sets and find that directedness p can affect the phase change interval. The phase transition interval increases when the probability of nodes with a large hyperdegree becoming a “point” is greater. Conversely, the phase transition interval becomes smaller when the probability of becoming a “point” is higher for nodes with a smaller hyperdegree. The above theoretical and real network simulations lead to the conclusion that there are two approaches if we want the viewpoint contagion process to be more accurate. The first one is to make all the nodes stubborn in the network. However, this situation may not be very satisfying. The second approach is to make the nodes with

higher-order influence in the network more stubborn. This approach makes the contagion process of the network more accurate.

Methods

Data description and processing

We considered three different datasets representing various human interaction scenarios: Infect-dublin³⁹, Vote⁴⁰, and Netscience⁴¹. The Infect-dublin dataset represents the human physical contact network, Vote contains voting data from Wikipedia since its inception until January 2008, and Netscience represents co-author relations among scientists. These datasets are considered simple networks without higher-order structures. To introduce higher-order effects, we extract all triangle interactions and represent them as a correlation matrix. Moreover, we focus on the effect of higher-order vectoriality, so we remove nodes that are not involved in higher-order interactions. The connectivity properties of these datasets are summarized in Table 1. $\langle k_1 \rangle$ denotes the average degree of pairwise connectivity, $\langle k_2 \rangle$ represents the average hyperdegree of second-order connectivity, k_2^{\max} indicates the maximum second-order hyperdegree, k_2^{\min} denotes the minimum second-order hyperdegree, and k_2^{number} signifies the total number of second-order hyperedges in the directed hypergraph. We tabulate the generality distributions for the three datasets in Supplementary Fig. 2 of Supplementary Note 2.

Data addition with directedness

Our research specifically focuses on directedness. The article categorizes directedness into two main types: node directedness and edge directedness. In analyzing the actual network, we adhere to the setup detailed in the article.

Directedness of nodes. How different nodes are a key aspect we are observing to examine. Directedness probabilities of node directedness can lead to various outcomes in the contagion process. Specifically, there are three subcategories:

1. Equal Probability: The probability of each node being a “point” or a “source” is set to be 0.5. In this case, all nodes have an equal chance of having either role within each hyperedge.
2. High Probability of “Point”: Nodes have a higher probability of becoming a “point” with a high hyperdegree. We model this using an exponential distribution:

$$f(x) = e^{-x}, \tag{18}$$

where x represents the hyperdegree of the node. The lower the value of x , the higher the probability. This is achieved by ordering the sequence of hyperdegrees in the higher-order network in descending order. The

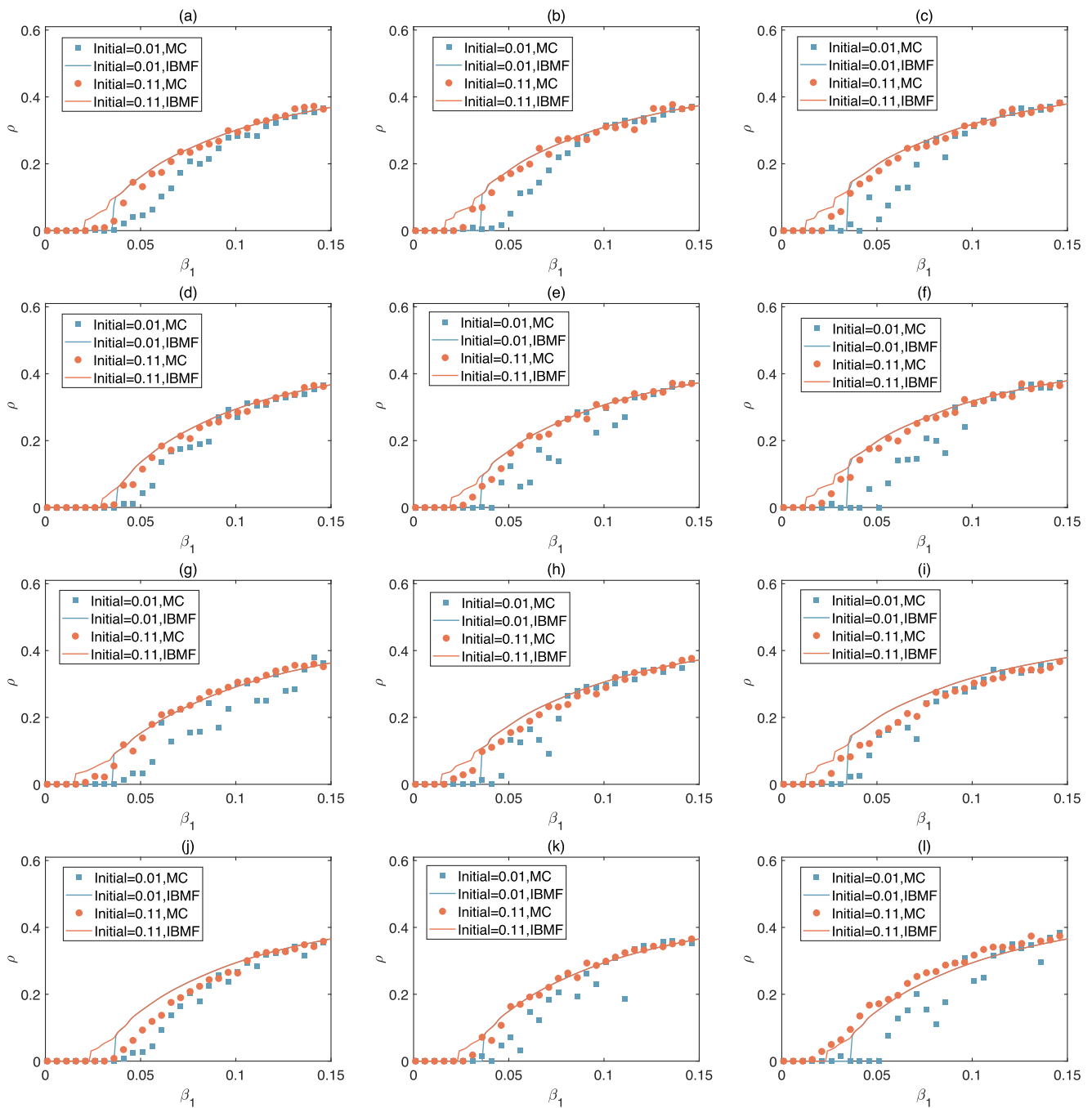


Fig. 8 | A comparison between Monte Carlo (MC) results and individual-based mean-filed (IBMF) theoretical results within the Infect-dublin network.

a–c Demonstrate scenarios of equal probabilities of node directedness from data addition with directedness of the “Methods” section, with directedness parameters at $p = 0.4$, $p = 0.7$, and $p = 1$, respectively. **d–f** Depict the low probability of “point”, while **g–i** represent the high probability of “point”, both settings derived from the

same section. **j–l** Focus on the directedness of edges. The parameter p is consistently maintained across vertical counterparts, meaning **a, d, g, j** corresponds to $p = 0.4$; **b, e, h, k** to $p = 0.7$; and **c, f, i, l** to $p = 1$. The red circles and blue squares indicate Monte Carlo simulations with initial infection ratios of 0.11 and 0.01, respectively, complemented by solid red and solid blue lines showing theoretical IBMF values for these ratios.

probability of a selection with the max hyperdegree becoming a “point” is set to $e^{-0.01}$. Subsequently, the interval between the sequences of hyperdegrees arranged in descending order is set to 0.02. For instance, the probability of a node with a hyperdegree of 1 becoming a “point” is set to $e^{-\infty} \approx 0$.

3. Low Probability of “Point”: Nodes have a lower probability of becoming a “point” with a high hyperdegree. Using a similar exponential distribution as above, we obtain the desired probability values.

The specific procedure is akin to the one described above, requiring only the arrangement of the hyperdegree’s degree sequence in ascending order. Subsequently, the probability of a node with a hyperdegree of 1 becoming a “point” is set to $e^{-0.01}$. All following operations remain consistent with this approach.

It is important to note that if we specify directedness as a node property, not require all hyperedges need to have the same type, such as one “point”

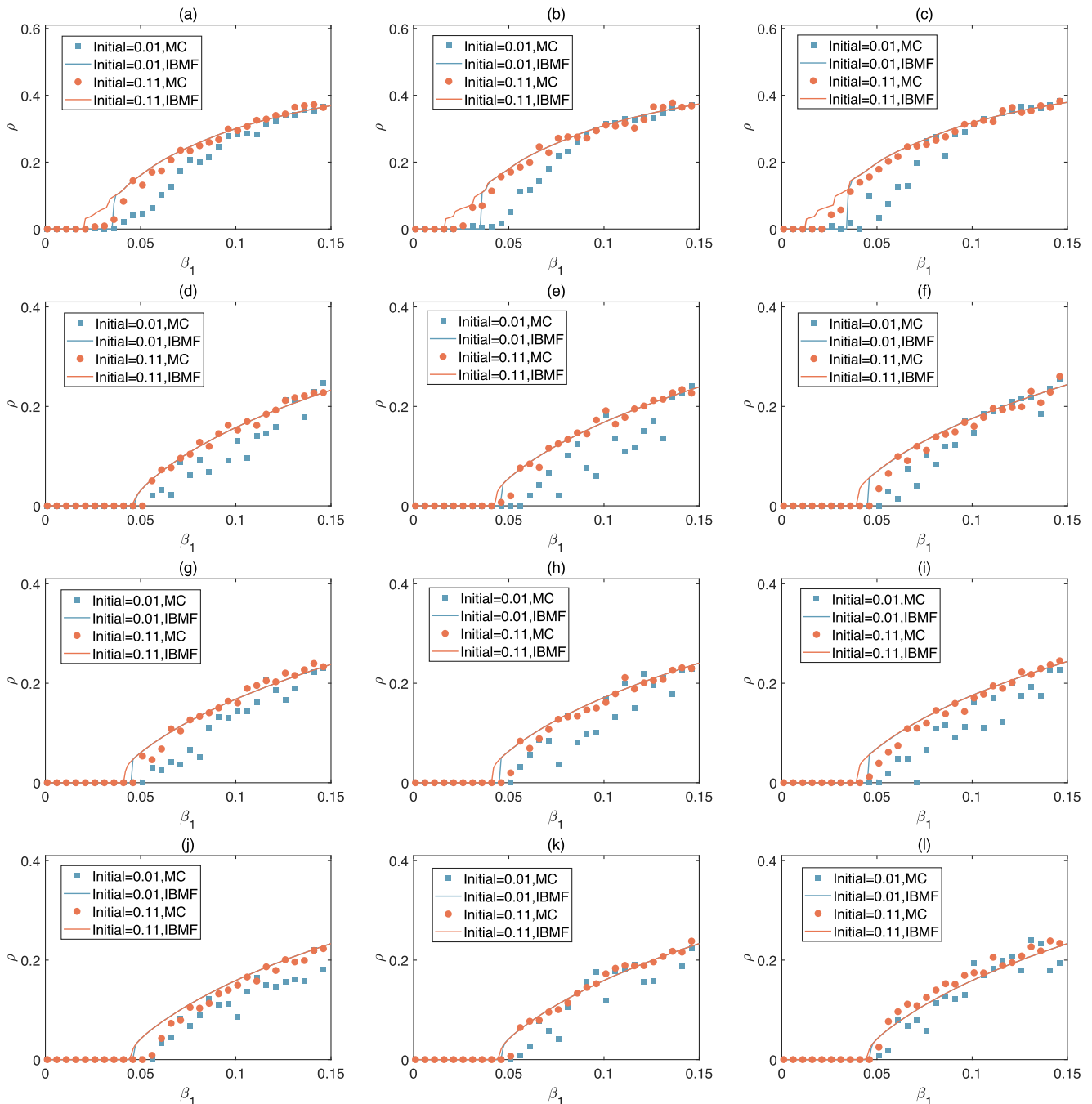


Fig. 9 | A comparison between Monte Carlo (MC) results and individual-based mean-field (IBMF) theoretical results within the Vote network. a–l Correspond directly to their counterparts in Fig. 8. The only difference lies in the network structure.

and two “sources” or two “points” and one “source”. Although the rule graph is guaranteed in the main text, a general hypergraph may not always be fully realized. This may also present a limitation of the node property. The edge property can further contribute to exploring this aspect.

Directedness of edges. The edge property is that each node in a hyperedge will exhibit its property, and nodes in different hyperedges may exhibit different properties. We investigate the probability that any node in a hyperedge becomes a “point” is 0.5.

Monte Carlo simulation process

The primary challenge in higher-order Monte Carlo simulations involves simplifying the tensor matrix associated with higher-order networks. We

suggest a simple method for this reduction before simulation. The first step is to remove duplicate hyperedges in each dimension. In an undirected hypergraph, a hyperedge is represented by six tensor elements, whereas in a directed hypergraph, it’s represented by two. However, in the contagion process, only one contagion is possible per hyperedge at any given moment. Therefore, the effect of duplicate hyperedges in the tensor must be addressed. Next, we reorganize the three-dimensional tensor matrix. For example, an $N * N * N$ tensor matrix is restructured into $N * N^2$ dimensions, where each $N * N$ matrix dimension in the i -dimension represents the hyperedge connected to the i th node. We then convert each $N * N$ matrix dimension into a $1 * N^2$ matrix to aid further simulation steps. As a result, we get an $N * N^2$ matrix, with the i th row indicating all

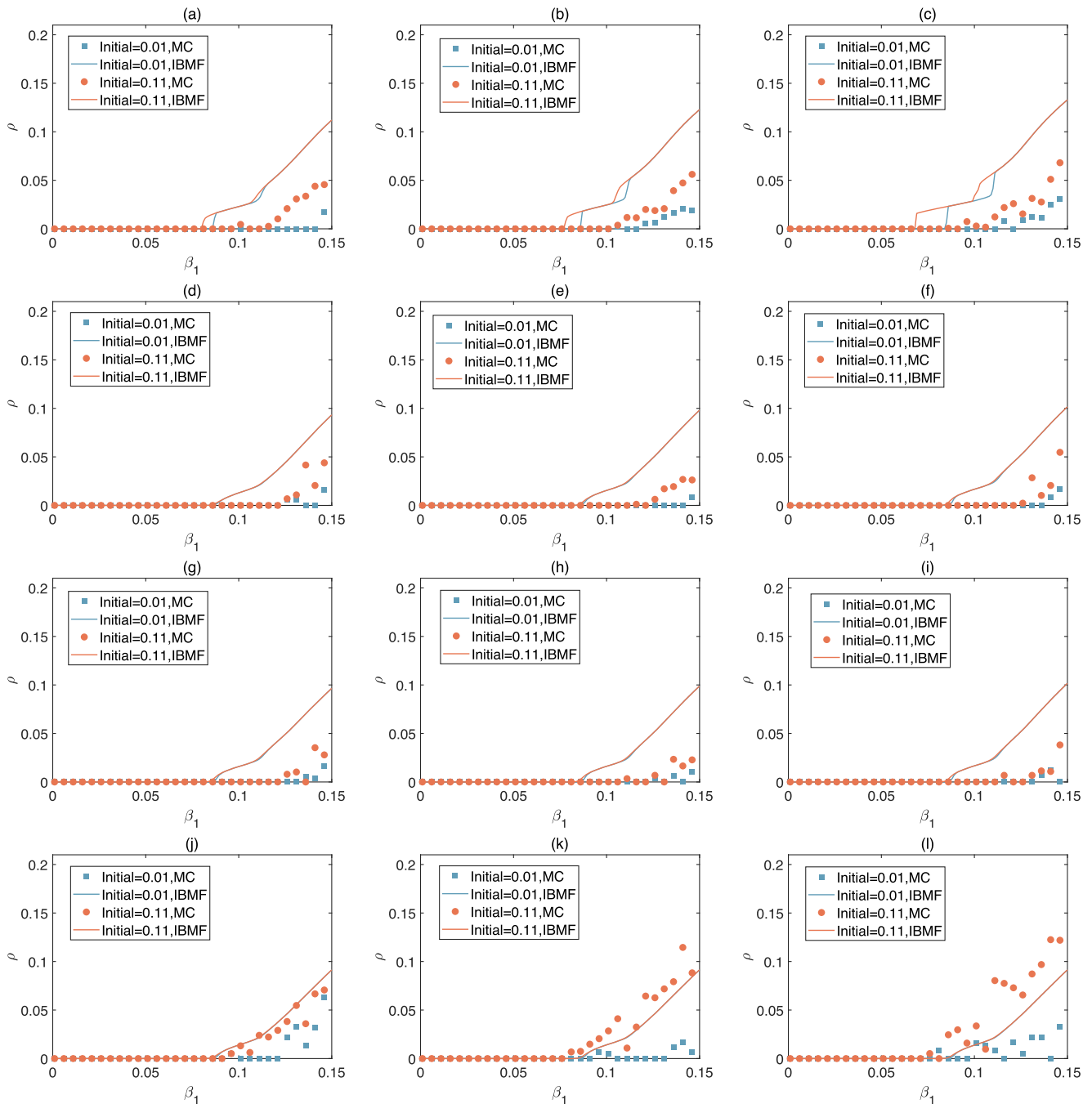


Fig. 10 | A comparison between Monte Carlo (MC) results and individual-based mean-field (IBMF) theoretical results within the Netscience network. a–l Correspond directly to their counterparts in Fig. 8. The only difference lies in the network structure.

Table 1 | Statistics of real-world directed hypergraph

Data set	context	$\langle k_1 \rangle$	$\langle k_2 \rangle$	k_2^{\max}	k_2^{\min}	k_2^{number}
Infect-Dublin	Social network	13.3503	108.0914	556	2	14196
Vote	Voting data	7.9590	25.8504	502	1	4205
Netscience	Collaboration network	4.9972	13.7521	136	1	1609

$\langle k_1 \rangle, \langle k_2 \rangle$: average degrees of pairwise and second-order connectivity;
 k_2^{\max}, k_2^{\min} : maximum and minimum second-order hyperdegrees;
 k_2^{number} : total number of second-order hyperedges.

hyperedges linked to the i th node. The Monte Carlo simulation is then performed using the methods detailed in Eqs. (12), (13), and (14). The key difference in a higher-order network is that node i can only be infected if the other two nodes in its hyperedges are also infected and

susceptible. The detailed Monte Carlo code is available in Supplementary Note 6.

Data availability

The SocioPatterns data sets were downloaded from <https://networkrepository.com/socfb>.

Code availability

Some of the codes in the manuscript have been placed in the Supplementary Information. The rest of the numerical simulation code is available upon request from the corresponding author.

Received: 15 September 2023; Accepted: 30 March 2024;

Published online: 15 April 2024

References

- Horsevad, N., Mateo, D., Kooij, R. E., Barrat, A. & Bouffanais, R. Transition from simple to complex contagion in collective decision-making. *Nat. Commun.* **13**, 1442 (2022).
- Tang, L., Wu, X., Lü, J., Lu, J.-a. & D'Souza, R. M. Master stability functions for complete, intralayer, and interlayer synchronization in multiplex networks of coupled rossler oscillators. *Phys. Rev. E* **99**, 012304 (2019).
- Li, N., Wu, X., Feng, J., Xu, Y. & Lü, J. Fixed-time synchronization of coupled neural networks with discontinuous activation and mismatched parameters. *IEEE Trans. Neural Netw. Learn. Syst.* **32**, 2470–2482 (2021).
- Wu, X. et al. Synchronization in multiplex networks. *Phys. Rep. Rev. Sec. Phys. Lett.* **1060**, 1–54 (2024).
- Fan, Z., Wu, X., Mao, B. & Lü, J. Output discernibility of topological variations in linear dynamical networks. *IEEE Trans. Automat. Control* **1–8** (2024).
- St-Onge, G., Sun, H., Allard, A., Hébert-Dufresne, L. & Bianconi, G. Universal nonlinear infection kernel from heterogeneous exposure on higher-order networks. *Phys. Rev. Lett.* **127**, 158301 (2021).
- Wei, X., Wu, X., Chen, S., Lu, J.-a. & Chen, G. Cooperative epidemic spreading on a two-layered interconnected network. *SIAM J. Appl. Dyn. Syst.* **17**, 1503–1520 (2018).
- Zhong, S., Wu, X., Li, Y. & Liu, C. Indirect transmission and disinfection strategies on heterogeneous networks. *Phys. Rev. E* **106**, 054309 (2022).
- Gómez, S. et al. Diffusion dynamics on multiplex networks. *Phys. Rev. Lett.* **110**, 028701 (2013).
- Tejedor, A., Longjas, A., Fofoula-Georgiou, E., Georgiou, T. T. & Moreno, Y. Diffusion dynamics and optimal coupling in multiplex networks with directed layers. *Phys. Rev. X* **8**, 031071 (2018).
- Wei, J., Wu, X., Lu, J.-a., Lü, J. & Chen, G. A topological mechanism of superdiffusion on duplex networks. *IEEE Trans. Control Netw. Syst.* **10**, 556–563 (2023).
- Wu, X., Mao, B., Wu, X. & Lü, J. Dynamic event-triggered leader-follower consensus control for multiagent systems. *SIAM J. Control Opt.* **60**, 189–209 (2022).
- Mao, B., Wu, X., Lü, J. & Chen, G. Predefined-time bounded consensus of multiagent systems with unknown nonlinearity via distributed adaptive fuzzy control. *IEEE T. Cybern.* **53**, 2622–2635 (2022).
- Jia, P., MirTabatabaei, A., Friedkin, N. E. & Bullo, F. Opinion dynamics and the evolution of social power in influence networks. *SIAM Rev.* **57**, 367–397 (2015).
- Iacopini, I., Petri, G., Barrat, A. & Latora, V. Simplicial models of social contagion. *Nat. Commun.* **10**, 2485 (2019).
- Dong, Y., Ding, Z., Martínez, L. & Herrera, F. Managing consensus based on leadership in opinion dynamics. *Inf. Sci.* **397–398**, 187–205 (2017).
- Wang, Z., Andrews, M. A., Wu, Z.-X., Wang, L. & Bauch, C. T. Coupled disease-behavior dynamics on complex networks: a review. *Phys. Life Rev.* **15**, 1–29 (2015).
- Giusti, C., Pastalkova, E., Curto, C. & Itskov, V. Clique topology reveals intrinsic geometric structure in neural correlations. *Proc. Natl. Acad. Sci.* **112**, 13455–13460 (2015).
- Benson, A. R., Gleich, D. F. & Leskovec, J. Higher-order organization of complex networks. *Science* **353**, 163–166 (2016).
- Bairey, E., Kelsic, E. D. & Kishony, R. High-order species interactions shape ecosystem diversity. *Nat. Commun.* **7**, 12285 (2016).
- Battiston, F. et al. Networks beyond pairwise interactions: structure and dynamics. *Phys. Rep.* **874**, 1–92 (2020).
- Grilli, J., Barabas, G., Michalska-Smith, M. J. & Allesina, S. Higher-order interactions stabilize dynamics in competitive network models. *Nature* **548**, 210 (2017).
- Matamalas, J. T., Gómez, S. & Arenas, A. Abrupt phase transition of epidemic spreading in simplicial complexes. *Phys. Rev. Res.* **2**, 012049 (2020).
- Petri, G. & Barrat, A. Simplicial activity driven model. *Phys. Rev. Lett.* **121**, 228301 (2018).
- Gambuzza, L. V. et al. Stability of synchronization in simplicial complexes. *Nat. Commun.* **12**, 1255 (2021).
- Ferraz de Arruda, G., Tizzani, M. & Moreno, Y. Phase transitions and stability of dynamical processes on hypergraphs. *Commun. Phys.* **4**, 24 (2021).
- Contisciani, M., Battiston, F. & De Bacco, C. Inference of hyperedges and overlapping communities in hypergraphs. *Nat. Commun.* **13**, 7229 (2022).
- Zhang, Y., Lucas, M. & Battiston, F. Higher-order interactions shape collective dynamics differently in hypergraphs and simplicial complexes. *Nat. Commun.* **14**, 1605 (2023).
- Piketty, T. & Saez, E. Inequality in the long run. *Science* **344**, 838–843 (2014).
- Scheffer, M., van Bavel, B., van de Leemput, I. A. & van Nes, E. H. Inequality in nature and society. *Proc. Natl. Acad. Sci. USA* **114**, 13154–13157 (2017).
- Gallo, L. et al. Synchronization induced by directed higher-order interactions. *Commun. Phys.* **5**, 263 (2022).
- Gallo, G., Longo, G., Pallottino, S. & Nguyen, S. Directed hypergraphs and applications. *Discrete. Appl. Math.* **42**, 177–201 (1993).
- de Arruda, G. F., Petri, G. & Moreno, Y. Social contagion models on hypergraphs. *Phys. Rev. Res.* **2**, 023032 (2020).
- Kramer, A. D. I., Guillory, J. E. & Hancock, J. T. Experimental evidence of massive-scale emotional contagion through social networks. *Proc. Natl. Acad. Sci.* **111**, 8788–8790 (2014).
- Wang, W., Liu, Q.-H., Liang, J., Hu, Y. & Zhou, T. Coevolution spreading in complex networks. *Phys. Rep.* **820**, 1–51 (2019).
- Gómez, S., Arenas, A., Borge-Holthoefer, J., Meloni, S. & Moreno, Y. Discrete-time markov chain approach to contact-based disease spreading in complex networks. *Europhys. Lett.* **89**, 38009 (2010).
- Matamalas, J. T., Arenas, A. & Gómez, S. Effective approach to epidemic containment using link equations in complex networks. *Sci. Adv.* **4**, eaau4212 (2018).
- Benson, A. R. Three hypergraph eigenvector centralities. *SIAM J. Math. Data Sci.* **1**, 293–312 (2019).
- SocioPatterns. *Infectious contact networks*, <http://www.sociopatterns.org/datasets/>.
- Leskovec, J., Huttenlocher, D. & Kleinberg, J. Signed networks in social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, <https://doi.org/10.1145/1753326.1753532> (2010).
- Newman, M. E. Finding community structure in networks using the eigenvectors of matrices. *Phys. Rev. E* **74**, 036104 (2006).

Acknowledgements

This work was supported by the Training Program of the Major Research Plan of the National Natural Science Foundation of China under Grant 92267101 and the National Natural Science Foundation of China under Grants 61973241 and 62141604.

Author contributions

Juyi Li, Xiaoqun Wu, Jinhu Lü, and Ling Lei designed the research; Juyi Li performed the research; Juyi Li analyzed data; Juyi Li wrote the supporting information; Juyi Li 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/s42005-024-01614-9>.

Correspondence and requests for materials should be addressed to Xiaoqun Wu or Jinhu Lü.

Peer review information *Communications Physics* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available

Reprints and permissions 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024