RESEARCH ARTICLE | OCTOBER 20 2020

The effect of heterogeneity on hypergraph contagion models

Nicholas W. Landry ; Juan G. Restrepo

30, 103117 (2020)
https://doi.org/10.1063/5.0020034

Chaos A L sc on o o n /c

The effect of heterogeneity on hypergraph contagion models

model has also been extended to scale-free uniform hypergraphs.[29] The fact that the network SIS model with more general higher-order interactions results in bistability has been proven rigorously in Ref. 26. However, so far, there is no general theory explaining how heterogeneity and correlations in the structure of higher-order interactions affect the onset of bistability.

In this paper, we present and analyze a degree-based mean-field description of the dynamics of the SIS model in networks with higher-order interactions. To describe higher-order interactions, we consider the SIS model on a p g p, formed by a set of nodes and a set of edges of multiple sizes (so that edges of size larger than two represent higher-order interactions). Our formulation allows us to consider heterogeneous structures in the organization of the edges of a given size and correlations between the structure of edges of different sizes. Using the illustrative case of networks with edges of sizes 2 and 3, we derive conditions for the appearance of bistability and hysteresis in terms of moments of the degree distribution of the pairwise interaction network. We find that the onset of bistability and hysteresis can be suppressed by heterogeneity in the pairwise interaction network and promoted by positive correlations between the number of pairwise and higher-order interactions a node has. We also consider the effect of healing by higher-order interactions (a "hipster effect").

The structure of the paper is as follows. In Sec. II, we present our hypergraph and contagion models. In Sec. III, we derive a mean-field description of the model and apply it to various illustrative cases. In Sec. IV, we study how model parameters affect the onset of bistability. Finally, we discuss our results and present our conclusions in Sec. V.

II. MODEL

In this section, we present our hypergraph and contagion models. Our model consists of SIS contagion spreading on a hypergraph via pairwise and higher-order interactions. While we focus on the SIS epidemic model, we note that our formalism could be extended to other models. In the context of epidemic spreading, pairwise interactions could represent, for example, face-to-face interactions leading to contagion via viral droplets, while higher-order interactions could represent, for example, contagion via the shared spaces by a group. In the context of opinion dynamics, higher-order contagion could model, for example, a majority-vote process common in caucusing. In the following, we provide details about the hypergraph model representing the higher-order interactions and the contagion models that we consider.

A. Hypergraph model

We consider a population of nodes labeled $= 1, 2, \ldots,$ coupled via undirected hyperedges of sizes $= 2, 3, \ldots, M$, where a hyperedge of size is a set of nodes, $\{ _1, _2, \ldots, \}$. We define the d d g of node , $^{( )}$, as the number of hyperedges of size to which the node belongs, and its p d g as the vector $\mathbf{k} = [^{(2)}, ^{(3)}, \ldots, ^{(M)}]$. The 2nd order degree of a node corresponds to the number of pairwise connections of the node, while higher-order degrees measure the node's participation in hyperedges of larger sizes. Figure 1 illustrates a hypergraph with

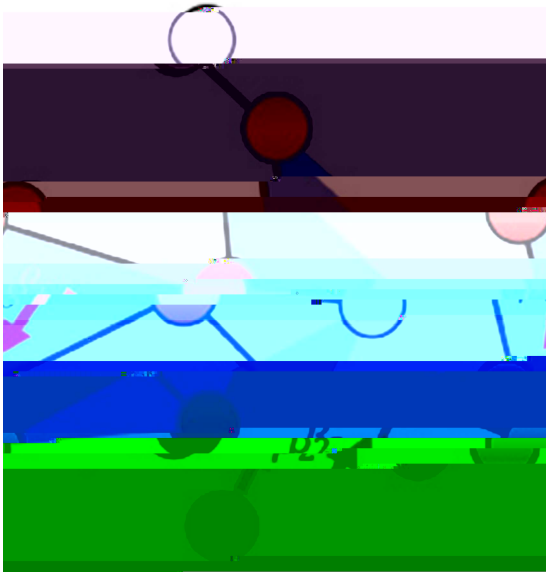


Illustration of a hypergraph Infected nodes red infect a healthy node gray via hyperedges of sizes and with rates β and β , respectively

hyperedges of sizes 2 and 3, which, for simplicity, we will henceforth denote as n and ng, respectively.

Extending degree-based descriptions of epidemic spreading on networks,[30,31] we will develop a mean-field theory for the propagation of epidemics based on the assumption that nodes with the same hyperdegree have the same statistical properties. For this purpose, we assume that the number of nodes with the hyperdegree $\mathbf{k}$, $(\mathbf{k})$, is given and that the probability that nodes with hyperdegrees $\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}$ belong to a hyperedge of size is given by $(\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k})$. This assumes that the statistical structure of the network is completely described by the hyperdegree distribution $(\mathbf{k})$ and the connection probabilities $(\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k})$. While this restriction rules out the possibility of assortative mixing by other node properties, it is straightforward to extend our formalism to include other node variables. Note that counting the number of hyperedges of size in two different ways, the connection probabilities must be normalized such that

$$\frac{1}{!} \sum_{\mathbf{k}_1, \ldots, \mathbf{k}} (\mathbf{k}_1), \ldots, (\mathbf{k}) \; (\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}) = \frac{1}{} \sum_{\mathbf{k}} {}^{( )} (\mathbf{k}). \tag{1}$$

For example, for the configuration model for networks without higher-order interactions (i.e., only hyperedges of size 2, $M = 2$), the hyperdegree of a node is just the number of links, $\mathbf{k} =$, connecting that node to other nodes and the connection probability is $_2(, ') = '/(\langle \rangle)$, where $\langle \rangle = \sum_{=1} / = \sum ( )/$. For networks with hyperedges of sizes 2 and 3, $_3(_1, _2, _3)$ is the probability that three nodes with degrees $_1$, $_2$, and $_3$ are connected by a hyperedge of size 3. The configuration model for hypergraphs

and its associated statistical properties has been studied in Refs. 32 and 33.

This framework allows us to study networks with heterogeneously distributed higher-order interactions and correlations between nodal degrees of different orders. In addition, it allows us to treat the case in which nodes belonging to a triangle are not necessarily connected by links, as is assumed in simplicial complex models.[24] We will study how the structure of higher-order interactions modifies some of the properties of epidemic spreading on networks with exclusively pairwise interactions (i.e., hyperedges of size 2 only), on which epidemic spreading has been studied extensively.[1]

B. Contagion model

Now, we describe the contagion models we will study. As mentioned above, we will focus on the SIS model, but other epidemic models could be treated using the same formalism. We assume that at any given time ≥ 0, each node can be in either the susceptible (S) or infected (I) state. Infected nodes heal and become susceptible again at rate γ. Now, we specify how hyperedges mediate the contagion process. In general, the probability of contagion by a hyperedge could be a function of the number of infected nodes in the hyperedge (e.g., as in Ref. 27). Here, we will consider the two extreme cases where contagion occurs if all the other members of the hyperedge are infected or if at least one member of the hyperedge is infected. More precisely, in the *co c con g on* case, a susceptible node that belongs to a hyperedge of size gets infected at rate β if the other members of the hyperedge are infected; in the *nd d con g on* case, the node gets infected at rate β if one member is infected. While we will analyze these two cases only, in principle, one could treat the case in which at least other nodes of the hyperedge need to be infected for contagion to occur using the techniques presented below. This case corresponds to a quorum of size , and there is evidence for such effects in collective behavior.[34,35] For hyperedges of size 2, i.e., links, both cases reduce to the usual contagion via pairwise interactions. The social contagion model of Ref. 24 corresponds to the collective contagion case. The contagion processes are illustrated in Fig. 1 for hyperedges of sizes 2 and 3. Table I summarizes the notation and variables used.

ABL Relevant notation

Variable	Definition
	Number of nodes
( )	Number of hyperedges of size a node belongs to
$\mathbf{k} = [{}^{(2)}, \ldots, {}^{(M)}]$	Hyperdegree
(k)	Number of nodes with hyperdegree k
γ	Rate of healing
β	Rate of infection by a hyperedge of size
$(\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}\ )$	Probability that nodes form a hyperedge of size
k	Fraction of nodes with hyperdegree k that are infected

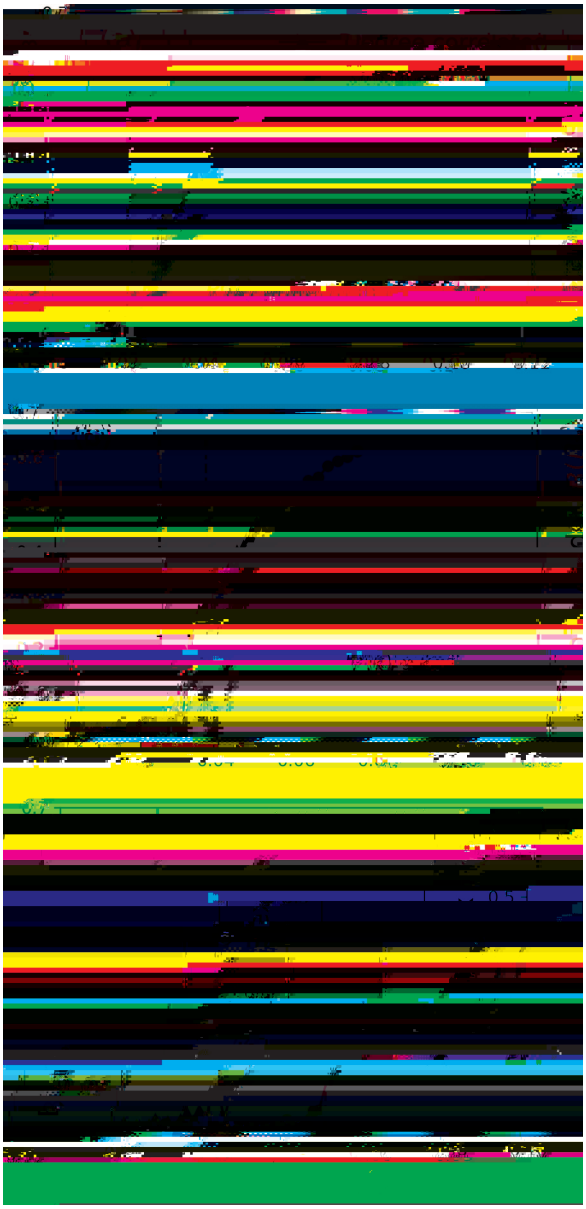
Now, we study the uncorrelated case where $_3(, _1, _2) = 2\langle \rangle / ^2$. In this case, Eq. (6) can be rewritten in terms of the fraction of infected nodes,

$$= \sum \frac{( )}{}, \tag{12}$$

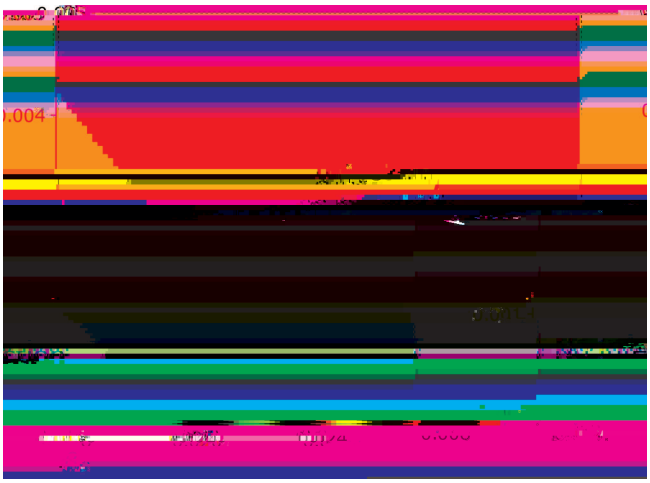
and the fraction of infected links . In terms of these quantities, Eq. (6) reads

d

the approaches proposed in Refs. 28 and 36 could be used to obtain better approximations. The qualitative aspects of interest, captured by the mean-field equations and the numerical solution of Eqs. (15) and (16), are the following. For small values of β_3 [Fig. 3(a), $\beta_3 = 0.0194$], the bifurcation from the state with no infection ($= 0$) to the infected state (> 0) is continuous. However, for larger values of β_3 [Fig. 3(c), $\beta_3 = 0.0582$], the transition is discontinuous: as β_2



Bistability index B as a function of β for a $P(k)$ constant for $< k$ <1 and otherwise, b $P(k) \propto k^{-}$ for 6 $< k <$



Phase diagram for the degree correlated, individual contagion model with parameters $\gamma =$ and $P(k) \propto k^{-}$ when 6 $< k <^{1}$ and otherwise

which has different first-order behavior than the degree-correlated case. Inserting this expression into Eqs. (7) and (12), we obtain

$$= \frac{1}{} \sum \frac{( )(\beta_2 \quad + 2\beta_3\langle \rangle \quad - \beta_3\langle \rangle^{2})}{\gamma + \beta_2 \quad + 2\beta_3\langle \rangle \quad - \beta_3\langle \rangle^{2}}, \tag{26}$$

$$= \frac{1}{\langle \rangle} \sum \frac{( )(\beta_2 \quad + 2\beta_3\langle \rangle \quad - \beta_3\langle \rangle^{2})}{\gamma + \beta_2 \quad + 2\beta_3\langle \rangle \quad - \beta_3\langle \rangle^{2}}. \tag{27}$$

Linearizing, we obtain the system

$$\delta \quad = \frac{\langle \rangle\beta_2}{\gamma}\delta \quad + \frac{2\langle \rangle\beta_3}{\gamma}\delta \quad , \tag{28}$$

$$\delta \quad = \frac{\langle^{2}\rangle\beta_2}{\langle \rangle\gamma}\delta \quad + \frac{2\langle \rangle\beta_3}{\gamma}\delta \quad . \tag{29}$$

Solving this system and canceling the zero solution, we find that the

hyperedges where

$(k, k_1,$

we find conditions under which there are at least two solutions to

$$(, \beta_2) = \frac{1}{\langle \ \rangle} \sum \frac{( )(\beta_2 \ + \beta_3 \ )}{\gamma + \beta_2 \ + \beta_3 \ \langle \ ^2 \rangle} - 1 = 0. \quad (35)$$

First, note that $(0, \beta_2) = \beta_2/\beta_2^c - 1$ and that $(1, \beta_2) < 0$. Therefore, if $\frac{\partial}{\partial}(0, \beta_2^c) > 0$, then by continuity, there will be at least two solutions for β_2 less than, but sufficiently close to, β_2^c. This condition gives

$$\frac{\beta_3^c}{\gamma} = \frac{\langle \ ^3 \rangle \langle \ \rangle^2}{\langle \ ^2 \rangle^3}, \quad (36)$$

which works well in predicting the onset of bistability for the degree-correlated case. The relative error with respect to the value obtained from directly solving Eq. (10) for all distributions tested is less than 2% (not shown).

The analysis for the degree-correlated case was based on the behavior of $(, \beta_2)$ near $= 0$. For the uncorrelated case, however, we find that a saddle-node bifurcation can occur at positive values of , and it is necessary to expand Eqs. (15) and (16) to higher order.

Expanding Eqs. (15) and (16) to second order, setting $\beta_2 = \beta_2^c = \gamma \langle \ \rangle / \langle \ ^2 \rangle$, and subtracting the two equations yield

$$= \frac{\langle \ \rangle^2}{\langle \ ^2 \rangle} + \left(\frac{\langle \ \rangle \langle \ ^3 \rangle}{\langle \ ^2 \rangle^2} - \frac{\langle \ \rangle^2}{\langle \ ^2 \rangle} \right) \ ^2, \quad (37)$$

which, when evaluated in

$$(, \beta_2) = \frac{1}{\langle \ \rangle} \sum \frac{( )(\beta_2 \ + \beta_3 \langle \ \rangle \ ^2)}{\gamma + \beta_2 \ + \beta_3 \langle \ \rangle \ ^2} - \ = 0 \quad (38)$$

and expanded to fourth order, again setting $\beta_2 = \beta_2^c$, yields

$$(, \beta_2^c) = (\alpha_0 + \alpha_1 \ + \alpha_2 \ ^2) \ ^2, \quad (39)$$

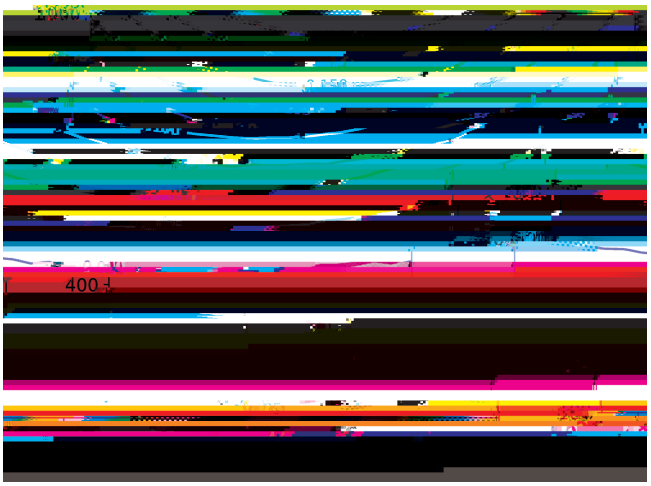
where

$$\alpha_0 = -\frac{\langle \ \rangle \langle \ ^3 \rangle}{\langle \ ^2 \rangle^2} + \frac{\langle \ \rangle^5 \beta_3}{\langle \ ^2 \rangle \gamma}, \quad (40)$$

$$\alpha_1 = \frac{\langle \ \rangle^2 \langle \ ^4 \rangle}{\langle \ ^2 \rangle^3} - 4\frac{\langle \ \rangle^5 \beta_3}{\langle \ ^2 \rangle^2 \gamma} + 2\frac{\langle \ \rangle^4 \langle \ ^3 \rangle \beta_3}{\langle \ ^2 \rangle^3 \gamma}, \quad (41)$$

$$\alpha_2 = -\frac{\langle \ \rangle^3 \langle \ ^5 \rangle}{\langle \ ^2 \rangle^4} + 5\frac{\langle \ \rangle^5 \beta_3}{\langle \ ^2 \rangle^2 \gamma} + 3\frac{\langle \ \rangle^6 \langle \ ^3 \rangle \beta_3}{\langle \ ^2 \rangle^4 \gamma} - 6\frac{\langle \ \rangle^4 \langle \ ^3 \rangle \beta_3}{\langle \ ^2 \rangle^3 \gamma} + \frac{\langle \ \rangle^3 \langle \ ^3 \rangle^2 \beta_3}{\langle \ ^2 \rangle^4 \gamma} - \frac{\langle \ \rangle^{10} \beta_3^2}{\langle \ ^2 \rangle^4 \gamma^2}. \quad (42)$$

For continuous transitions to epidemics, there is only one equilibrium for at $\beta_2 = \beta_2^c$, namely, $= 0$. The onset of bistability occurs when a second solution appears, which corresponds to the first appearance of a root of (39) in the interval $(0, 1)$. Such a root can appear at $= 0$ in a transcritical bifurcation or at > 0 as a pair of roots in a saddle-node bifurcation. A pair of roots appears when the discriminant of the quadratic equation $\alpha_0 + \alpha_1 \ + \alpha_2 \ ^2 = 0$ is zero. However, this bifurcation is physically meaningless if it occurs for values of outside the interval [0, 1]. Therefore, we impose the constraint that the value of β_3 found by solving $\alpha_1^2 - 4\alpha_0\alpha_2 = 0$ must satisfy the inequality $0 \leq -\alpha_1/2\alpha_2 \leq 1$. In addition, we



Relative error in the value of β^c/β^c obtained from Eq compared with the numerically obtained value shown in Fig b

note that because of continuity, the sign of the α_2 term must be negative because otherwise, $\frac{\partial}{\partial}(0, \beta_2^c) > 0$ and the bifurcation has already occurred. The transcritical bifurcation occurs when a root crosses from a negative value to a positive value, which occurs when one root of $\alpha_0 + \alpha_1 \ + \alpha_3 \ ^2 = 0$ is $= 0$, implying that $\alpha_0 = 0$ and $\beta_3^c = \gamma \langle \ ^3 \rangle / \langle \ \rangle^4$. Using these conditions, we can construct a piecewise definition of β_3^c

$$\beta_3^c = \begin{cases} \text{Solve}(\alpha_1^2 - 4\alpha_0\alpha_2 = 0), & \alpha_2 < 0, 0 \leq -\frac{\alpha_1}{2\alpha_2} \leq 1, \\ \frac{\langle \ ^3 \rangle}{\langle \ \rangle^4}\gamma, & \text{else.} \end{cases} \quad (43)$$

The relative error in the value of β_3^c/β_2^c obtained from Eq.

c

pairwise contact network, suppress discontinuous transitions. Conversely, heterogeneity in the degree distribution of hyperedges of higher order promotes such transitions.

At $= 0$, the network is randomly and uniformly seeded with a small fraction ($p = 0.001$) of infected nodes, and at each subsequent step, the current state is iterated as described above and the population average, $= \sum_{=1} X /$, is stored. To avoid the absorbing state $\mathrm{X} =$, we infect a single randomly chosen node if the population becomes completely healthy. To mitigate the effect of variability in the stochastic simulation, we average the time response of over a sufficient time window (determined from the average infected response curves) after it reached the steady-state. In this study, we ran the simulation for a fixed set of parameters $\{\gamma, \beta_2, \beta_3\}$