Supplementary Information for Transformability of Higher-order Network Dynamics

Ming Xie[1], Shibo He[1], Aming Li[2], Zi-Ke Zhang[3], Qihao Huang[1], Youxian Sun[1], Jiming Chen[1]

[1]State Key Laboratory of Industrial Control Technology, Zhejiang University, No. 38, Zheda Road, Hangzhou, 310027, Zhejiang, P.R.China.
[2]Center for Systems and Control, College of Engineering, Peking University, No. 5, Yiheyuan Road, Beijing, 100871, Beijing, P.R.China.
[3]College of Media and International Culture, Zhejiang University, No. 866, Yuhangtang Road, Hangzhou, 310058, Zhejiang, P.R.China.

Contributing authors: mingxie@zju.edu.cn; s18he@zju.edu.cn; liaming@pku.edu.cn; zkz@zju.edu.cn; 22332052@zju.edu.cn; sunyx@zju.edu.cn; cjm@zju.edu.cn;

1 Supplementary Note 1: DATA DESCRIPTION

We conduct our experiments using both synthetic networks and the networks generated from the real-world data. Firstly, we list the description of four higher-order networks, namely, Email-Enron, NDC-classes, Contact primary school and Contact-high school used in our experiments.

- **Email-Enron** [1]: A dataset characterizes a higher-order network of email communications within the Enron organization. In this context, higher-order interactions are represented by emails sent to multiple recipients simultaneously. Nodes in the network correspond to email addresses of Enron employees, while each interaction consists of the sender and all recipients of a given email. The dataset focuses on a core group of employees, ensuring a consistent and representative portrayal of the organization's internal communication structure.
- **NDC-Classes** [1]: A higher-order system in the realm of pharmacy, which is derived from the National Drug Code (NDC) Directory, maintained by the U.S. Food and Drug Administration under the Drug Listing Act. Each higher-order interaction represents a drug, and the nodes correspond to the classification labels assigned to it. This dataset captures the evolving structure of drug classifications over time, providing valuable insights into pharmaceutical trends and regulations.
- **Primary School** [1, 2]: The dataset captures higher-order interactions among individuals within a primary school setting. Each interaction is defined as the maximal group of individuals detected interacting within a specific time interval, recorded via wearable sensors. These sensors operate with a 20-second interval, aggregating all interactions occurring within the preceding period. Nodes in the network represent individuals, while the interactions delineate the social groups formed during the observation. This dataset provides a nuanced portrayal of dynamic social interactions within a controlled environment, offering a detailed representation for the patterns and structure of social group behavior.
- **High School** [1, 3]: Analogous to the Primary School dataset, the High School dataset is constructed in a comparable manner, capturing a higher-order social network of interactions among individuals within a high school setting. Wearable sensors record interactions at a 20-second resolution, similar to the Primary School dataset. In this network, nodes represent people within the high school, while

higher-order interactions characterize social groups over time. Note that in the processing of this dataset, duplicate social groups were excluded to ensure that each interaction uniquely represents a distinct group formation.

We begin our experiments by constructing the integrated higher-order network, where duplicate interactions are removed from the dataset, ensuring that each higher-order interaction is unique. Subsequently, we compute the number of the interactions, denoted as $|E|$, for the integrated network. In analyzing the linkages among various dynamic processes, we address the challenge posed by the parameter settings of infection rates, particularly when $M > 2$. To manage this complexity, we restrict the maximum size of hyperedges to 3 in our experiments, as this provides a representative case [4, 5]. Furthermore, we quantify the number of 3-sized hyperedges, $|E_3|$, in the hypergraph derived from the integrated higher-order network. The results, including the values of $|E_3|$, are summarized in Supplementary Table 1.

To quantify the contributions of the two driving factors, namely, structural and dynamical aspects, we construct simplicial complexes with diverse topologies and conduct the validation. The construction of these networks follows a configuration model, a widely adopted network generation model in the study of simplicial contagions. Such approach ensures a controlled variation of topological features while preserving key network properties, including $\langle k \rangle$ and $\langle k_\Delta \rangle$. Detailed topological characteristics of the generated simplicial complexes are presented in Supplementary Table 2.

In addition, we construct a series of hypergraphs to validate our findings in opinion dynamics. These hypergraphs are generated using the HyperCL algorithm [6], with the hyperdegree distribution $\mathcal{P}(k_h) \sim k_h^{-\gamma}$ controlled by the exponent γ [6–8]. The constructed hypergraphs $\mathcal{H}(V, E)$ consist of $V, E = \{100, 200, 500, 800, 1000\}$ hyperedges, with hyperedge sizes randomly chosen from 2 to 10. The resulting hyperdegree distributions are illustrated in Supplementary Fig. 1. To illustrate the scale-free property commonly observed in empirical systems, the exponent is set to $\gamma = 2$ in the HyperCL algorithm.

Table 1 Details of the higher-order networks in real-world datasets. N represents the network size, $|E|$ denotes the number of higher-order interactions in integrated higher-order networks, $|E_3|$ indicates the number of the 3-sized hyperedges restricted from the integrated higher-order networks.

Higher-order Networks	Nodes	Higher-order Interactions	N	$\lvert E\rvert$	$\lvert E_3\rvert$
Email-Enron	Email address	Email recipients	143	1542	6577
NDC-classes	Class labels	Drugs	1161	1224	37810
Primary school	people in primary school	Individual contacts	242	12799	5138
High school	people in high school	Individual contacts	327	7937	2369

Table 2 Topological properties of synthetic higher-order networks. $\langle k \rangle$ and $\langle k^* \rangle$ represents the average node degrees with/without excluding binary links contained in simplices. $\langle k_\Delta \rangle$ represents the average number of 2-simplex in the simplicial complex. $\langle C_1 \rangle$ and $\langle C_2 \rangle$ denotes the average clustering coefficient in 1- and 2-hops.

$\langle k^* \rangle$	$\langle k \rangle$	$\langle k_\Delta \rangle$	$\langle C_1 \rangle$	$\langle C_2 \rangle$
8	10	3	0.2325	0.2020
10	12	4	0.1847	0.1762
15	10	5	0.2009	0.2020
20	4	8	0.0552	0.0516

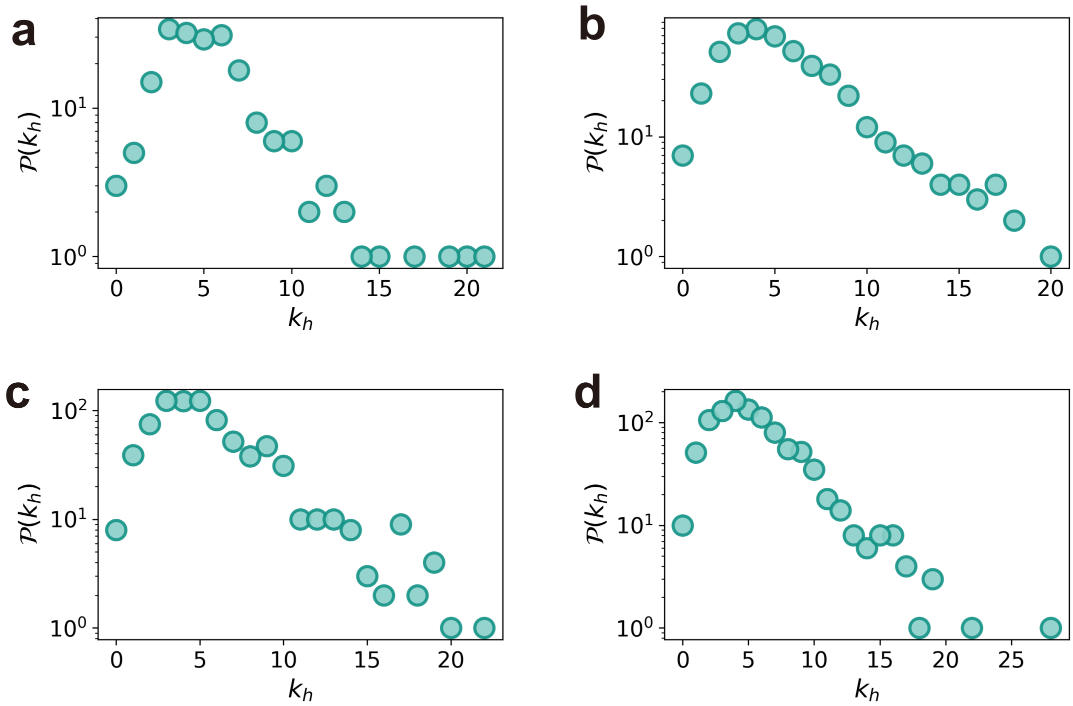

Fig. 1 The hyperdegree distribution $\mathcal{P}(k_h) \sim k_h^{-\gamma}$ of the synthetic hypergraphs of **a** $\mathcal{H}(V = 200, E = 200)$, **b** $\mathcal{H}(V = 500, E = 500)$, **c** $\mathcal{H}(V = 800, E = 800)$ and **d** $\mathcal{H}(V = 1000, E = 1000)$, respectively. Note that the exponent is set to be $\gamma = 2$.

2 Supplementary Note 2: MEAN-FIELD SOLUTION FOR DYNAMICS ON HIGHER-ORDER NETWORKS, MULTI-EDGE GRAPHS AND SIMPLE GRAPHS

We utilize the mean-field approximation [9–11] to investigate the dynamics on a higher-order network, Υ, alongside its dyadic representations, including the multi-edge dyadic graph $\tilde{\Upsilon}$ and the simple dyadic graph $\hat{\Upsilon}$.

As a representative example, we consider the SI contagion process on a simplicial complex Υ. In this framework, we derive expressions to estimate the infection rate at which an S-state node i transitions to the I-state within a 2-simplex (triangle) and along a binary link, denoted as $\Gamma_i^{\text{tri}}(t)$ and $\Gamma_i^{\text{bi}}(t)$, respectively. These rates are given by:

$$\Gamma_i^{\text{tri}}(t) = \sum_{i \in \Delta} [\beta I_{j_1}(t) + \beta I_{j_2}(t) + \beta_\Delta I_{j_1}(t) \cdot I_{j_2}(t)], \tag{1}$$

$$\Gamma_i^{\text{bi}}(t) = \sum_{i \in (i, j_3)} \beta I_{j_3}(t). \tag{2}$$

Here, nodes j_1 and j_2 represent two neighboring infected nodes (i.e., node with I-state) for i within a 2-simplex, while node j_3 is one neighboring I-state node in a binary link.

Consider a node i that, on average, is adjacent to$\langle k_\Delta \rangle$ 2-simplices and $\langle k \rangle$ single links, where $\langle k \rangle$ excludes the contributions from links contained within 2-simplices. We then derive the total infection rate as follows:

$$\Gamma_i(t) \tag{3}$$
$$= \langle k_\Delta \rangle \cdot \Gamma_i^{\text{tri}}(t) + \langle k \rangle \cdot \Gamma_i^{\text{bi}}(t) \tag{4}$$
$$= \langle k_\Delta \rangle \cdot ([\beta I_{j_1}(t) + \beta I_{j_2}(t) + \beta_\Delta I_{j_1}(t) \cdot I_{j_2}(t)]) + \langle k \rangle \cdot \beta I_{j_3}(t) \tag{5}$$

Then, the rate of change of the infected nodes can be described as follows:

$$\frac{dI_\Upsilon(t)}{dt} = S(t) \cdot \Gamma(t) \tag{6}$$
$$= (1 - I(t)) \cdot \Gamma(t) \tag{7}$$
$$= (1 - I(t)) \cdot [\langle k_\Delta \rangle \cdot \Gamma_i^{\text{tri}}(t) + \langle k \rangle \cdot \Gamma_i^{\text{bi}}(t)] \tag{8}$$
$$= (1 - I(t)) \cdot [\langle k_\Delta \rangle \cdot (2\beta I(t) + \beta_\Delta I(t) \cdot I(t)) + \langle k \rangle \cdot \beta I(t)]. \tag{9}$$

In the context of $\Gamma(t)$, we assume that $I_j(t) = I(t)$ within the framework of mean-field theory.

Building on this, we extend the master approximation equation to a multi-edge dyadic representation $\tilde{\Upsilon}$, which can be expressed as:

$$\frac{dI_{\tilde{\Upsilon}}(t)}{dt} = (1 - I(t)) \cdot (\langle k \rangle + 2 \langle k_\Delta \rangle) \cdot \beta I(t) \tag{10}$$

In the Methods section, we provide the mean-field solution for dynamics in dyadic graphs under SI contagion. Here, we apply this solution to the projected simple graph $\hat{\Upsilon}$ of Υ:

$$\frac{dI_{\hat{\Upsilon}}(t)}{dt} = (1 - I(t)) \cdot (\langle k \rangle + 2 \langle k_\Delta \rangle - \langle \Theta \rangle) \cdot \beta I(t). \tag{11}$$

3 Supplementary Note 3: INTERPLAY OF DYNAMICS FROM THE PERSPECTIVE OF INFECTION RATES

Here, we show the interconnections between dynamical processes through the perspective of infection rates, as depicted in Supplementary Fig. 2. The figure depicts the connections between two distinct complex contagion processes—dynamics on hypergraphs and simplicial complexes—as well as the transition from complex contagion to simple contagion dynamics. We demonstrate that, for any arbitrary hypergraph, the infection rate of an S-state node in the corresponding mapped M-adic network, $\mathcal{H}_M$, (where M represents the maximum hyperedge size) can be systematically determined. Such a determination relies on both the number and the size of higher-order interactions adjacent to the susceptible node, as well as the states of its neighboring nodes.

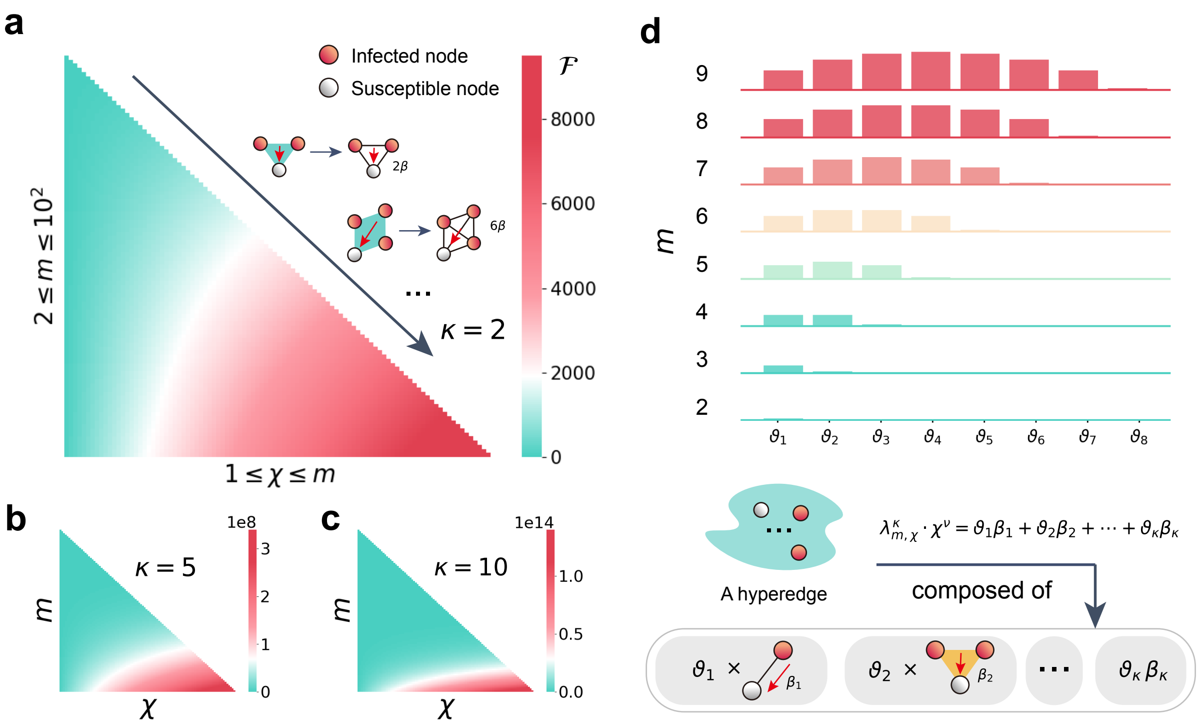

Fig. 2 Interrelationships among dynamics through infection rates. **a** Infection rates when contagion on hypergraphs is mapped onto dyadic graphs. This mapping is quantified by the term $\mathcal{F}(m, \chi, \kappa) \cdot \beta$, which denotes the infection rate experienced by an S-state node when dynamics within an m-sized hyperedge of the hypergraph are projected onto a dyadic graph. Here, $\mathcal{F}(m, \chi, \kappa) = \sum_{i=\tau}^{\tau^*} i \cdot \binom{\chi}{i}\binom{m-\chi-1}{\kappa-i}$, derived from C2. $\lambda^{\kappa}_{m,\chi} \cdot, \chi^{\nu}$ corresponds to the infection rate in an m-sized hypergraph with mapping ratio κ, containing χ I-state nodes. **a** illustrates the variation of $\mathcal{F}(m, \chi, \kappa = 2)$ as m and χ change, while **b** and **c** show corresponding results for $\kappa = 5$ and $\kappa = 10$, respectively. The analysis reveals that as κ increases, χ becomes the dominant factor influencing the infection rate relationship between hypergraph and dyadic graph dynamics, establishing a clear connection between complex and simple contagion processes. **d** The relationships of infection rates between dynamics on hypergraphs and simplicial complexes. The plot decomposes the contributions from simplices of varying orders within hyperedges of size $m = 2$ to 9, with $\kappa = 8$ and $\chi = m - 1$. The coefficient ϑ_δ quantifies the contribution of a δ-simplex (infection rate β_δ) to the S-state node.

4 Supplementary Note 4: PROOFS FOR THE GENERALIZED CONDITION FOR THE INTERPLAY OF DYNAMICS

In the main text, we present the infection probability $P^i(t)$ in a simplicial complex $\mathcal{C}$ as follows:

$$P^i_{\mathcal{C}}(t) = 1 - \prod_{\delta=2}^{\max|u|, u\in U} \prod_{j_1^\delta,\cdots,j_{\delta-1}^\delta \in I(t)} (1 - \beta_\delta A^\delta_{i,j_1^\delta,\cdots,j_{\delta-1}^\delta}). \tag{12}$$

Building on this, we aim to determine the condition that allows Eq. 12 to be transformed into the following expression, which captures $P^i(t)$ in a dyadic graph and can be written as:

$$P^i_{\mathcal{G}}(t) = 1 - \prod_{j\in I(t)} (1 - \beta A_{ij}). \tag{13}$$

If we let $\beta_\delta = 0(\delta > 1)$, we find that the expression for $P^i_{\mathcal{C}}(t)$ simplifies to:

$$P^i_{\mathcal{C}}(t) = 1 - \textstyle\prod_{j_1\in I(t)}(1 - \beta_1 A^1_{ij}), \tag{14}$$

Given that $\beta_1 = \beta$ and $A^1_{ij} = A_{ij}$, we derive:

$$\begin{aligned} P^i_{\mathcal{C}}(t) &= 1 - \textstyle\prod_{j\in I(t)}(1 - \beta A_{ij}) && (15)\\ &= P^i_{\mathcal{G}}(t). && (16) \end{aligned}$$

Thus, the condition for mapping the dynamics from a simplicial complex to a contagion on its projected multi-edge dyadic graph is $\beta_\delta = 0$ for $\delta > 1$.

Subsequently, we examine the relationship between the dynamics on a hypergraph $\mathcal{H}$ and its corresponding simplicial complex $\mathcal{C}$. In the main text, the infection probability $P^i_{\mathcal{H}}(t)$ is given by:

$$P^i_{\mathcal{H}}(t) = 1 - \prod_{m=2}^{\max|h|, h\in E} \prod_{h\in\mathcal{E}_i} (1 - \lambda\chi_h^\nu A_h^m). \tag{17}$$

Next, we decompose the formula in terms of various hyperedge sizes m, the number of infected nodes χ, and the mapping ratio κ, as detailed in the main text. This leads to the following expression:

$$P^i_{\mathcal{H}}(t) = 1 - \prod_{m=2}^{M} \prod_{h\in\mathcal{E}_i^m} (1 - \lambda\cdot\chi^\nu\cdot A_h^m) = 1 - \prod_{m=2}^{M} \prod_{\chi=1}^{m-1} (1 - \lambda^\kappa_{m,\chi}\cdot\chi^\nu\cdot A_h^m), \tag{18}$$

where M represents the maximum observed value of m.

Subsequently, the conditions are categorized into two distinct cases: (i) $\kappa < m$ and (ii) $\kappa \geq m$. We first examine the behavior of $\lambda^\kappa_{m,\chi}$ in each case separately, and then integrate the results into a unified framework.

In the case where $\kappa < m$, the m-sized hyperedge should first be mapped to κ-simplices. Let χ represent the number of infected nodes within the hyperedge. The number of infected nodes in the mapped κ-simplices can range from i to χ, where the minimum number of infected nodes, $\min(i)$, is given by $\kappa + 1 - (m - \chi)$, and the maximum number, $\max(i)$, is κ. For a susceptible (S-state) node a in the hyperedge, we must select an additional $\kappa - i$ nodes to complete the κ-simplex. Given that i infected nodes have already been chosen from the hyperedge, the remaining task is to select $\kappa - i$ susceptible nodes from the remaining $m - 1 - \chi$ susceptible nodes in the hyperedge. The infection rate in the resulting κ-simplex with i infected nodes is then determined as follows:

$$\sum_{j=1}^{i} \binom{i}{j} \beta_j. \tag{19}$$

Note that, the infection rate is represented using a linear sum formulation, where product terms are expanded and higher-order terms in the mathematical expansion are omitted to focus on the dominant contributions. Therefore, the condition for the infection rate node a can be expressed as:

$$\lambda_{m,\chi}^{\kappa} \cdot \chi^{\nu} = \sum_{i=\kappa+1-(m-\chi)}^{\kappa} \binom{\chi}{i} \binom{m-\chi-1}{\kappa-i} \sum_{j=1}^{i} \binom{i}{j} \beta_j. \tag{20}$$

Next, we consider the second case, where $\kappa \geq m$. In this scenario, the hyperedge does not require mapping, and the infection rate for a susceptible (S-state) node a within the hyperedge can be expressed as follows:

$$\lambda_{m,\chi}^{\kappa} \cdot \chi^{\nu} = \sum_{i=1}^{\chi} \binom{\chi}{i} \beta_i. \tag{21}$$

We combine Eqs. 20 and 21, leading to the consolidated expression:

$$\lambda_{m,\chi}^{\kappa} = \frac{1}{\chi^{\nu}} \left[\mathbb{I}_{\kappa<m} \cdot \sum_{i=\max(1,\kappa+1-(m-\chi))}^{\min(\kappa,\chi)} \binom{\chi}{i} \binom{m-\chi-1}{\kappa-i} \sum_{j=1}^{i} \binom{i}{j} \beta_j + \mathbb{I}_{\kappa>=m} \cdot \sum_{i=1}^{\chi} \binom{\chi}{i} \beta_i] \right] \tag{22}$$

We then define $\tau = \max(1, \kappa+1-(m-\chi))$ and $\tau^* = \min(\kappa, \chi)$. The equation of Eq. 22 can then be consolidated into the following form:

$$\lambda_{m,\chi}^{\kappa} = \frac{1}{\chi^{\nu}} \sum_{i=\tau}^{\tau^*} \binom{\chi}{i} \binom{m-\chi-1}{\kappa-i} \sum_{j=1}^{i} \binom{i}{j} \beta_j. \tag{23}$$

Building upon this, we derive the generalized condition for the coupling of dynamics between hypergraphs and simplicial complexes. Furthermore, under the condition $\beta_\delta = 0 (\delta > 1)$, the transformable condition for the dynamics between hypergraphs and its dyadic multi-edge projection is given by:

$$\lambda_{m,\chi}^{\kappa} = \frac{1}{\chi^{\nu}} \sum_{i=\tau}^{\tau^*} i \cdot \binom{\chi}{i} \binom{m-\chi-1}{\kappa-i} \beta. \tag{24}$$

Thus, the generalized conditions for the coupling across various dynamics are established.

5 Supplementary Note 5: DYNAMIC DIFFERENCES IN INFORMATION ENTROPY ACROSS NETWORK ORDERS

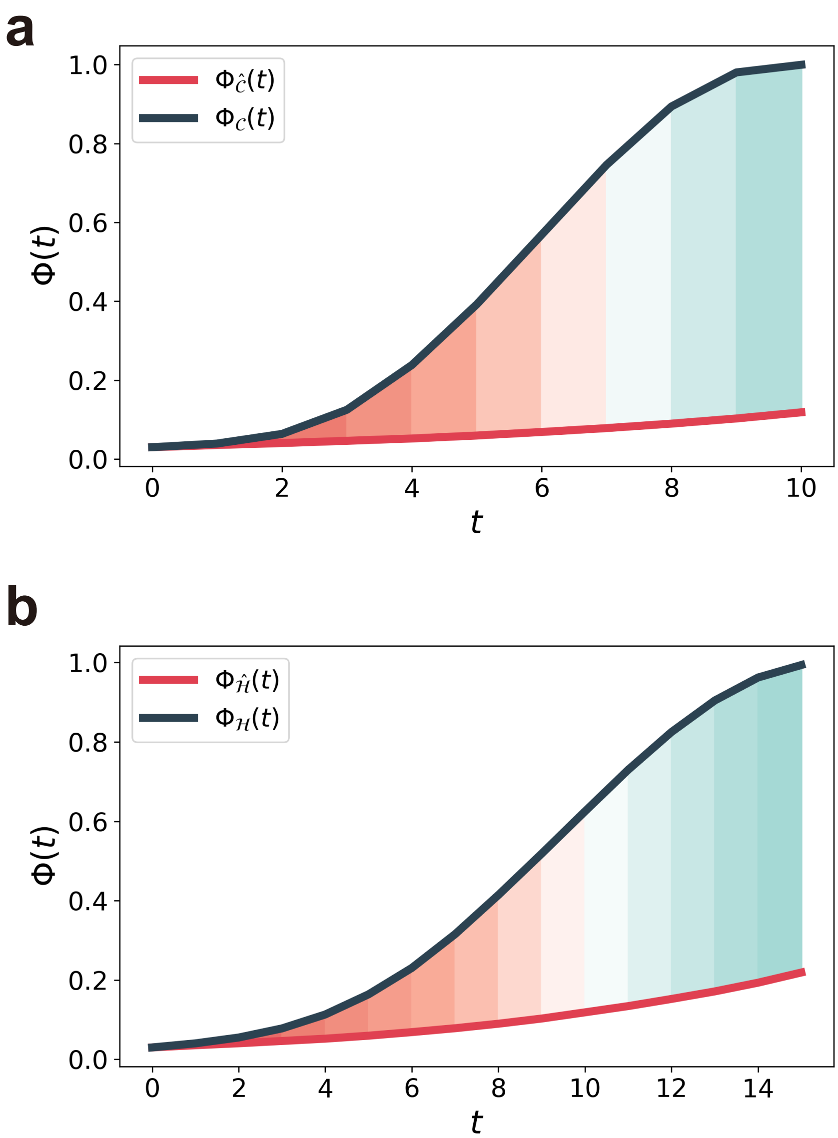

Fig. 3 Encoding empirical systems with higher-order interactions as dyadic graphs leads to information loss in dynamics. The variance of information entropy Φ during contagion dynamics is analyzed in (**a**) a hypergraph $\mathcal{H}$, (**b**) a simplicial complex $\mathcal{C}$, and their projected simple graphs. Notably, the hypergraph and simplicial complex are derived from real datasets of High-school and Primary-school. As time steps increase, a growing disparity in Φ emerges between dynamics on hypergraphs or simplicial complexes (red) and their projected graphs (black). Each time step t represents the average over 500 simulations of the SI contagion process.

6 Supplementary Note 6: PROOFS FOR THE UNDERLYING TWO CAUSES FOR THE TRANSFORMATION OF DYNAMICS ACROSS NETWORK ORDERS

Furthermore, we prove that both dynamical and structural impacts independently drive the transformation of dynamics across network orders. First, we quantify the total information differences between higher-order networks and their projected simple graphs, as well as the impacts arising from both structural and dynamical factors. To illustrate this, we consider the problem of exploring the transition of dynamics between a simplicial complex and its projected simple graph. Specifically, the task is to prove:

$$\Phi_{\mathcal{C}}^{\mathcal{Q}}(t) - \Phi_{\hat{\mathcal{C}}}(t) = \Phi_{\mathcal{C}}^{\mathcal{Q}}(t) - \Phi_{\mathcal{C}}^{\mathcal{Q}^*}(t) + \Phi_{\tilde{\mathcal{C}}}(t) - \Phi_{\hat{\mathcal{C}}}(t) \tag{25}$$

Based on the relation $\Phi(t) = \Phi(I(t))$, as described in the Methods section in the main text, and under the condition $I_{\tilde{\mathcal{C}}}(t) = I_{\hat{\mathcal{C}}}(t) = I_{\mathcal{C}}^{Q}(t) = I_{\mathcal{C}}^{\mathcal{Q}^*}(t) = I_0$, the problem is reduced to proving the following expression:

$$\frac{dI_{\mathcal{C}}^{\mathcal{Q}}(t)}{dt} - \frac{dI_{\mathcal{C}}^{\mathcal{Q}^*}(t)}{dt} + \frac{dI_{\tilde{\mathcal{C}}}(t)}{dt} - \frac{dI_{\hat{\mathcal{C}}}(t)}{dt} = \frac{dI_{\mathcal{C}}^{\mathcal{Q}}(t)}{dt} - \frac{dI_{\hat{\mathcal{C}}}(t)}{dt} \tag{26}$$

Using the mean-field theory outlined in Supplementary Note 2, we subsequently evaluate $\Delta I_1(t)$ and $\Delta I_2(t)$, respectively:

$$\begin{aligned}
&\Delta I_1(t) && (27)\\
&= \frac{dI_{\mathcal{C}}^{\mathcal{Q}}(t)}{dt} - \frac{dI_{\mathcal{C}}^{\mathcal{Q}^*}(t)}{dt} && (28)\\
&= (1 - I(t)) \cdot [\langle k_\Delta \rangle \cdot (2\beta I(t) + \beta_\Delta I^2(t)) + \langle k \rangle \cdot \beta I(t)] - (1 - I(t)) \cdot [\langle k_\Delta \rangle \cdot 2\beta I(t) + \langle k \rangle \cdot \beta I(t)] && (29)\\
&= (1 - I(t)) \cdot [\langle k_\Delta \rangle \cdot \beta_\Delta I^2(t)]. && (30)
\end{aligned}$$

$$\begin{aligned}
&\Delta I_2(t) && (31)\\
&= \frac{dI_{\tilde{\mathcal{C}}}(t)}{dt} - \frac{dI_{\hat{\mathcal{C}}}(t)}{dt} && (32)\\
&= (1 - I(t)) \cdot (\langle k \rangle + 2\langle k_\Delta \rangle) \cdot \beta I(t) - (1 - I(t)) \cdot (\langle k \rangle + 2\langle k_\Delta \rangle - \langle \Theta \rangle) \cdot \beta I(t) && (33)\\
&= \langle \Theta \rangle (1 - I(t)) \cdot \beta. && (34)
\end{aligned}$$

Here, in the context of the higher-order network of simplicial complexes, $\mathcal{Q}^*$ represents the value of $\mathcal{Q}$ under the condition $\beta_\delta = 0$ for $\delta > 1$. Subsequently, we evaluate $\Delta I^*(t)$ as follows:

$$\begin{aligned}
&\Delta I^*(t) && (35)\\
&= \frac{dI_{\mathcal{C}}^{\mathcal{Q}}(t)}{dt} - \frac{dI_{\hat{\mathcal{C}}}(t)}{dt} && (36)\\
&= (1 - I(t)) \cdot [\langle k_\Delta \rangle \cdot (2\beta I(t) + \beta_\Delta I^2(t)) + \langle k \rangle \cdot \beta I(t)] - (1 - I(t)) \cdot (\langle k \rangle + 2\langle k_\Delta \rangle - \langle \Theta \rangle) \cdot \beta I(t) && (37)\\
&= (1 - I(t)) \cdot [\langle k_\Delta \rangle \cdot \beta_\Delta I^2(t) + \langle \Theta \rangle (1 - I(t)) \cdot \beta] && (38)\\
&= \Delta I_1(t) + \Delta I_2(t). && (39)
\end{aligned}$$

We find that the value of ΔI^* is equivalent to the sum of ΔI_1 and ΔI_2. This indicates that the two independent contributing factors—namely, the dynamical and structural aspects—jointly account for the information differences between the dynamics on higher-order and pairwise networks.

7 Supplementary Note 7: VALIDATION OF THE QUANTITATIVE MODEL

In the main text, we present the results under the SI contagion dynamics. Here, we provide a detailed comparison between the simulation and theoretical results of our proposed quantitative model across different network topologies, under the SI, SIS, and SIR contagion dynamics, see Supplementary Fig. 4-6. The topological structures of these networks are provided in Table S2. We note that each experimental data point in the figures is derived over 1000 realizations. We observe that the proposed model exhibits robust quantitative performance across a variety of network structures and contagion parameters. This holds true for the SI model as well as for the SIS and SIR models, demonstrating consistent agreement between theoretical predictions and simulation outcomes.

However, we observed that under specific parameter settings, the model's quantitative predictions did not align satisfactorily with the actual simulation results. To rule out the possibility that network size was a contributing factor, we generated a $\mathcal{H}_3$ hypergraph (i.e., $M = 3$, also representing a simplicial complex with the observation of 2-simplices) using a configuration model with $N = 2000$ nodes, and evaluated the model's fit to the infected ratio, $\overline{I}(t)$, as shown in Supplementary Fig. 7. In the early stages of the epidemic, the model demonstrated significant accuracy in quantifying both driving factors and closely captured the phase transition of the contagion. Furthermore, the model provided a reliable estimation of the overall epidemic outbreak scale. Nonetheless, in the later stages, the model exhibited noticeable discrepancies in capturing the evolution of $\overline{I}(t)$.

To address this issue, we developed two sets of calibration mechanisms to improve the fitting accuracy of the infection proportion $\overline{I}(t)$ and thereby reduce the quantification error of $\Phi(t)$. First, for both simple graphs and multi-graphs, we introduced the concept of effective degree. Prior to defining the effective degree, we first define the notion of redundancy. Redundancy refers to the connections between the neighbors of a node during the propagation process that do not contribute to the spread. Mathematically, redundancy is expressed as:

$$\langle k_{\mathrm{rd}}\rangle = \langle C_1\rangle(\langle k^*\rangle - 1) + \langle C_2\rangle(\langle k^*\rangle - 1), \tag{40}$$

where $\langle C_1\rangle$ denotes the average clustering coefficient of the nodes, $\langle C_2\rangle$ represents the average clustering coefficient of second-order neighbors, and $\langle k^*\rangle$ is the average degree of the nodes. The average effective degree is then defined as the difference between the average degree and the average redundancy:

$$\langle k_{\mathrm{eff}}\rangle = \langle k^*\rangle - \langle k_{\mathrm{rd}}\rangle \tag{41}$$

In the field equations for both simple graphs and multi-graphs, we replace the average node degree $\langle k^*_{\hat{\mathcal{H}}_3}\rangle = \langle k\rangle + 2\langle k_\Delta\rangle - \langle\Theta\rangle$ and $\langle k^*_{\tilde{\mathcal{H}}_3}\rangle = \langle k\rangle + 2\langle k_\Delta\rangle$ with the corresponding effective degrees $\langle k_{\mathrm{eff}_{\hat{\mathcal{H}}_3}}\rangle$ and $\langle k_{\mathrm{eff}_{\tilde{\mathcal{H}}_3}}\rangle$, respectively. By incorporating local topological information, the field equations provide a more refined consideration of the effective local spreading range of nodes. The calibration results are depicted in Supplementary Fig. 8.

Additionally, for higher-order networks, we extend the mean-field equations by incorporating degree heterogeneity. Specifically, the original equations are transformed into difference equations that account for the degree distribution of nodes. These are then compared with simulation results, as shown in Supplementary Fig. 8. We find that our proposed quantitative model for propagation dynamics across different network types achieves better performance for quantification, particularly in larger-scale networks.

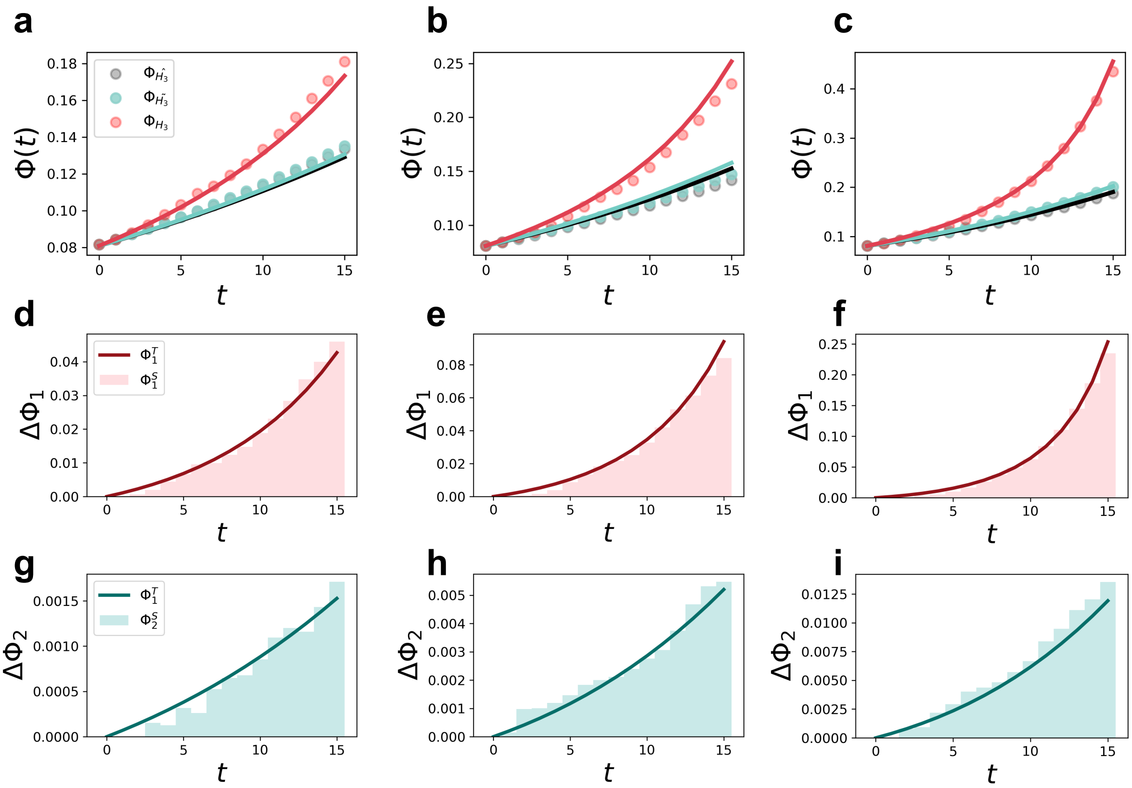

Fig. 4 Comparison of the theoretical (lines) and simulated values (scatters or bars) of **a-c** $\Phi(t)$, **d-f** $\Delta\Phi_1(t)$ and **g-i** $\Delta\Phi_2(t)$ in networks with different topologies, including $\mathcal{H}_3(\langle k^* \rangle = 8, \langle k_\Delta \rangle = 3)$, $\mathcal{H}_3(\langle k^* \rangle = 10, \langle k_\Delta \rangle = 4)$, and $\mathcal{H}_3(\langle k^* \rangle = 15, \langle k_\Delta \rangle = 5)$. Here, $N = 100$, and the contagion dynamics are conducted using the SI model. The infected rate of β and β_Δ are set to be 0.005 and 0.5, respectively. All results are averaged over 500 numerical simulations.

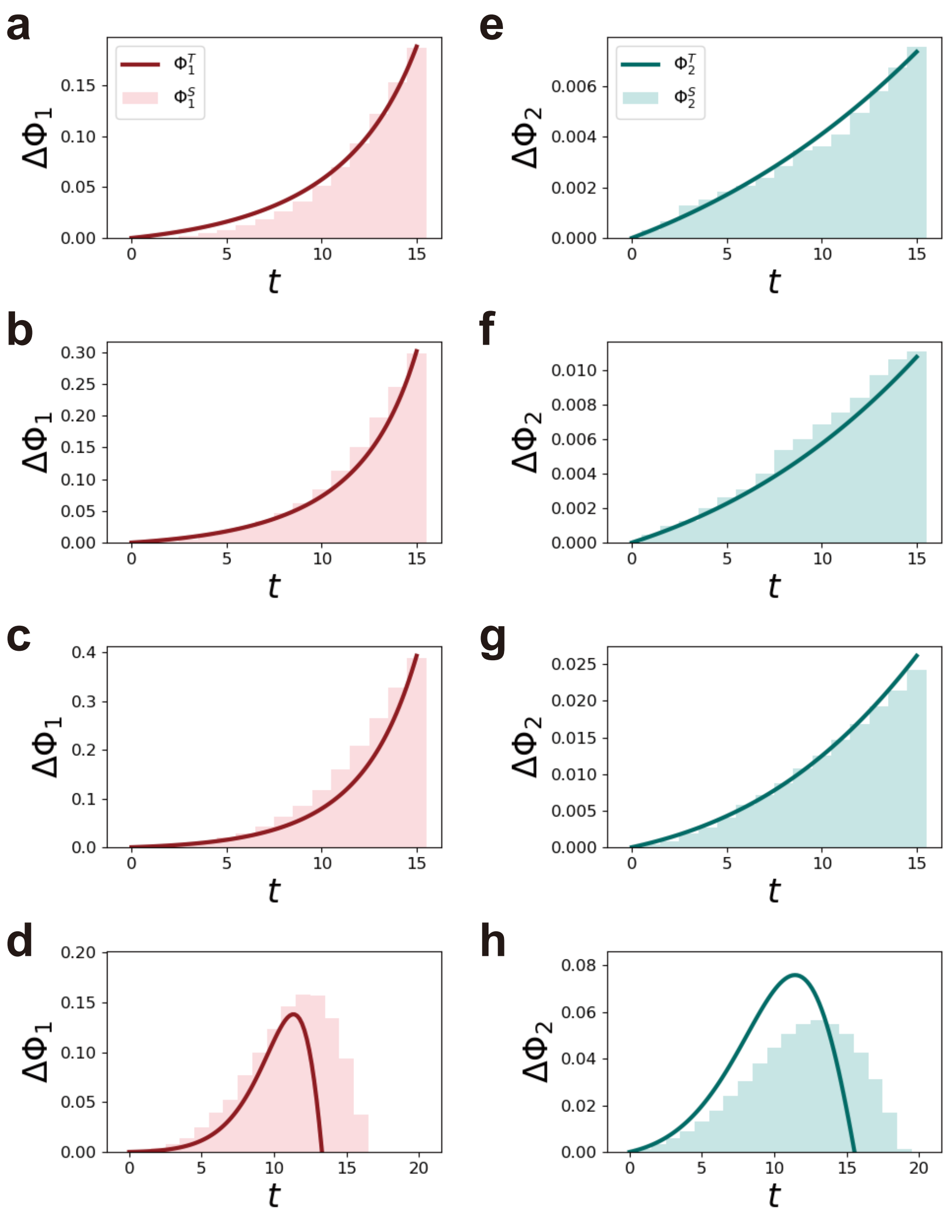

Fig. 5 Comparison of the theoretical (lines) and simulated values (bars) of **a**-**d** $\Delta\Phi_1(t)$ (red) and **e**-**h** $\Delta\Phi_2(t)$ (green) in the network of $\mathcal{H}_3(\langle k^* \rangle = 15, \langle k_\Delta \rangle = 5)$, where $N = 100$. The contagion dynamics are conducted on SIS model. The infected rate of β and β_Δ and recovery rate of μ are set to be {0.004, 0.005, 0.008, 0.02}, {0.6, 0.6, 0.4, 0.1} and {0.01, 0.01, 0.01, 0.005}, respectively. All results are averaged over 500 numerical simulations.

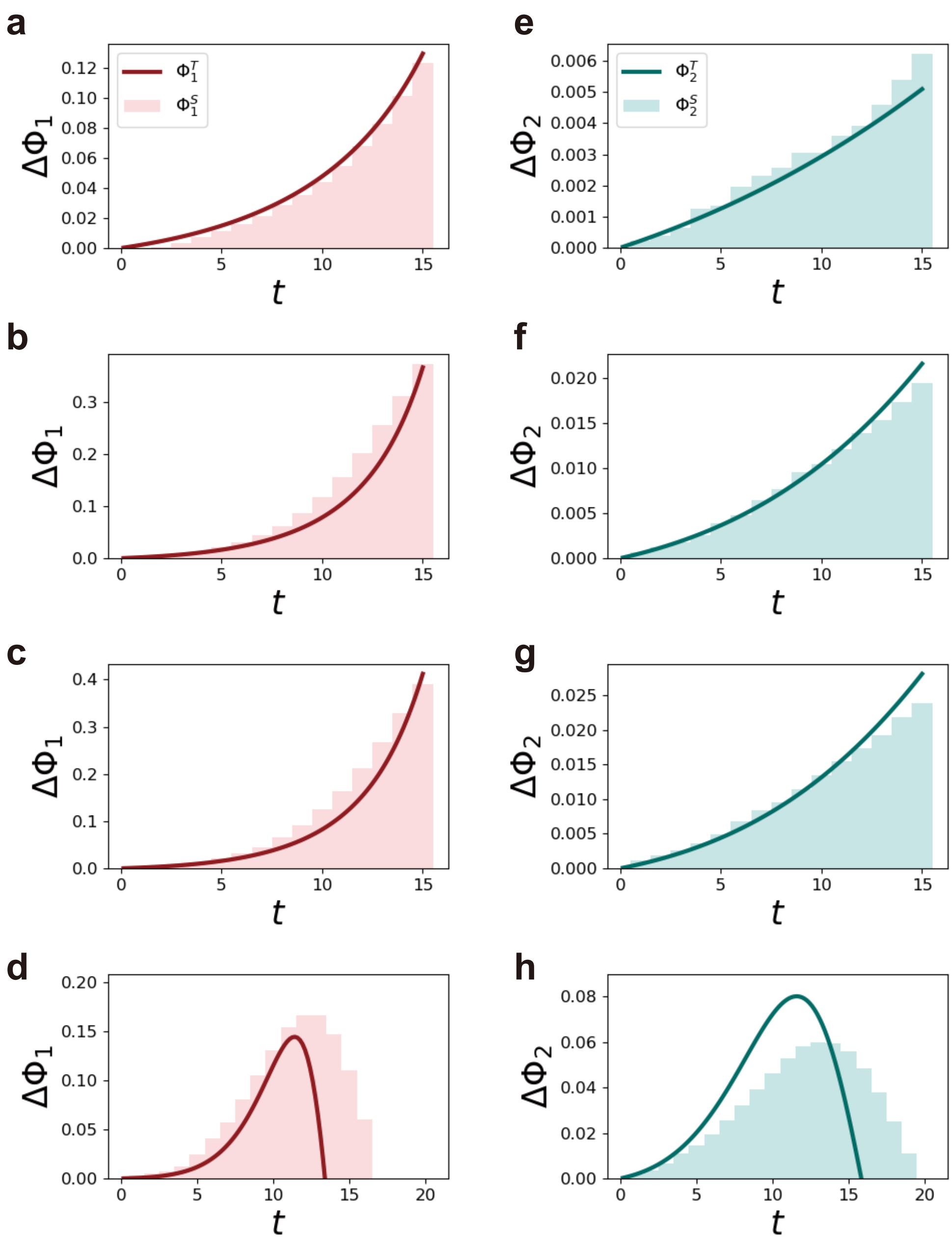

Fig. 6 Comparison of the theoretical (lines) and simulated values (bars) of **a-d** $\Delta\Phi_1(t)$ (red) and **e-h** $\Delta\Phi_2(t)$ (green) in the network of $\mathcal{H}_3(\langle k^* \rangle = 15, \langle k_\Delta \rangle = 5)$, where $N = 100$. The contagion dynamics are conducted on SIR model. The infected rate of β and β_Δ and recovery rate of μ are set to be {0.003, 0.007, 0.008, 0.02}, {0.6, 0.45, 0.4, 0.1} and {0.01, 0.01, 0.01, 0.005}, respectively. All results are averaged over 500 numerical simulations.

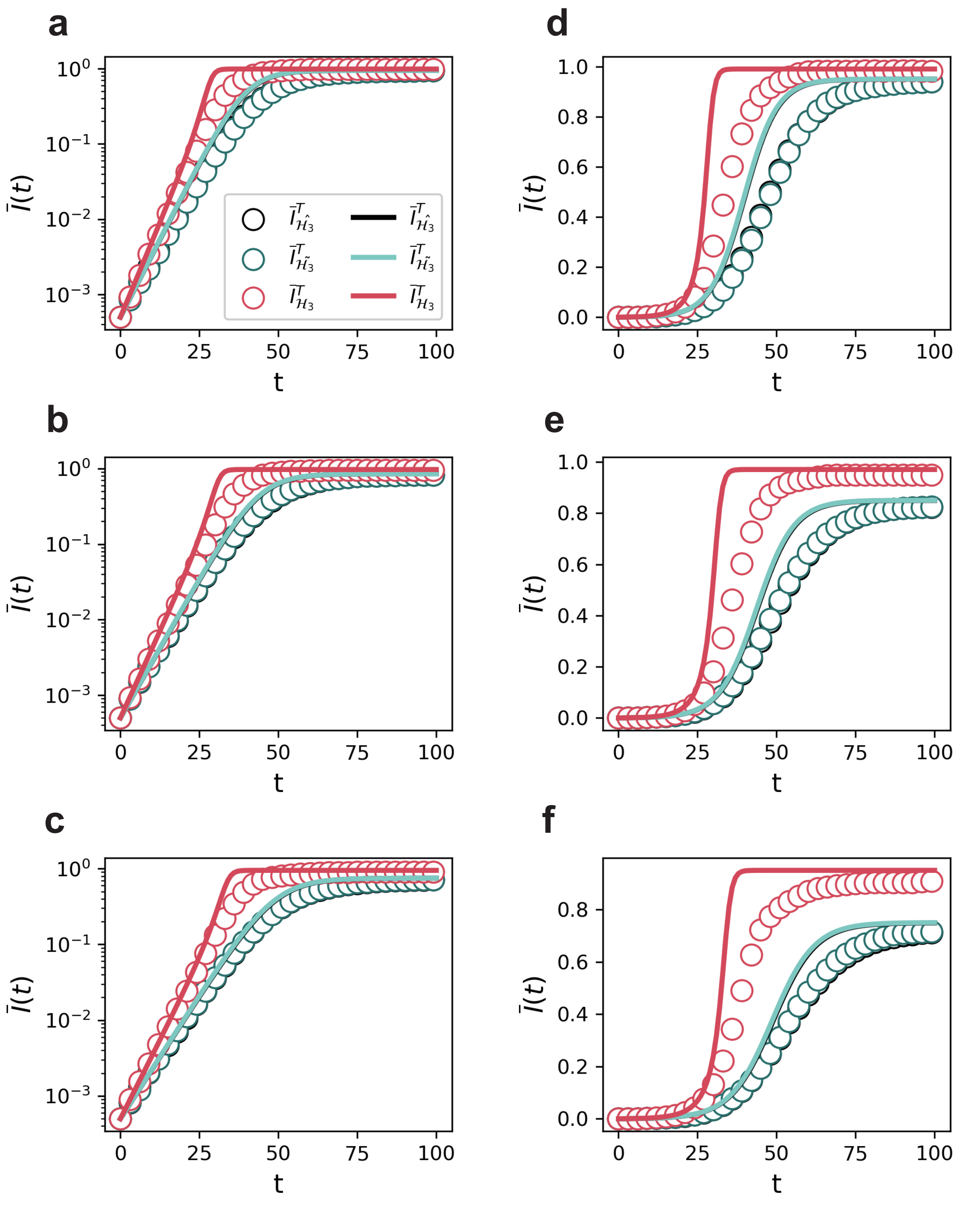

Fig. 7 Comparison of the theoretical values (lines) and simulated values (scatters) of **a-c** $\overline{I}(t)$, and **e-f** their corresponding numerical values in logarithmic coordinates. The experiment is conducted in the network of $\mathcal{H}_3(\langle k^* \rangle = 20, \langle k_\Delta \rangle = 8)$, where $N = 2000$. The contagion dynamics are conducted on SIS model. The infected rate of β and β_Δ are set to be 0.01 and 0.1, respectively. The recovery rate μ are set to be **a** 0.01, **b** 0.03 and **c** 0.05, respectively. All results are averaged over 1000 numerical simulations.

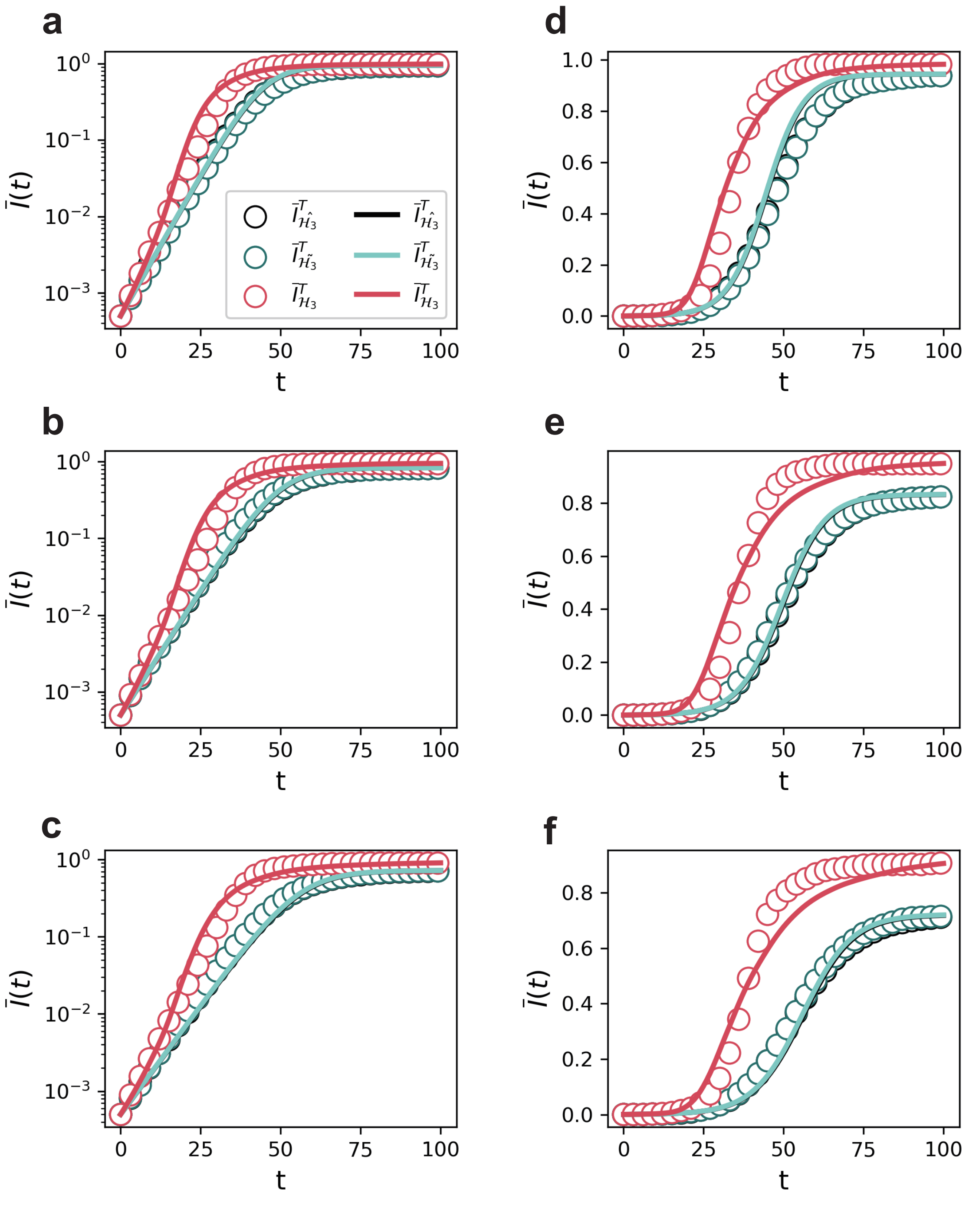

Fig. 8 Comparison of the calibrated theoretical (lines) and simulated values (scatters) of **a-c** $\bar{I}(t)$, and **e-f** their corresponding numerical values in logarithmic coordinates. The experiment is conducted in the network of $\mathcal{H}_3(\langle k^* \rangle = 20, \langle k_\Delta \rangle = 8)$, where $N = 2000$. The contagion dynamics are conducted on SIS model. The infected rate of β and β_Δ are set to be 0.01 and 0.1, respectively. The recovery rate μ are set to be **a** 0.01, **b** 0.03 and **c** 0.05, respectively. All results are averaged over 1000 numerical simulations.

8 Supplementary Note 8: DYNAMICAL AND STRUCTURAL EFFECTS ON OPINION DYNAMICS

As the main text gives the validation in synthetic hypergraphs of $\mathcal{H}(V = 100, E = 100)$ and the empirical hypergraphs of Email-Enron. We provide the results of the validation in various synthetic and empirical systems. Firstly, we conduct the validation in other series of synthetic networks generated by HyperCl algorithm, as shown in Supplementary Fig. 9.

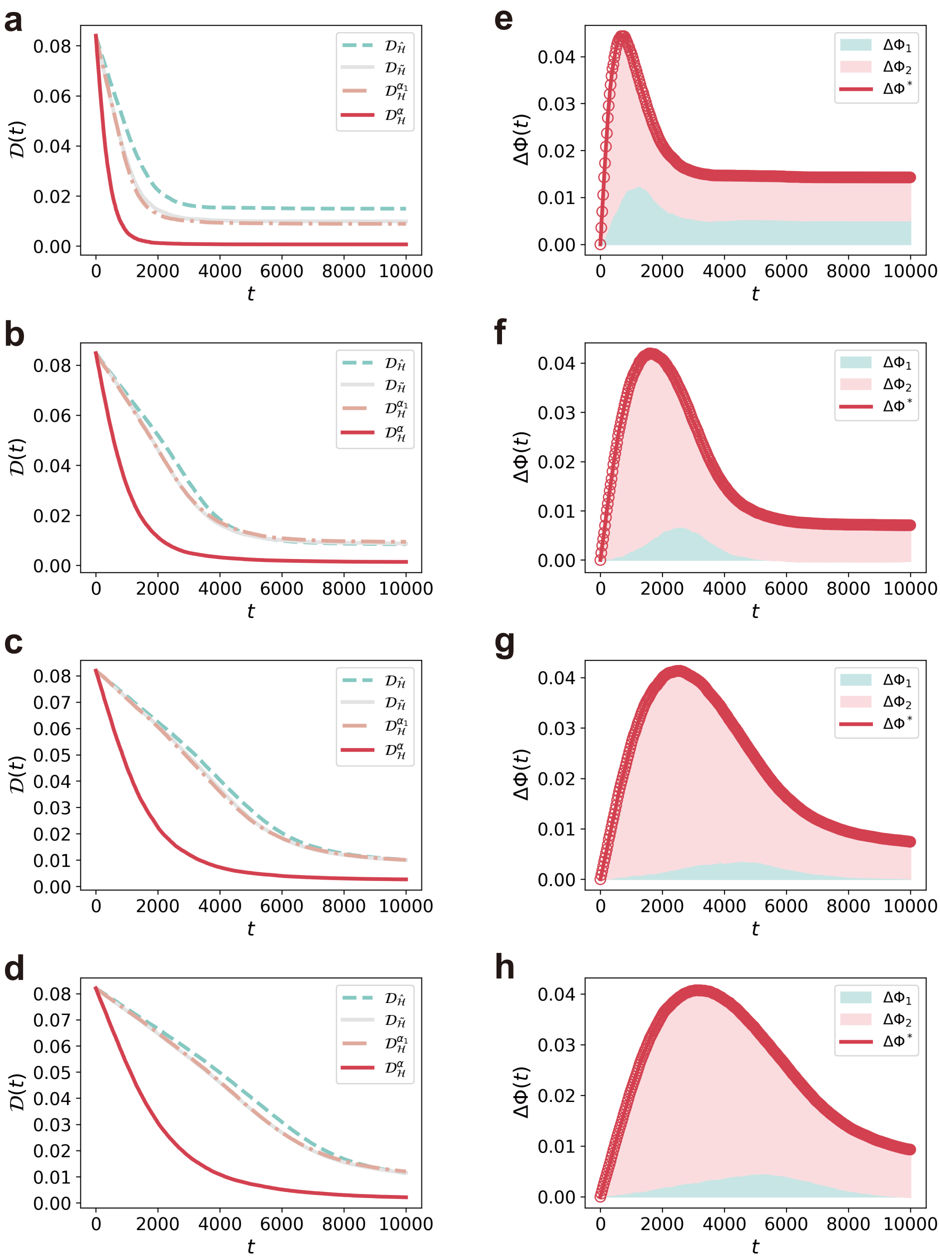

Fig. 9 Declining curves for $\mathcal{D}(t)$ observed in the DW dynamical process in synthetic hypergraphs of **a** $\mathcal{H}(V = 200, E = 200)$, **b** $\mathcal{H}(V = 500, E = 500)$, **c** $\mathcal{H}(V = 800, E = 800)$ and **d** $\mathcal{H}(V = 1000, E = 1000)$. Note that dynamics on higher-order networks with $\alpha = 1/2$ and $\alpha = 1$, as well as those on multi-edge graphs and simple graphs, denoted as $\mathcal{D}^{\alpha}_{\mathcal{H}}$, $\mathcal{D}^{\alpha_1}_{\mathcal{H}}$, $\mathcal{D}_{\tilde{\mathcal{H}}}$, and $\mathcal{D}_{\hat{\mathcal{H}}}$, respectively. **e-h** The relationships between the dynamical loss $\Delta\Phi_1 = \mathcal{D}^{\alpha}_{\mathcal{H}}(t) - \mathcal{D}^{\alpha_1}_{\mathcal{H}}(t)$, the structural loss $\Delta\Phi_2 = \mathcal{D}_{\tilde{\mathcal{H}}}(t) - \mathcal{D}_{\hat{\mathcal{H}}}(t)$, and the total information loss $\Delta\Phi^*$ are examined. All results are averaged over 10 numerical simulations with the time steps $t = 10^4$.

Subsequently, we conduct the validation of our insights on opinion dynamics using three other empirical systems, including High-school, Primary-school and NDC-class, as shown in Supplementary Fig. 10. We find that our insights that decompose the information loss in dynamical and structural aspects are also applicable to opinion dynamics in various real scenarios.

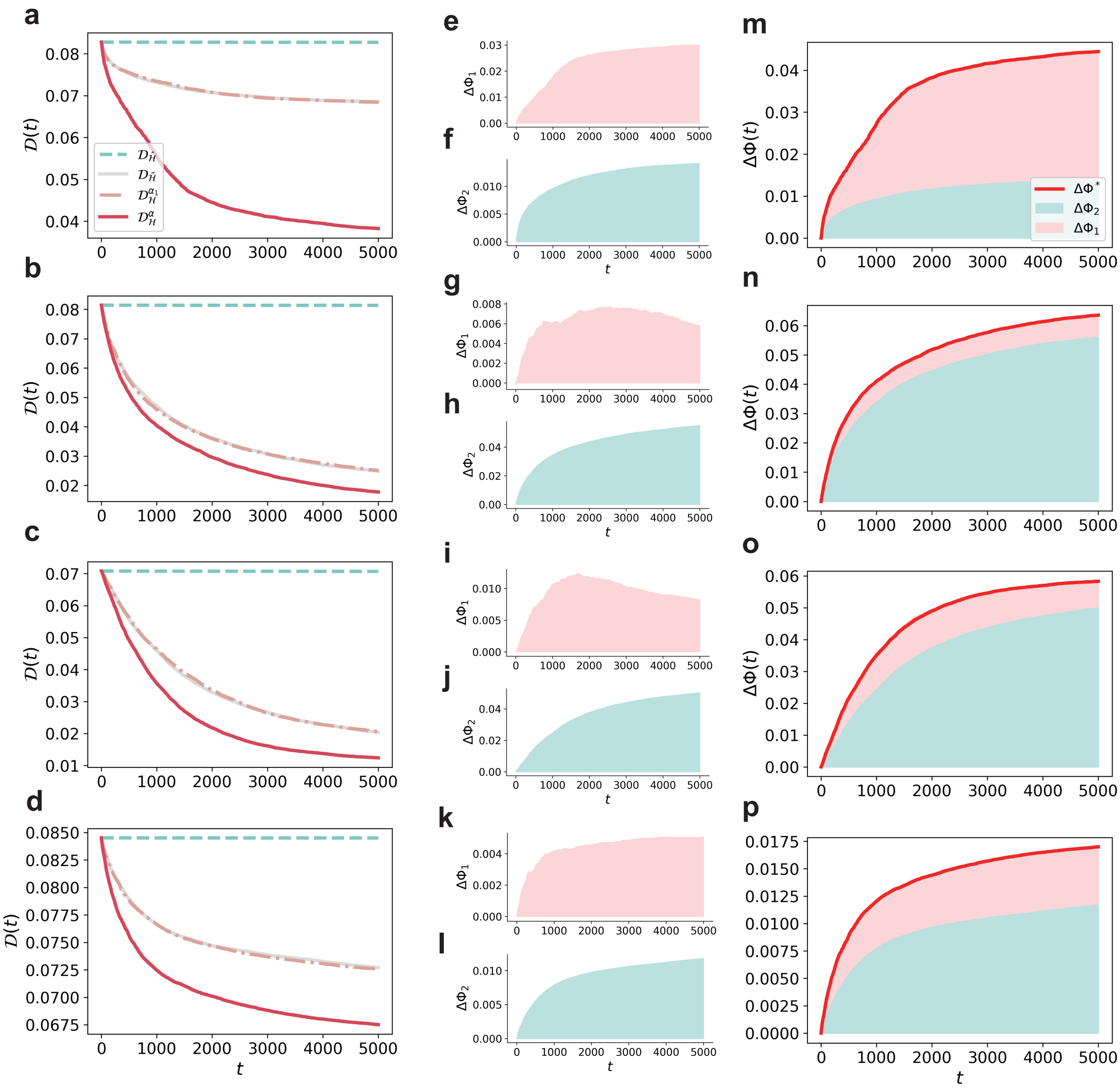

Fig. 10 Declining curves for $\mathcal{D}(t)$ observed in the DW dynamical process in empirical systems of **a** Email-Enron, **b** contact-high-school, **c** contact-primary-school and **d** NDC-class. Note that dynamics on higher-order networks with $\alpha = 1/2$ and $\alpha = 1$, as well as those on multi-edge graphs and simple graphs, denoted as $\mathcal{D}^{\alpha}_{\mathcal{H}}$, $\mathcal{D}^{\alpha 1}_{\mathcal{H}}$, $\mathcal{D}_{\tilde{\mathcal{H}}}$, and $\mathcal{D}_{\hat{\mathcal{H}}}$, respectively. **e-l** Both the values of dynamical and structural effects observed from each dataset are presented. **m-p** The relationships between the dynamical impact $\Delta\Phi_1$, the structural impact $\Delta\Phi_2$, and the total differences of system instability $\Delta\Phi^*$ are examined. All results are averaged over 10 numerical simulations with the time steps $t = 5 \times 10^3$.

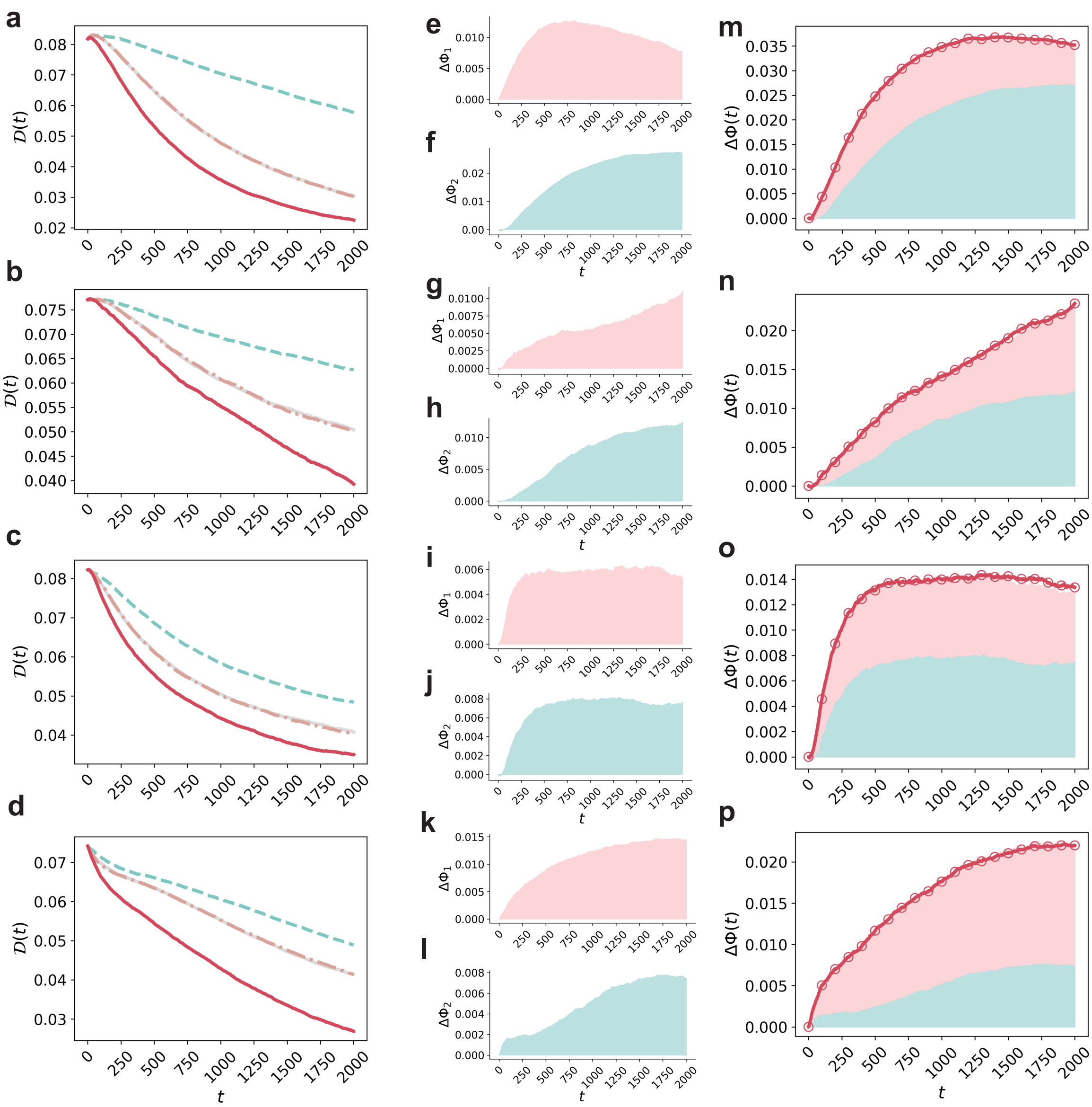

Fig. 11 Declining curves for $\mathcal{D}(t)$ observed in the Voter dynamical process in empirical systems of **a** Email-Enron, **b** contact-high-school, **c** contact-primary-school and **d** NDC-class. Note that dynamics on higher-order networks with $\alpha = 1/2$ and $\alpha = 1$, as well as those on multi-edge graphs and simple graphs, denoted as $\mathcal{D}^{\alpha}_{\mathcal{H}}$, $\mathcal{D}^{\alpha_1}_{\mathcal{H}}$, $\mathcal{D}_{\tilde{\mathcal{H}}}$, and $\mathcal{D}_{\hat{\mathcal{H}}}$, respectively. **e-l** Both the values of dynamical and structural impacts observed from the datasets. **m-p** The relationships between the dynamical impact $\Delta\Phi_1$, the structural impact $\Delta\Phi_2$, and the total differences of system instability $\Delta\Phi^*$ are examined. All results are averaged over 2×10^3 numerical simulations with the time steps $t = 5 \times 10^3$.

References

1. Benson, A. R., Abebe, R., Schaub, M. T., Jadbabaie, A. & Kleinberg, J. Simplicial closure and higher-order link prediction. *Proc. Natl Acad. Sci.* ISSN: 0027-8424 (2018).
2. Stehlé, J., Voirin, N., Barrat, A., Cattuto, C., Isella, L., Pinton, J.-F., Quaggiotto, M., den Broeck, W. V., Régis, C., Lina, B. & Vanhems, P. High-Resolution Measurements of Face-to-Face Contact Patterns in a Primary School. *PLoS ONE* **6** (ed Viboud, C.) e23176. https://doi.org/10.1371/journal.pone.0023176 (2011).
3. 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** (ed Viboud, C.) e0136497. https://doi.org/10.1371/journal.pone.0136497 (2015).
4. Iacopini, I., Petri, G., Barrat, A. & Latora, V. Simplicial models of social contagion. *Nat. Commun.* **10**, 2485 (2019).
5. Cencetti, G., Contreras, D. A., Mancastroppa, M. & Barrat, A. Distinguishing simple and complex contagion processes on networks. *Phys. Rev. Lett.* **130**, 247401 (2023).
6. Lee, G., Choe, M. & Shin, K. *How do hyperedges overlap in real-world hypergraphs?-patterns, measures, and generators* in *Proc. Web Conf. 2021* (2021), 3396–3407.
7. Kim, S., Lee, S. Y., Gao, Y., Antelmi, A., Polato, M. & Shin, K. *A survey on hypergraph neural networks: an in-depth and step-by-step guide* in *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining* (2024), 6534–6544.
8. Lee, G., Bu, F., Eliassi-Rad, T. & Shin, K. A survey on hypergraph mining: Patterns, tools, and generators. *arXiv preprint arXiv:2401.08878* (2024).
9. De Arruda, G. F., Petri, G. & Moreno, Y. Social contagion models on hypergraphs. *Phys. Rev. Res.* **2**, 023032 (2020).
10. Malizia, F., Gallo, L., Frasca, M., Kiss, I. Z., Latora, V. & Russo, G. A pair-based approximation for simplicial contagion. *arXiv preprint arXiv:2307.10151* (2023).
11. Lucas, M., Iacopini, I., Robiglio, T., Barrat, A. & Petri, G. Simplicially driven simple contagion. *Phys. Rev. Res.* **5**, 013201 (2023).