

The effect of heterogeneity on hypergraph contagion models

Nicholas W. Landry^{1, a)} and Juan G. Restrepo^{1, b)}

Department of Applied Mathematics, University of Colorado at Boulder, Boulder, Colorado 80309

(Dated: 15 September 2020)

The dynamics of network social contagion processes such as opinion formation and epidemic spreading are often mediated by interactions between multiple nodes. Previous results have shown that these higher-order interactions can profoundly modify the dynamics of contagion processes, resulting in bistability, hysteresis, and explosive transitions. In this paper, we present and analyze a hyperdegree-based mean-field description of the dynamics of the SIS model on hypergraphs, i.e. networks with higher-order interactions, and illustrate its applicability with the example of a hypergraph where contagion is mediated by both links (pairwise interactions) and triangles (three-way interactions). We consider various models for the organization of link and triangle structure, and different mechanisms of higher-order contagion and healing. We find that explosive transitions can be suppressed by heterogeneity in the link degree distribution, when links and triangles are chosen independently, or when link and triangle connections are positively correlated when compared to the uncorrelated case. We verify these results with microscopic simulations of the contagion process and with analytic predictions derived from the mean-field model. Our results show that the structure of higher-order interactions can have important effects on contagion processes on hypergraphs.

Including group interactions in network models of contagion can significantly affect epidemic behavior. By studying the susceptible-infected-susceptible epidemic model on networks with higher-order interactions, we observe that for certain parameters there is a bistable regime, where above a critical number of infected individuals the contagion spreads until it becomes an epidemic, and below this critical number the epidemic dies out. We find that heterogeneity in the individual and group contact structure of a social network determines the existence of such tipping point events, and derive conditions for their appearance. Lastly, we comment on how three group contagion mechanisms – collective contagion, infection by individuals, and the “hipster effect” – affect the onset of epidemics and the existence of bistability.

I. INTRODUCTION

The study of contagion processes is a fundamental problem in network science, with applications including epidemics^{1–7}, social media⁸, opinion formation⁹, idea diffusion^{10,11}, sudden changes in social convention^{12,13}, and many more. Contagion processes can be of many types, ranging from discrete-state models such as the susceptible-infected-susceptible (SIS) model, to continuous models of opinion formation, to realistic models of disease such as those currently used to model the spread of COVID-19^{14,15}. Modeling the dynamics of such processes on pairwise interaction networks has been a hallmark of network science, providing many insights into the effect of network structure on the propagation of disease and information. Recently, the role of *complex contagion* mechanisms (i.e., contagion processes that can not be described solely by pairwise interactions) has received much

attention¹⁶. It has been shown that higher-order interactions in networks (i.e., interactions involving multiple nodes) can have profound effects on dynamical network processes¹⁷ such as opinion formation¹⁸, synchronization^{19–21} and population dynamics²². Efforts to map higher-order interactions in real-world networks have uncovered rich structure²³ which is only now starting to be appreciated. In the context of contagion processes, it was recently shown²⁴ that the addition of higher-order interactions to the SIS epidemic model on Erdős-Rényi networks result in bistability, hysteresis, and explosive transitions to an endemic disease state (see also Refs.^{25–28}). The simplicial SIS model has also been extended to scale-free uniform hypergraphs²⁹. 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 *hypergraph*, 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 structure 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

^{a)}Electronic mail: nicholas.landry@colorado.edu.

^{b)}Electronic mail: juanga@colorado.edu

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 N nodes labeled $i = 1, 2, \dots, N$ coupled via undirected hyperedges of sizes $m = 2, 3, \dots, M$, where a hyperedge of size m is a set of m nodes, $\{i_1, i_2, \dots, i_m\}$. We define the m -th order degree of node i , $k_i^{(m)}$, as the number of hyperedges of size m to which the node belongs, and its *hyperdegree* as the vector $\mathbf{k}_i = [k_i^{(2)}, k_i^{(3)}, \dots, k_i^{(M)}]$. 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 hyperedges of sizes 2 and 3, which, for simplicity,

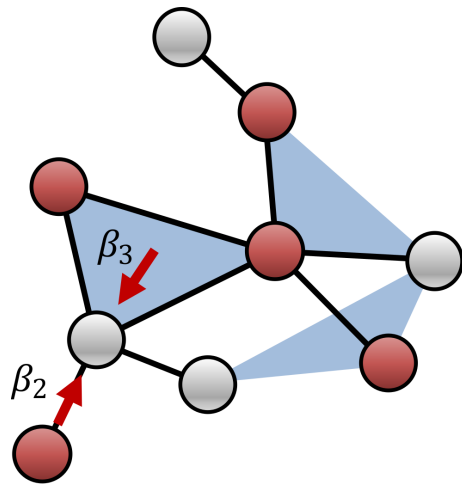


FIG. 1. Illustration of a hypergraph. Infected nodes (red) infect a healthy node (grey) via hyperedges of sizes 2 and 3 with rates β_2 and β_3 respectively.

we will henceforth denote as *links* and *triangles* 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 hyperdegree \mathbf{k} , $P(\mathbf{k})$, is given, and that the probability that nodes with hyperdegrees $\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_m$ belong to a hyperedge of size m is given by $f_m(\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_m)$. This assumes that the statistical structure of the network is completely described by the hyperdegree distribution $P(\mathbf{k})$ and the connection probabilities $f_m(\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_m)$. 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 m in two different ways, the connection probabilities must be normalized such that

$$\frac{1}{m!} \sum_{\mathbf{k}_1, \dots, \mathbf{k}_m} P(\mathbf{k}_1) \dots P(\mathbf{k}_m) f_m(\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_m) = \frac{1}{m} \sum_{\mathbf{k}} k^{(m)} P(\mathbf{k}). \quad (1)$$

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, $\mathbf{k} = k$, connecting that node to other nodes and the connection probability is $f_2(k, k') = kk' / (N \langle k \rangle)$, where $\langle k \rangle = \sum_{i=1}^N k_i / N = \sum_k k P(k) / N$. For networks with hyperedges of sizes 2 and 3, $f_3(k_1, k_2, k_3)$ is the probability that three nodes with degrees k_1, k_2 , and k_3 are connected by a hyperedge of size 3. The configuration model for hypergraphs and its associated statistical properties has been studied in Refs.^{32,33}.

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 models²⁴. 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 extensively¹.

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 $t \geq 0$, 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 *collective contagion* case,

a susceptible node that belongs to a hyperedge of size m gets infected at rate β_m if *all* the other members of the hyperedge are infected; in the *individual contagion* case the node gets infected at rate β_m if *at least* one member is infected. While we will analyze these two cases only, in principle one could treat the case in which at least j 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 j 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 Figure 1 for hyperedges of sizes 2 and 3. Table I summarizes the notation and variables used.

III. MEAN-FIELD ANALYSIS

In this section we present a mean-field analysis of the epidemic dynamics on a network specified by the hyperdegree distribution $P(\mathbf{k})/N$ and the hyperedge connection probabilities $f_m(\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_m)$. Assuming that all nodes with the same

hyperdegree behave similarly, we focus on $x_{\mathbf{k}}$, the fraction of nodes with hyperdegree \mathbf{k} that are infected. The mean-field equation describing the evolution of $x_{\mathbf{k}}$ is

Variable	Definition
N	Number of nodes
$k^{(m)}$	number of hyperedges of size m a node belongs to
$\mathbf{k} = [k^{(2)}, \dots, k^{(M)}]$	hyperdegree
$P(\mathbf{k})$	Number of nodes with hyperdegree \mathbf{k}
γ	Rate of healing
β_m	Rate of infection by a hyperedge of size m
$f_m(\mathbf{k}_1, \mathbf{k}_2, \dots, \mathbf{k}_m)$	Probability that m nodes form a hyperedge of size m
$x_{\mathbf{k}}$	fraction of nodes with hyperdegree \mathbf{k} that are infected

TABLE I. Relevant notation

$$\frac{dx_{\mathbf{k}}}{dt} = -\gamma x_{\mathbf{k}} + (1 - x_{\mathbf{k}}) \sum_{m=2}^M \beta_m \frac{1}{(m-1)!} \sum_{\mathbf{k}_1, \dots, \mathbf{k}_{m-1}} \prod_{l=1}^{m-1} P(\mathbf{k}_l) f_m(\mathbf{k}, \mathbf{k}_1, \dots, \mathbf{k}_{m-1}) G(x_{\mathbf{k}_1}, \dots, x_{\mathbf{k}_{m-1}}) \quad (2)$$

$$G(x_{\mathbf{k}_1}, \dots, x_{\mathbf{k}_{m-1}}) = \begin{cases} \prod_{l=1}^{m-1} x_{\mathbf{k}_l}, & \text{collective contagion,} \\ 1 - \prod_{l=1}^{m-1} (1 - x_{\mathbf{k}_l}), & \text{individual contagion.} \end{cases} \quad (3)$$

The first term on the right-hand side of Eq. (2) corresponds to healing at rate γ and the second term accounts for infection by hyperedges. The number of hyperedges of size m that can pass an infection to a node with hyperdegree \mathbf{k} is calculated by considering all the possible hyperdegrees of the other $m-1$ nodes participating in the hyperedge $(\mathbf{k}, \mathbf{k}_1, \dots, \mathbf{k}_{m-1})$, counting how many such combinations there are not counting permutations $[P(\mathbf{k}_1) \cdots P(\mathbf{k}_{m-1}) / (m-1)!]$, calculating what fraction of such combinations form a hyperedge with the node in consideration $[f_m(x_{\mathbf{k}}, x_{\mathbf{k}_1}, \dots, x_{\mathbf{k}_{m-1}})]$, multiplying by the probability that the hyperedge can transmit the infection $[G(x_{\mathbf{k}_1} \cdots x_{\mathbf{k}_{m-1}})]$, and summing over all hyperdegree combinations. The probability that the hyperedge can transmit the infection, given by (3), depends on whether the collective contagion or individual contagion model is assumed. Note that the form for G taken above, and the mean field treatment in general, assume that the states of nodes are independent. A better approximation that includes correlations between connected nodes has been implemented in Ref.²⁸ for the case of unstructured hyperedges of sizes 2 and 3, leading to improved quantitative agreement with the results of numerical simula-

tions. Since our interest is in the effects of higher-order structures on qualitative aspects of the epidemic dynamics, we will use the mean field approximation in Eq. (2). A similar mean-field equation for a node-based description of the contagion process was recently formulated in Ref.²⁶. In the following we will apply the mean-field description to illustrative cases.

A. Hyperedges of sizes 2 and 3 with collective contagion

Here we focus on the case where the hyperedge sizes are either 2 or 3, i.e., $M = 3$. This corresponds to a network like in Fig. 1, with hyperedges of size 2 (links) and 3 (triangles). For simplicity, here we denote the number of links per node as k , i.e., $k = k^{(2)}$, and the number of triangles a node belongs to by q , i.e., $q = k^{(3)}$. In addition, we will consider the case where the connection probabilities depend only on the node links, i.e., $f_m(\mathbf{k}, \mathbf{k}_1, \dots, \mathbf{k}_{m-1}) = f_m(k, k_1, \dots, k_{m-1})$. With these assumptions, and using the collective contagion rule in Eq. (3), Eq. (2) becomes

$$\frac{dx_{k,q}}{dt} = -\gamma x_{k,q} + (1 - x_{k,q})\beta_2 \sum_{k_1,q_1} P(k_1,q_1)f_2(k,k_1)x_{k_1,q_1} + (1 - x_{k,q})\frac{\beta_3}{2} \sum_{k_1,q_1,k_2,q_2} P(k_1,q_1)P(k_2,q_2)f_3(k,k_1,k_2)x_{k_1,q_1}x_{k_2,q_2}, \quad (4)$$

where the first term on the right hand side represents healing, the second represents contagion by links, and the third represents contagion by triangles.

Since the connection probabilities do not depend on q , we can reduce the dynamics to the fraction of nodes with degree k that are infected,

$$x_k = \frac{\sum_q P(k,q)x_{k,q}}{P(k)}, \quad (5)$$

where $P(k) = \sum_q P(k,q)$ is the number of nodes with degree k . Multiplying Eq. (4) by $P(k,q)$, summing over q and dividing by $P(k)$, we obtain

$$\begin{aligned} \frac{dx_k}{dt} = & -\gamma x_k + (1 - x_k)\beta_2 \sum_{k_1} P(k_1)f_2(k,k_1)x_{k_1} \\ & + (1 - x_k)\frac{\beta_3}{2} \sum_{k_1,k_2} P(k_1)P(k_2)f_3(k,k_1,k_2)x_{k_1}x_{k_2}. \end{aligned} \quad (6)$$

For the link connection probability $f_2(k,k_1)$, we will take $f_2(k,k_1) = kk_1/(N\langle k \rangle)$, which corresponds to nodes being connected completely at random according to their degree as in the configuration model. For the triangle connection probability f_3 , we will consider two cases: the *uncorrelated* case and the *degree-correlated* case. In the degree-correlated case, we assume that the connection probability is given by $f_3(k,k_1,k_2) = 2kk_1k_2/(N\langle k \rangle)^2$, so that nodes that have a higher number of links also belong to more triangles. In the uncorrelated case, we assume instead that $f_3(k,k_1,k_2) = 2\langle k \rangle/N^2$, so that triangles are formed independent of the nodal degrees. The normalization is chosen using Eq. (1) so that the mean number of triangles per node, $\langle q \rangle = \sum_{i=1}^N k_i^3/N$, in each case is equal to $\langle k \rangle$. We note that the model for triangle formation in Ref.²⁴ corresponds to the uncorrelated case. We can choose the mean triangle degree independent of the mean network degree by scaling $f_3(k,k_1,k_2)$ by $\langle q \rangle/\langle k \rangle$, but for simplicity, we assume $\langle q \rangle = \langle k \rangle$. Figure 2 illustrates the difference between the two cases in a small network, where in the degree-correlated case, the triangles cluster around nodes with high pairwise degree, and in the uncorrelated case, the triangles are distributed uniformly at random on the network.

We can also specify the distribution of triangle degrees by defining $f_2(q,q_1)$ and $f_3(q,q_1,q_2)$ and then reducing Eq. (4) by multiplying by $P(k,q)$, dividing by $P(q)$, and summing over k to reduce the dynamics to the fraction of infected nodes with triangle degree q . For the triangle connection probability, we take $f_3(q,q_1,q_2) = 2qq_1q_2/(N\langle q \rangle)^2$ which corresponds to three nodes being connected at random according to the configuration model for triangles³². For the pairwise links, we define the degree-correlated and uncorrelated cases as before, where in the degree-correlated case, $f_2(q,q_1) = qq_1/(N\langle q \rangle)$,

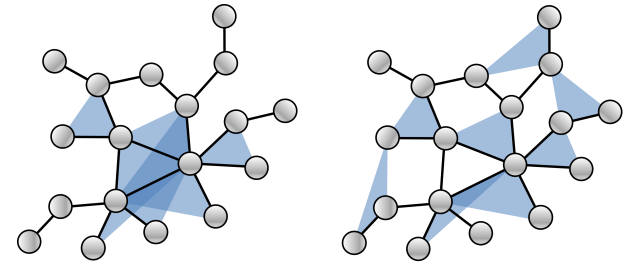


FIG. 2. Schematic illustration of the degree-correlated and uncorrelated cases. In the degree-correlated case (left), nodes with more links are more likely to belong to a triangle. In the uncorrelated case (right), triangles connect nodes with a probability independent of their degree.

and for the uncorrelated case, $f_2(q,q_1) = \langle q \rangle/N$. From there, we can use the same formalism as our approach when specifying the pairwise degree.

Now we consider separately the degree-correlated and uncorrelated cases. In the correlated case, where $f_3(k,k_1,k_2) = kk_1k_2/(N\langle k \rangle)^2$, Eq. (6) can be rewritten in terms of the fraction of infected links

$$V = \sum_k \frac{kP(k)x_k}{N\langle k \rangle} \quad (7)$$

as

$$\frac{dx_k}{dt} = -\gamma x_k + \beta_2(1 - x_k)kV + \beta_3(1 - x_k)kV^2. \quad (8)$$

In this case, the dynamics of nodes of degree k is determined by the global variable V . To study the qualitative characteristics of the dynamics, we find the steady-state solutions. The fixed point of Eq. (8) is

$$x_k = \frac{\beta_2kV + \beta_3kV^2}{\gamma + \beta_2kV + \beta_3kV^2}. \quad (9)$$

Inserting this in (7), we obtain a nonlinear equation that determines the fraction of infected links V :

$$V = \frac{1}{N\langle k \rangle} \sum_k \frac{kP(k)(\beta_2kV + \beta_3kV^2)}{\gamma + \beta_2kV + \beta_3kV^2}. \quad (10)$$

The state with no infection, $V = 0$, is a solution to (10). However, it is linearly unstable for $\beta_2 > \beta_2^c = \gamma\langle k \rangle/\langle k^2 \rangle$, as can be seen by linearizing Eq. (8) about $V = 0$, multiplying by $kP(k)/(N\langle k \rangle)$, and summing over k , which yields the linearized equation for the evolution of the perturbation δV

$$\frac{d\delta V}{dt} = -\gamma\delta V + \beta_2 \frac{\langle k^2 \rangle}{\langle k \rangle} \delta V. \quad (11)$$

The nonzero solutions of Eq. (10) represent states with a nonzero fraction of infected nodes.

Now we study the uncorrelated case where $f_3(k, k_1, k_2) = 2\langle k \rangle / N^2$. In this case, Eq. (6) can be rewritten in terms of the fraction of infected nodes

$$U = \sum_k \frac{P(k)x_k}{N}, \quad (12)$$

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

$$\frac{dx_k}{dt} = -\gamma x_k + \beta_2(1 - x_k)kV + \beta_3(1 - x_k)\langle k \rangle U^2. \quad (13)$$

As in the prior case, the equilibrium is

$$x_k = \frac{\beta_2 k V + \beta_3 \langle k \rangle U^2}{\gamma + \beta_2 k V + \beta_3 \langle k \rangle U^2}. \quad (14)$$

Evaluating this expression in Eqs. (7) and (12) we obtain the coupled equations

$$U = \frac{1}{N} \sum_k \frac{P(k)(\beta_2 k V + \beta_3 \langle k \rangle U^2)}{\gamma + \beta_2 k V + \beta_3 \langle k \rangle U^2}, \quad (15)$$

$$V = \frac{1}{N \langle k \rangle} \sum_k \frac{k P(k)(\beta_2 k V + \beta_3 \langle k \rangle U^2)}{\gamma + \beta_2 k V + \beta_3 \langle k \rangle U^2}. \quad (16)$$

The state with no infection, $U = 0$, $V = 0$, is a solution of (15)-(16). By considering perturbations δU , δV from this solution, linearizing Eq. (13), and evaluating in Eq. (7) for the first equation and Eq. (12) for the second equation, we obtain the linear system

$$\frac{d\delta V}{dt} = -\gamma \delta V + \beta_2 \frac{\langle k^2 \rangle}{\langle k \rangle} \delta V, \quad (17)$$

$$\frac{d\delta U}{dt} = -\gamma \delta U + \beta_2 \langle k \rangle \delta V, \quad (18)$$

which shows that the no infection state is linearly unstable for $\beta_2 > \gamma \langle k \rangle / \langle k^2 \rangle$, which is the same threshold we obtained for the correlated case.

In summary, nonzero solutions of Eq. (10) and Eqs. (15)-(16) for the degree-correlated and uncorrelated cases, respectively, represent states with a nonzero number of infected nodes. Figure 3 shows the fraction of infected nodes U for the uncorrelated case as a function of the normalized pairwise infectivity β_2/β_2^c for three values of the triangle infectivity β_3 obtained from numerical solution of Eqs. (15)-(16) with $P(k) \propto k^{-4}$ for $67 < k < 1000$ and 0 otherwise. Different solutions are plotted as solid and dashed lines to indicate stability or instability, respectively. The connected circles are obtained from numerical simulations of the full stochastic microscopic model. In these simulations β_2 was slowly increased in small steps up to a maximum value, and subsequently decreased back to its initial value. For each β_2 , the average number of infected nodes after transient effects disappeared is shown as a filled circle. For more details about the simulations, see Appendix A.

The behavior of the microscopic simulations is captured qualitatively by the mean field equations. The quantitative disagreement is likely due to the assumptions inherent to the mean-field approximation. In fact, Ref.²⁸ has shown that, for the particular case of uncorrelated triangles on an Erdős-Rényi network, the disagreement almost disappears when pair correlations are taken into account. Since our interest in this paper is on the qualitative dynamics, we use the mean-field theory, but note that the approaches proposed in^{28,36} 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)-(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 ($U = 0$) to the infected state ($U > 0$) is continuous. However, for larger values of β_3 [Fig. 3(c), $\beta_3 = 0.0582$], the transition is discontinuous: as β_2 increases past a critical value β_2^c , the fraction of infected links increases explosively towards an epidemic equilibrium (upward arrow). If β_2 is subsequently decreased, the fraction of infected links remains high until β_2 decreases past the value at which the epidemic equilibrium solution disappears, and then it decreases to zero (downward arrow). For such values of β_3 , there is hysteresis, bistability, and explosive transitions. At a critical value $\beta_3 = \beta_3^c$, which will be the focus of our interest, there is a transition from the type of bifurcation shown in Fig 3(a) to the type of bifurcation shown in Fig. 3(c). Fig. 3(b) shows U as a function of β_2 for a value $\beta_3 = 0.0388 \approx \beta_3^c$. We are interested in exploring how the presence of this bistable regime is affected by the degree distribution $P(k)$ and other parameters of the model, in particular the triangle infectivity, β_3 .

Figure 4 shows the phase diagram in the (β_2, β_3) plane for the degree-correlated, collective contagion model. The plot was obtained by counting the number of solutions of Eq. (10) as a function of β_2 and β_3 for $\gamma = 2$, and $P(k) \propto k^{-4}$ when $67 < k < 1000$ and 0 otherwise (all subsequent phase diagram plots are calculated using the same parameters). Light pink indicates that there is only the solution $V = 0$ corresponding to a stable state with no contagion. Orange indicates two solutions, the unstable $V = 0$ solution and another stable solution with $V > 0$. Finally, dark red indicates a bistable regime with three solutions: the stable $V = 0$ solution, and a pair of stable and unstable solutions with positive V . As noted in Ref.^{24,26}, this regime is only present for large enough triangle infectivity, i.e., for $\beta_3 > \beta_3^c$. The phase space for the uncorrelated case (not shown) is qualitatively similar to the one in Fig. 4, but the transition to bistable behavior occurs at a larger value of β_3 .

To quantify how the onset of bistability depends on the hypergraph parameters, we define the *bistability index* $B(\beta_3)$ as the maximum separation, over all values of β_2 , between the largest and smallest stable solutions for the fraction of infected nodes U . The bistability index can be calculated from microscopic simulations of the contagion process such as those used to produce Fig. 3, or from numerical solution of Eq. (10) for the correlated case and Eqs. (15)-(16) for the uncorrelated case. In Figure 5 we plot the bistability index B as a function of β_3 computed from microscopic simulations for three choices of the link degree distribution $P(k)$, all with a mean degree of 100: (a) $P(k)$ constant for $50 < k < 150$ and 0 other-

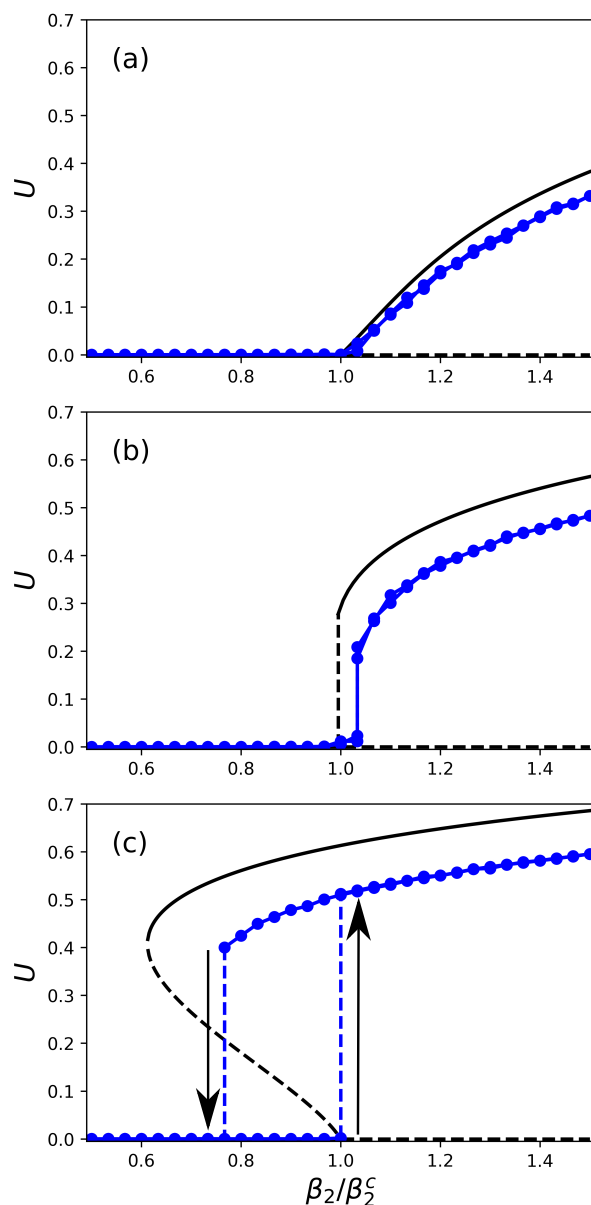


FIG. 3. Fraction of infected nodes U versus link infectivity β_2 obtained from the mean field equations (15)-(16) (solid and dashed lines) and from microscopic simulations (connected circles) using $P(k) \propto k^{-4}$ on $[67, 1000]$, $\gamma = 2$, and $N = 10000$ for $\beta_3 = 0.0194$ (a), 0.0388 (b), and 0.05482 (c). Refer to the text for explanation of the discrepancy between the mean field equations and microscopic simulations.

wise, (b) $P(k) \propto k^{-4}$ for $67 < k < 1000$ and 0 otherwise, and (c) $P(k) \propto k^{-3}$ for $53 < k < 1000$ and 0 otherwise. For each distribution, we considered the uncorrelated case (orange connected circles) and the degree correlated case (blue connected triangles). The dashed lines indicate the value β_3^c at which we expect the onset of bistability, obtained from numerical solution of Eqs. (10) and (15)-(16) for the degree correlated and uncorrelated cases, respectively (in Sec. IV we provide analytical expressions for these values). As the degree distribu-

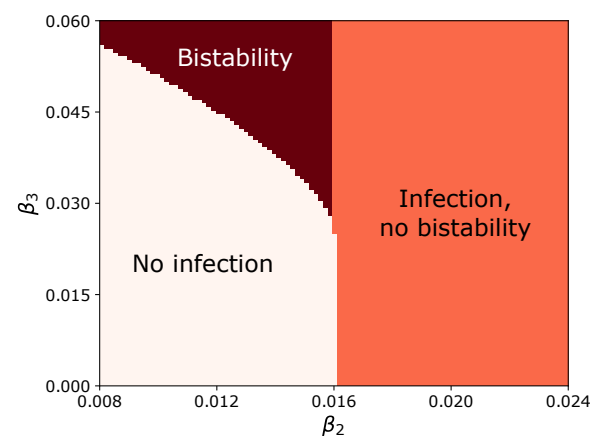


FIG. 4. Phase diagram for the degree-correlated, collective contagion model. The light pink region labeled “No infection” corresponds to 1 solution of Eq. (10), the orange region labeled “Infection, no bistability” to 2 solutions, and the region labeled “Bistability” to 3 solutions. The parameters are $\gamma = 2$ and $P(k) \propto k^{-4}$ when $67 < k < 1000$ and 0 otherwise.

tion of the pairwise interaction network $P(k)$ becomes more heterogeneous from (a) to (c), the value of β_3 at which the onset of bistability occurs increases for the uncorrelated case, while it remains almost unchanged for the degree-correlated case. A heuristic interpretation of this phenomenon is the following: in the uncorrelated case, the triangle interactions do not depend on the heterogeneity of the link degree distribution. Therefore, as the link degree distribution $P(k)$ becomes more heterogeneous, contagion becomes dominated by hubs of the pairwise interaction network, a mechanism which does not result in bistability. Therefore, bistability is suppressed in the uncorrelated case. On the other hand, for the degree correlated case, both triangle and link contagion mechanisms increase their effectiveness in tandem as the heterogeneity of the link degree distribution is increased. It is important to note that the increase in β_3^c with heterogeneity, which is shown here in absolute terms, still occurs if one considers it relative to the value of β_2^c (i.e., β_3^c/β_2^c also increases with heterogeneity), as we will show later.

Another interesting aspect seen in Figure 5 is that the transition to bistable behavior seems sharper in the uncorrelated case for the more heterogeneous networks. As we will see in Sec. IV, the nature of the bifurcation is indeed different for the uncorrelated case and heterogeneous networks.

Finally, we have to point out that the numerical calculation of the bistability index from numerical simulations can be challenging. When the unstable solution is small, finite size effects can cause transitions to the nonzero stable solution from the stable zero solution, making the numerical determination of the stable fixed points difficult and the bistability index plots noisy. Nevertheless, the mean field theory predicts well the onset of bistability.

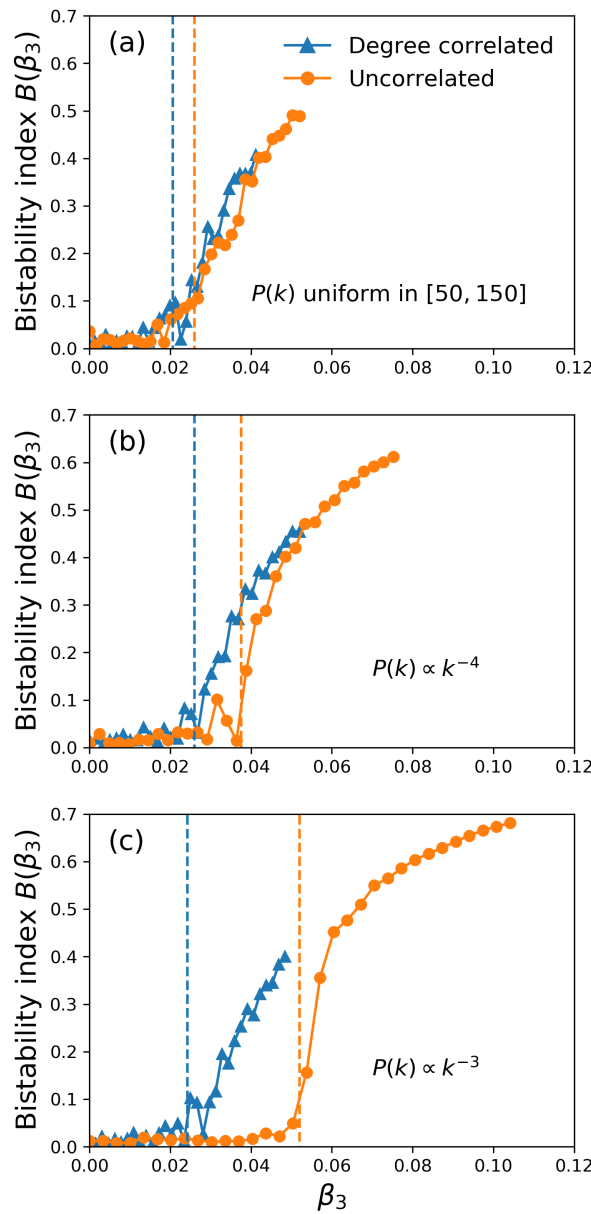


FIG. 5. Bistability index B as a function of β_3 for (a) $P(k)$ constant for $50 < k < 150$ and 0 otherwise, (b) $P(k) \propto k^{-4}$ for $67 < k < 1000$ and 0 otherwise, and (c) $P(k) \propto k^{-3}$ for $53 < k < 1000$ and 0 otherwise. For each distribution, we considered the uncorrelated case (orange connected circles) and the degree correlated case (blue connected triangles). The dashed lines indicate the value β_3^c at which we expect the onset of bistability, obtained from numerical solution of the mean field equations (12) and (15)-(16).

B. Hyperedges of sizes 2 and 3 with individual contagion

Now we consider the case of individual contagion, in which an m -hyperedge infects a susceptible node with rate β_m when at least one member of the hyperedge is infected. For simplicity, we will still consider only links and triangles ($M = 3$) with infection rates of β_2 and β_3 respectively.

The analogue to Eq. (6) for the individual contagion case is

$$\frac{dx_k}{dt} = -\gamma x_k + (1 - x_k) \beta_2 \sum_{k_1} P(k_1) f_2(k, k_1) x_{k_1} + (1 - x_k) \frac{\beta_3}{2} \sum_{k_1, k_2} P(k_1) P(k_2) f_3(k, k_1, k_2) [1 - (1 - x_{k_1})(1 - x_{k_2})]. \quad (19)$$

For the correlated case, $f_3(k, k_1, k_2) = 2kk_1k_2/(N\langle k \rangle)^2$, this can be rewritten as

$$\frac{dx_k}{dt} = -\gamma x_k + (\beta_2 + 2\beta_3)(1 - x_k)kV - \beta_3(1 - x_k)kV^2, \quad (20)$$

with fixed point

$$x_k = \frac{(\beta_2 + 2\beta_3)kV - \beta_3kV^2}{\gamma + (\beta_2 + 2\beta_3)kV - \beta_3kV^2}. \quad (21)$$

Inserting this into Eq. (7) like before, we obtain

$$V = \frac{1}{N\langle k \rangle} \sum_k \frac{kP(k)[(\beta_2 + 2\beta_3)kV - \beta_3kV^2]}{\gamma + (\beta_2 + 2\beta_3)kV - \beta_3kV^2}. \quad (22)$$

Linearizing about the $V = 0$ equilibrium, we find that the epidemic threshold is given by the condition

$$\beta_2 + 2\beta_3 = \gamma \frac{\langle k \rangle}{\langle k^2 \rangle}, \quad (23)$$

which defines a linear relationship between β_2 and β_3 for fixed γ , in contrast to the collective contagion mechanism which does not alter the epidemic threshold $\beta_2^c = \gamma\langle k \rangle / \langle k^2 \rangle$. This relationship can be understood heuristically by noting that, close to the $V = 0$ solution, the probability that two nodes in a hyperedge are simultaneously infected can be neglected. Under that assumption, infection of a susceptible node by a triangle when at least one other node is infected is equivalent to independent infection by either of the two other nodes in the triangle with rate β_3 . Since in the correlated case a node belongs, on average, to the same number of links and triangles, the individual contagion model reduces to the traditional SIS model with contagion rate $\beta_2^{\text{eff}} = \beta_2 + 2\beta_3$ in the linear regime (we emphasize, however, that the nonlinear behavior can be different).

In Fig. 6 we plot the (β_2, β_3) phase space for this scenario, with light pink indicating one solution ($V = 0$) to Eq. (22) and orange indicating two solutions, the unstable $V = 0$ solution and a stable $V > 0$ solution.

Considering the uncorrelated case where $f(k, k_1, k_2) = 2\langle k \rangle / N^2$, and expressing Eq. (19) in terms of U and V , we obtain

$$\frac{dx_k}{dt} = -\gamma x_k + \beta_2(1 - x_k)kV + 2\beta_3(1 - x_k)\langle k \rangle U - \beta_3(1 - x_k)\langle k \rangle U^2, \quad (24)$$

with equilibrium

$$x_k = \frac{\beta_2kV + 2\beta_3\langle k \rangle U - \beta_3\langle k \rangle U^2}{\gamma + \beta_2kV + 2\beta_3\langle k \rangle U - \beta_3\langle k \rangle U^2}, \quad (25)$$

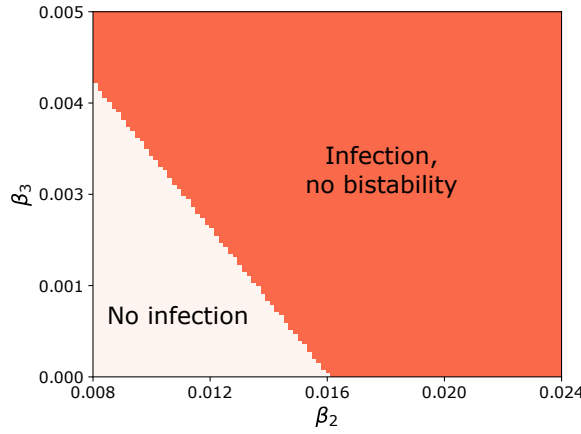


FIG. 6. Phase diagram for the degree-correlated, individual contagion model with parameters $\gamma = 2$ and $P(k) \propto k^{-4}$ when $67 < k < 1000$ and 0 otherwise.

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

$$U = \frac{1}{N} \sum_k \frac{P(k)(\beta_2 k V + 2\beta_3 \langle k \rangle U - \beta_3 \langle k \rangle U^2)}{\gamma + \beta_2 k V + 2\beta_3 \langle k \rangle U - \beta_3 \langle k \rangle U^2}, \quad (26)$$

$$V = \frac{1}{N \langle k \rangle} \sum_k \frac{k P(k)(\beta_2 k V + 2\beta_3 \langle k \rangle U - \beta_3 \langle k \rangle U^2)}{\gamma + \beta_2 k V + 2\beta_3 \langle k \rangle U - \beta_3 \langle k \rangle U^2}. \quad (27)$$

Linearizing, we obtain the system

$$\delta U = \frac{\langle k \rangle \beta_2}{\gamma} \delta V + \frac{2 \langle k \rangle \beta_3}{\gamma} \delta U, \quad (28)$$

$$\delta V = \frac{\langle k^2 \rangle \beta_2}{\langle k \rangle \gamma} \delta V + \frac{2 \langle k \rangle \beta_3}{\gamma} \delta U. \quad (29)$$

Solving this system and canceling the zero solution, we find that the epidemic threshold is defined by a non-linear relationship between the three epidemic parameters

$$\beta_2 = \frac{\langle k \rangle \gamma^2 - 2 \langle k^2 \rangle^2 \gamma \beta_3}{\langle k^2 \rangle \gamma - 2 \langle k^2 \rangle - \langle k \rangle^2 \langle k \rangle \beta_3} \quad (30)$$

This relationship implies that there is a singularity when $\beta_3 = \beta_3^* = \gamma \langle k^2 \rangle / [2(\langle k^2 \rangle - \langle k \rangle^2) \langle k \rangle]$. However, one can check that β_2 is negative at $\beta_3 = \beta_3^*$, and therefore the singularity is not physically relevant. Note that when $\langle k^2 \rangle = \langle k \rangle^2$ in the case of a k -regular network, the threshold reduces to that of the degree-correlated case.

C. Higher-order healing: hipster effect

Here we consider the effect of higher-order healing for both collective and individual contagion. By *higher-order healing* we refer to a situation where infected nodes that belong to a hyperedge of size $m > 2$ with other infected nodes heal at rate β_m . This can be thought of as a “hipster effect” where if an

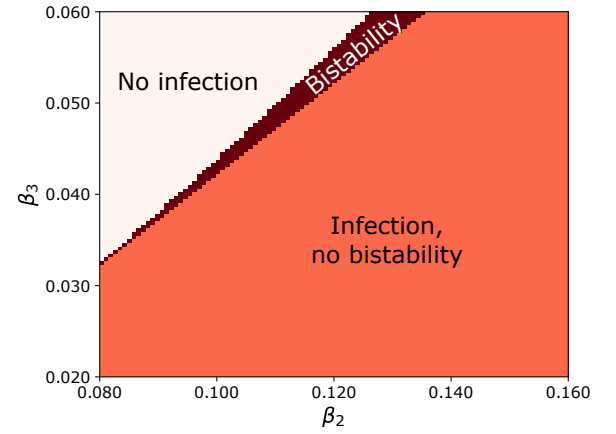


FIG. 7. Phase diagram for the degree-correlated, higher-order healing with individual contagion with parameters $\gamma = 2$ and $P(k) \propto k^{-4}$ when $67 < k < 1000$ and 0 otherwise.

idea or trend is popular in groups, then this makes an individual *less* likely to adopt the trend, but the individual can be convinced to adopt the trend by their pairwise connections³⁷. For both the collective and individual contagion cases, we comment on the existence of bistability based on numerical phase plots.

When the contagion is collective, the model including higher-order healing can be written as Eq. (6) with the sign of the third term changed, and because the triangle healing mechanism is solely higher-order, there is no effect on the epidemic threshold which is obtained by the linearization of the 0 solution. However, we find that explosive transitions do not occur for $\beta_2, \beta_3 \geq 0$.

Likewise, for the individual contagion model, higher-order healing can be written as Eq. (19) with the third term negative. In this case, the epidemic threshold for both the degree-correlated and uncorrelated case can be obtained by substituting $-\beta_3$ for β_3 in Eq. (23) and Eq. (30) respectively. Higher-order healing in individual contagion enables explosive transitions to occur for ranges of $\beta_2, \beta_3 \geq 0$, as can be seen in Fig. 7 which shows the phase space (β_2, β_3) for the degree-correlated case. As one might expect, for large enough higher-order healing β_3 there is no infection, but there is a narrow band of bistable behavior separating the regions of no infection and monostable infection.

D. Unfortunate series of events

So far we have considered hypergraphs with hyperedges of sizes 2 and 3 only. We now briefly discuss contagion in networks with hyperedges of all sizes, i.e. $M = N$. In the context of epidemic spreading, hyperedges could be interpreted as participation in social events such as parties, conferences, concerts, and sports events. For simplicity, we will focus on a

hypergraph with degree-correlated hyperedges where

$$f_m(\mathbf{k}, \mathbf{k}_1, \dots, \mathbf{k}_{m-1}) = \frac{(m-1)! k k_1 k_2 \dots k_{m-1} \langle k^{(m)} \rangle}{(N \langle k \rangle)^{m-1} \langle k \rangle}, \quad (31)$$

such that the average number of hyperedges of size m a node belongs to is $\langle k^{(m)} \rangle$. In this case, by repeating the calculations of Sec. III B, the fraction of infected nodes of degree k evolves in terms of the fraction of infected edges V (7) as

$$\frac{dx_k}{dt} = -\gamma + k(1-x_k) \sum_{m=2}^M \frac{\beta_m \langle k^{(m)} \rangle}{\langle k \rangle} [1 - (1-V)^{m-1}], \quad (32)$$

for individual contagion and

$$\frac{dx_k}{dt} = -\gamma + k(1-x_k) \sum_{m=2}^M \frac{\beta_m \langle k^{(m)} \rangle}{\langle k \rangle} V^{m-1}, \quad (33)$$

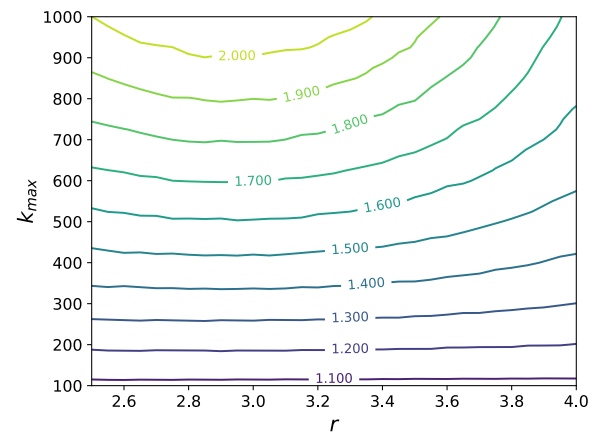
for collective contagion. In the case of collective contagion, larger hyperedges can cause the emergence of new stable fixed points which can lead to richer consensus dynamics²⁴. We focus, however, on the case of individual contagion. Linearizing, we find that the solution $x_k = 0$ becomes unstable when

$$\sum_{m=2}^M \frac{(m-1)\beta_m \langle k^{(m)} \rangle}{\langle k \rangle} > \frac{\gamma \langle k \rangle}{\langle k^2 \rangle}. \quad (34)$$

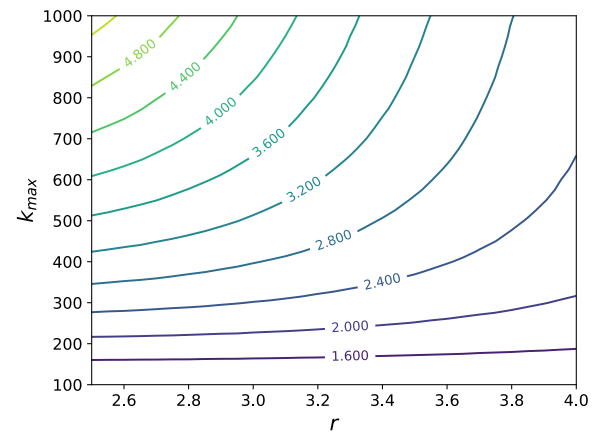
If the sum yields a value larger than $\gamma \langle k \rangle / \langle k^2 \rangle$ propagating social contagion will result. Social event restrictions implemented as a truncation of the series by prohibiting events larger than a certain size, or practices that reduce contagion in social events and reduce β_m (such as enforcing physical separation) can reduce the value of the sum so that contagion does not propagate^{38,39}.

IV. THE EFFECT OF DEGREE DISTRIBUTION ON β_3^c

In Section III we expressed the epidemic threshold β_2^c in terms of moments of the degree distribution of the underlying network structure. Similarly, we would like to express the critical value of β_3 at which the explosive transitions appear, β_3^c , as a function of hypergraph structure. Explosive transitions and bistability occur when there are two stable steady-state solutions to Eqs. (6). For the degree-correlated and uncorrelated cases, this occurs when there are two non-zero solutions to Eq. (10) and the coupled system of Eqs. (15)-(16) respectively. We can compute the critical value of β_3 by finding the numerical solution of these mean field equations and determining the value of β_3^c at which bistability appears. This method is much more efficient than using stochastic microscopic simulations of the contagion model to infer the onset of explosive transitions and to map the phase space. Fig. 8 shows the predicted value of β_3^c normalized by β_2^c for the correlated (a) and uncorrelated (b) cases as a function of the power-law exponent r and the maximum degree k_{max} , where Eqs. (10) and (15)-(16) were solved using $P(k) \propto k^{-r}$ if $50 \leq k \leq k_{max}$ and $P(k) = 0$ otherwise. Larger values of r and k_{max} correspond to



(a) Degree-correlated hyperedges



(b) Uncorrelated hyperedges

FIG. 8. β_3^c / β_2^c as a function of power-law distribution parameters for the degree-correlated case (a) and the uncorrelated case (b). β_3^c was calculated numerically from the mean field equations (see Appendix A 3), and $\beta_2^c = \gamma \langle k \rangle / \langle k^2 \rangle$. The parameters are $P(k) \propto k^{-r}$ if $50 \leq k \leq k_{max}$ and $P(k) = 0$ and $\gamma = 2$.

larger heterogeneity of the degree distribution. We note that for the most homogeneous network – the k -regular network – β_3^c / β_2^c is 1, and we see in Figs. 8(a) and (b) that β_3^c increases relative to β_2^c as r or k_{max} increase, except for small values of r and large values of k_{max} in the degree-correlated case. Thus, heterogeneity in the degree distribution of the pairwise interaction network appears to suppress explosive transitions. However, this effect is much more pronounced for the uncorrelated case (b) than for the degree-correlated case (a), as we discussed previously. In Appendix A 3, we describe in more detail the algorithm employed to find β_3^c from the mean field equations.

Although this method works well in predicting the value of β_3^c , it does not provide a direct relationship between the network structure and the onset of explosive transitions and is more computationally expensive than an analytical expression. For this reason, we present closed form approximations

to β_3^c and describe the parameter regimes over which they are accurate. Starting with the degree-correlated case and canceling the zero solution of Eq. (10), we find conditions under which there are at least two solutions to

$$h(V, \beta_2) = \frac{1}{N\langle k \rangle} \sum_k \frac{kP(k)(\beta_2 k + \beta_3 kV)}{\gamma + \beta_2 kV + \beta_3 kV^2} - 1 = 0. \quad (35)$$

First, note that $h(0, \beta_2) = \beta_2/\beta_2^c - 1$ and that $h(1, \beta_2) < 0$. Therefore, if $\frac{\partial h}{\partial V}(0, \beta_2^c) > 0$, 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

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

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 $h(V, \beta_2)$ near $V = 0$. For the uncorrelated case, however, we find that a saddle-node bifurcation can occur at positive values of V , and it is necessary to expand Eqs. (15)-(16) to higher order.

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

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

which, when evaluated in

$$h(V, \beta_2) = \frac{1}{N\langle k \rangle} \sum_k \frac{kP(k)(\beta_2 kV + \beta_3 \langle k \rangle U^2)}{\gamma + \beta_2 kV + \beta_3 \langle k \rangle U^2} - V = 0 \quad (38)$$

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

$$h(V, \beta_2^c) = (a_0 + a_1 V + a_2 V^2) V^2, \quad (39)$$

where

$$a_0 = -\frac{\langle k \rangle \langle k^3 \rangle}{\langle k^2 \rangle^2} + \frac{\langle k \rangle^5 \beta_3}{\langle k^2 \rangle \gamma}, \quad (40)$$

$$a_1 = \frac{\langle k \rangle^2 \langle k^4 \rangle}{\langle k^2 \rangle^3} - 4 \frac{\langle k \rangle^5 \beta_3}{\langle k^2 \rangle^2 \gamma} + 2 \frac{\langle k \rangle^4 \langle k^3 \rangle \beta_3}{\langle k^2 \rangle^3 \gamma}, \quad (41)$$

$$a_2 = -\frac{\langle k \rangle^3 \langle k^5 \rangle}{\langle k^2 \rangle^4} + 5 \frac{\langle k \rangle^5 \beta_3}{\langle k^2 \rangle^2 \gamma} + 3 \frac{\langle k \rangle^6 \langle k^3 \rangle \beta_3}{\langle k^2 \rangle^4 \gamma} - 6 \frac{\langle k \rangle^4 \langle k^3 \rangle \beta_3}{\langle k^2 \rangle^3 \gamma} + \frac{\langle k \rangle^3 \langle k^3 \rangle^2 \beta_3}{\langle k^2 \rangle^4 \gamma} - \frac{\langle k \rangle^{10} \beta_3^2}{\langle k^2 \rangle^4 \gamma^2}. \quad (42)$$

For continuous transitions to epidemics, there is only one equilibrium for V at $\beta_2 = \beta_2^c$, namely $V = 0$. 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 $V = 0$ in a transcritical bifurcation, or at $V > 0$ as a pair of roots in a saddle-node bifurcation. A pair of roots appears when the discriminant of the

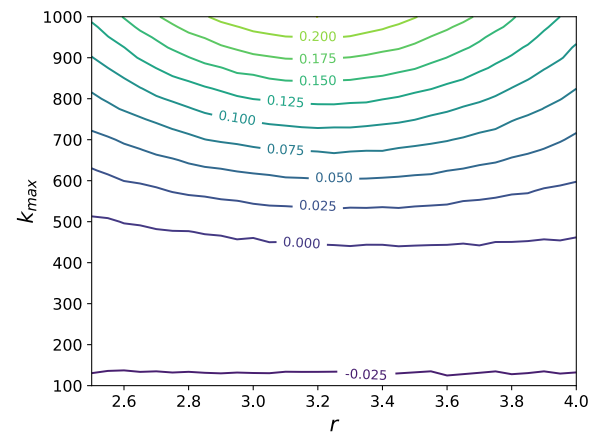


FIG. 9. Relative error in the value of β_3^c/β_2^c obtained from Eq. (43) compared with the numerically obtained value shown in Fig. 8(b).

quadratic equation $a_0 + a_1 V + a_2 V^2 = 0$ is zero. However, this bifurcation is physically meaningless if it occurs for values of V outside the interval $[0, 1]$. Therefore, we impose the constraint that the value of β_3 found by solving $a_1^2 - 4a_0a_2 = 0$ must satisfy the inequality $0 \leq -a_1/2a_2 \leq 1$. In addition, we note that because of continuity, the sign of the a_2 term must be negative, because otherwise $\frac{\partial h}{\partial V}(0, \beta_2^c) > 0$ 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 $a_0 + a_1 V + a_2 V^2 = 0$ is $V = 0$, implying that $a_0 = 0$ and $\beta_3^c = \gamma\langle k^3 \rangle / \langle k \rangle^4$. Using these conditions, we can construct a piecewise definition of β_3^c

$$\beta_3^c = \begin{cases} \text{Solve}(a_1^2 - 4a_0a_2 = 0), & a_2 < 0, 0 \leq -\frac{a_1}{2a_2} \leq 1, \\ \frac{\langle k^3 \rangle}{\langle k \rangle^4} \gamma, & \text{else.} \end{cases} \quad (43)$$

The relative error in the value of β_3^c/β_2^c obtained from Eq. (43) compared with the numerically obtained value shown in Fig. 8(b) is shown in Fig. 9. In principle, one can expand to higher order to gain accuracy for the most heterogeneous of distributions. However, there is limited utility in increasing the order of the expansion further, because the resulting conditions become extremely complicated.

V. DISCUSSION

In this paper we studied the SIS model of social contagion on hypergraphs with heterogeneous structure. The mean field description in Eq. (2) allowed us to explore the effects of hyperedge organization on the epidemic onset and the onset of bistability and explosive transitions. One of our main findings is that with increasing heterogeneity of the pairwise network degree distribution, the onset of explosive transitions is postponed when the pairwise and higher order interactions have independent structure. More generally, when considering a hypergraph contagion model, the group infection and pairwise infection are competing mechanisms by which con-

tagion spreads. Factors that promote contagion via pairwise infection, such as a heterogeneous degree distribution of the pairwise contact network, suppress discontinuous transitions. Conversely, heterogeneity in the degree distribution of hyperedges of higher order promotes such transitions.

We considered two ways in which the structure of hyperedges of different sizes could be organized: the uncorrelated case, in which they are independent, and the correlated case, in which hyperedges of different sizes connect preferentially to the same nodes. While the organization of hyperedges in real world networks is surely much more complicated, these cases can be considered as a null models against which the structure of real-world hypergraphs can be compared.

We studied various forms of higher-order contagion and healing: (i) collective contagion, in which all other members of the hyperedge need to be infected for contagion to occur, (ii) individual contagion, in which at least one member of the hyperdegree needs to be infected, and (iii) higher-order healing, in which pairwise interactions are infectious while higher-order interactions heal. Other forms of higher-order contagion could in principle be studied with the same methodology, but we leave these studies for future research.

Now we mention some of the limitations of our study. First, since we focused on the simplest contagion model, an important question left for future research is whether our results remain valid for more realistic epidemiological models (e.g. such as those used to model COVID-19^{14,40}). Our model also does not apply to non-Markovian contagion dynamics, which are important when modeling real-world epidemics. From a technical standpoint, another limitation is that we used a mean-field description of the dynamics, and it is known that such a description is not quantitatively accurate for moderate values of the infected population value^{28,41}. Since we were mainly interested in the behavior close to the onset of epidemics, the mean-field approximation was enough for our purposes. However, more precise descriptions could be obtained as in Refs.^{28,42}. Another important limitation of our hypergraph model is that we assume that the probability that two nodes belong to the same hyperedge is a function of their hyperdegrees. While this assumption can be relaxed by considering additional nodal variables, it is possible that such a model might be inadequate to describe some real-world networks. Finally we note that our model relies on knowledge of the functions f_m , which encode the organization of hyperedges across different hyperedge sizes. These functions have not yet been estimated from real-world networks, but as progress is made towards understanding the organization of higher-order interactions²³, the determination of these functions could be a natural next step.

While in this paper we applied our hyperdegree based mean-field equation to the SIS epidemic model, the same formalism could be applied to other dynamical processes on hypergraphs, such as synchronization, opinion formation, and other types of epidemic models. We believe that this method-

ology will be useful to study the effect of heterogeneity on these hypergraph dynamical processes.

VI. DATA AVAILABILITY STATEMENT

The data that support the findings of this study are openly available in <https://github.com/nwlandry/SimplexSIS>.

ACKNOWLEDGMENTS

Nicholas Landry acknowledges a helpful conversation with Danny Abrams and Lluís Arola-Fernández on the “hipster effect”.

Appendix A: Numerical Experiments

1. Microscopic simulation of the hypergraph SIS model

We simulated the stochastic SIS model on a given hypergraph as a discrete-time Markov process on the nodes with transitions to infected and healthy states through the modalities described in Section II B, a variant of which was simulated in Ref.⁴⁴. We denote the binary states of the nodes at a time step t by a vector $\mathbf{X}^t = (X_1^t, X_2^t, \dots, X_N^t)$ where $X_i^t = 0$ if node i is healthy and $X_i^t = 1$ if it is infected. In this model, we assume that the events that a hyperedge infects node i and that a pairwise connection infects node i are independent. Likewise, we assume that an infected neighbor, whether through a pairwise or group connection, infects a node independently of any other neighboring node. The probability that a single infected node infects its pairwise neighbor in the time interval $[t, t + \Delta t]$ is $\beta_2 \Delta t$, and so the probability that no neighboring node infects a given node is

$$(1 - \beta_2 \Delta t)^{(A\mathbf{X})_i}$$

where A is the adjacency matrix with entries $A_{ij} = 1$ if nodes i and j are connected by a link and 0 otherwise.

In the collective contagion model, the probability that a triangle infects a node in the time interval $[t, t + \Delta t]$ is $\beta_3 \Delta t$ provided the other two nodes are infected. Therefore, the probability of no triangles infecting node i can be written as

$$(1 - \beta_3 \Delta t)^{\sum_{\{i_1, i_2, i\}} X_{i_1}^t X_{i_2}^t},$$

where the sum is over all triangles $\{i_1, i_2, i\}$ with node i as a member. Lastly, the rate of healing is constant and independent of the infection status of any neighboring nodes so the probability that an infected node heals in a time interval $[t, t + \Delta t]$ is $\gamma \Delta t$.

The Markov process can then be described as

$$P(X_i^{t+1} = 1 \mid X_i^t = 0) = 1 - (1 - \beta_2 \Delta t)^{(AX^t)_i} (1 - \beta_3 \Delta t)^{\sum_{\{i_1, i_2, i\}} X_{i_1}^t X_{i_2}^t}, \quad (A1)$$

$$P(X_i^{t+1} = 0 \mid X_i^t = 1) = \gamma \Delta t. \quad (A2)$$

In our simulations, we updated the status of the nodes synchronously at times $t = 0, \Delta t, 2\Delta t, \dots, n\Delta t$ where $\Delta t = 0.1$.

Our specific implementation is described in what follows. We note that for all mechanisms of infection and healing described next, $u_i \sim \text{Uniform}(0, 1)$ and this variable is drawn independently for each modality and each node i . At each time step, we iterate through every node and follow the following conditional logic. If a node i is already infected, it is healed if $u_i < \gamma \Delta t$ and remains infected otherwise. Next, if the node i is currently healthy, it is infected by its pairwise neighbors if $u_i < 1 - (1 - \beta_2 \Delta t)^{(AX^t)_i}$ and remains healthy otherwise. If node i still remains healthy after being subjected to pairwise infection, the node is infected by its triangle neighbors if

$$u_i < 1 - (1 - \beta_3 \Delta t)^{\sum_{\{i_1, i_2, i\}} X_{i_1}^t X_{i_2}^t}$$

and remains healthy otherwise. Note that each infection mechanism is only dependent on the prior time step so the order of these steps does not matter.

At $t = 0$, 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, $x^t = \sum_{i=1}^N X_i^t / N$ is stored. To avoid the absorbing state $\mathbf{X}^t = \mathbf{0}$, 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 x^t 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 $\{\gamma, \beta_2, \beta_3\}$ for 1000 time steps and averaged over the last 300 time steps.

To find the bistability index, we initialize the simulation with a small fraction of infected agents for a fixed β_3 value and incrementally increase β_2 from a sufficiently small value (typically $\beta_2^c/2$) to a value above the critical value of β_2 (typically $3\beta_2^c/2$), and then incrementally decrease the value of β_2 down to its original value. As described previously, if the equilibrium value while increasing the value of β_2 is distinct from the equilibrium value while decreasing the value of β_2 for the same β_2 values, this indicates the presence of bistability. We simulated several equilibrium curves corresponding to different β_3 values to observe the value of β_3 at which the response curve starts to show bistability, and thus infer the value of β_3^c .

2. Network models

We exclusively considered networks generated using the configuration model in order to isolate the effect of the degree

distribution. Although the configuration model has the potential to contain both self-loops and multi-edges, in practice, the fraction of these types of edges is small⁴¹ and in our numerical experiments, the number of self-loops was approximately 1% of the total number of nodes.

We used networks of size $N = 10^4$ in the simulation of the hypergraph SIS model because this was sufficiently large enough to reduce the finite-size effects. Because the network realization was relatively large, we did not average over an ensemble of these random graphs as in Ref.²⁴. We have described in Section III A the particular distributions examined.

We generated the triangles in two different ways corresponding to the two separate cases; degree-correlated and uncorrelated. For the first case, we used the same degree sequence as used to generate the network using the configuration model and extended the configuration model to triangles as has been done in prior work^{32,33}. Because this is analogous to the construction of the network configuration model, there is also the possibility for self-loops and multi-edges, but this probability is low. For the independently-distributed triangles, we drew with replacement a fixed number of triples (enforcing the mean triangle degree) containing node indices and assigned these nodes to a triangle. Again, as with the standard configuration model, there is the possibility for self-loops and multi-edges, but the probability of either occurring is small.

3. The numerical computation of β_3^c

In Section IV, we plotted the numerical solution of β_3^c for truncated power law distributions as a function of the maximum degree and power-law exponent. In this section, we discuss the specific methodology in generating these results.

First, we describe the process for finding the bistability index accurately from the mean-field equations (10) and (15)-(16) for the correlated and uncorrelated cases respectively. Since the $V = 0$ solution becomes linearly unstable at $\beta_2 = \beta_2^c$ and the stable $V > 0$ solution is monotonically increasing with β_2 , the bistability index $B(\beta_3)$ coincides with the value of the largest root of Eq. (10) for the correlated case [or Eqs. (15)-(16) for the uncorrelated case] at $\beta_2 = \beta_2^c$, as shown schematically in Fig. 10. Therefore, using our analytical knowledge of β_2^c , we set $B(\beta_3) \approx V_\varepsilon^*$, where V_ε^* is the largest root at $\beta_2 = \beta_2^c - \varepsilon$ with $\varepsilon = 10^{-5}$ a small number added for numerical robustness. (We verified that this method gives numerically accurate results when compared with other methods which do not require knowledge of β_2^c but are more computationally intensive.)

Being able to compute $B(\beta_3)$, we find $\beta_3^c = \sup\{\beta_3 \mid B(\beta_3) = 0\}$ by bisection: starting with an in-

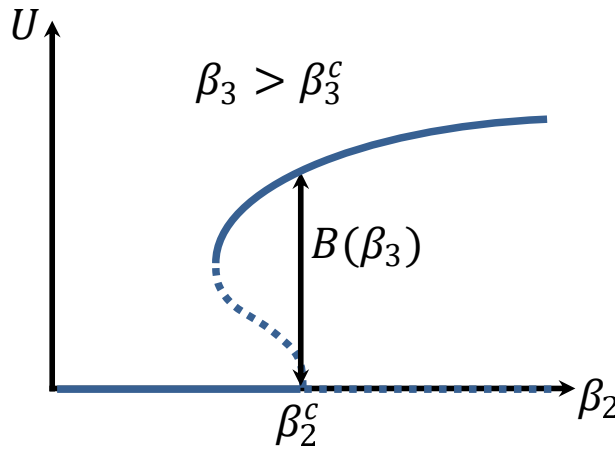


FIG. 10. Illustration of the bistability index with respect to the solutions to the mean-field equation in the bistable regime

interval $[\beta_3^{\min,0}, \beta_3^{\max,0}]$ such that $B(\beta_3^{\min,0}) = 0$, $B(\beta_3^{\max,0}) > 0$, we recursively define the interval $[\beta_3^{\min,i+1}, \beta_3^{\max,i+1}]$ as $[\beta_3^{\min,i}, \tilde{\beta}^i]$ if $B(\tilde{\beta}^i) > 0$ and $[\tilde{\beta}^i, \beta_3^{\max,i}]$ if $B(\tilde{\beta}^i) = 0$, where $\tilde{\beta}^i = (\beta_3^{\min,i} + \beta_3^{\max,i})/2$. When the length of the interval $[\beta_3^{\min,i}, \beta_3^{\max,i}]$ is less than the tolerance 10^{-4} , we set $\beta_3^c = \beta_3^{\min,i}$.

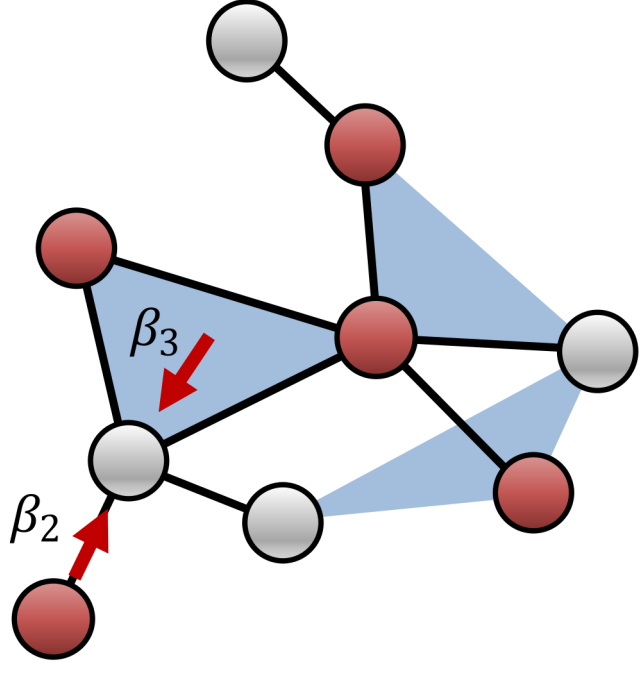
- ¹R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks," *Reviews of modern physics* **87**, 925 (2015).
- ²P. Trapman, "On analytical approaches to epidemics on networks," *Theoretical population biology* **71**, 160–173 (2007).
- ³I. Z. Kiss, J. C. Miller, P. L. Simon, *et al.*, "Mathematics of epidemics on networks," Cham: Springer **598** (2017).
- ⁴T. House, "Modelling epidemics on networks," *Contemporary Physics* **53**, 213–225 (2012).
- ⁵M. E. Newman, "Spread of epidemic disease on networks," *Physical review E* **66**, 016128 (2002).
- ⁶R. Pastor-Satorras and A. Vespignani, "Epidemic dynamics in finite size scale-free networks," *Physical Review E* **65**, 035108 (2002).
- ⁷Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, "Epidemic spreading in real networks: An eigenvalue viewpoint," in *22nd International Symposium on Reliable Distributed Systems, 2003. Proceedings.* (IEEE, 2003) pp. 25–34.
- ⁸F. Xiong and Y. Liu, "Opinion formation on social media: an empirical approach," *Chaos: An Interdisciplinary Journal of Nonlinear Science* **24**, 013130 (2014).
- ⁹D. J. Watts and P. S. Dodds, "Influentials, networks, and public opinion formation," *Journal of consumer research* **34**, 441–458 (2007).
- ¹⁰R. Cowan and N. Jonard, "Network structure and the diffusion of knowledge," *Journal of economic Dynamics and Control* **28**, 1557–1575 (2004).
- ¹¹T. W. Valente, *Network models of the diffusion of innovations*, 303.484 V3 (Hampton Press, 1995).
- ¹²M. Gladwell, *The tipping point: How little things can make a big difference* (Little, Brown, 2006).
- ¹³D. Centola, J. Becker, D. Brackbill, and A. Baronchelli, "Experimental evidence for tipping points in social convention," *Science* **360**, 1116–1119 (2018).
- ¹⁴A. Arenas, W. Cota, J. Gomez-Gardenes, S. Gómez, C. Granell, J. T. Matamalas, D. Soriano-Panos, and B. Steinegger, "A mathematical model for the spatiotemporal epidemic spreading of COVID19," *MedRxiv* (2020).
- ¹⁵B. Banerjee, P. K. Pandey, and B. Adhikari, "A model for the spread of an epidemic from local to global: A case study of COVID-19 in India," *arXiv preprint arXiv:2006.06404* (2020).

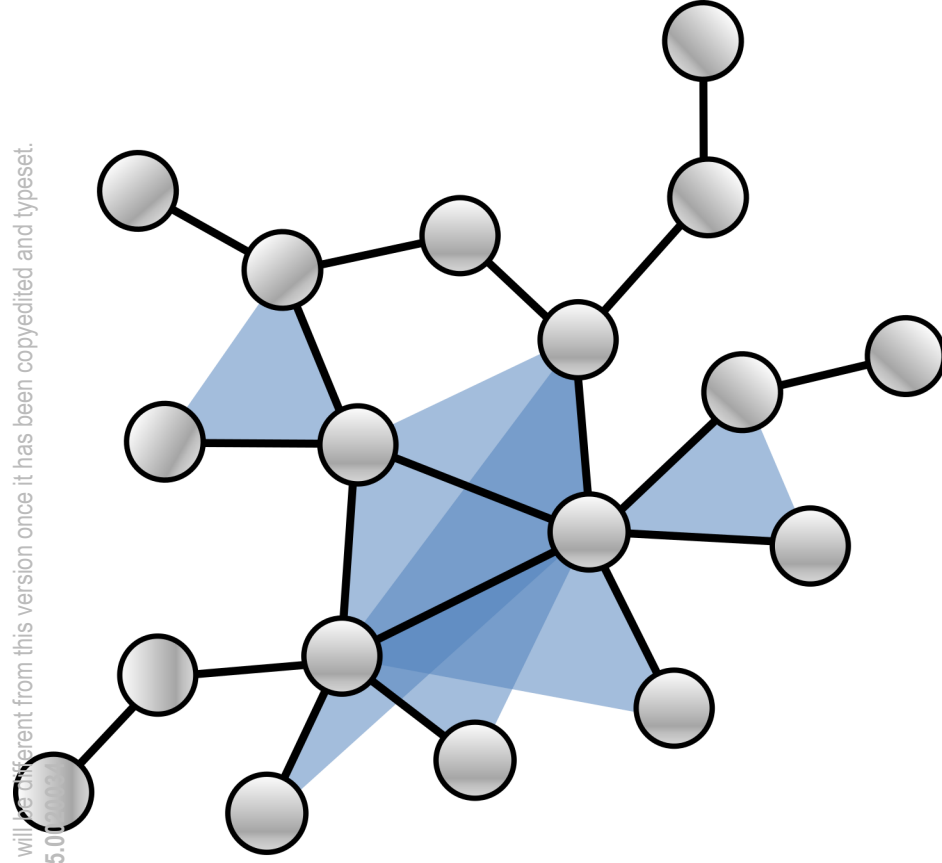
- ¹⁶F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, "Networks beyond pairwise interactions: structure and dynamics," *Physics Reports* (2020).
- ¹⁷M. A. Porter, "Nonlinearity+ networks: A 2020 vision," *arXiv preprint arXiv:1911.03805* (2019).
- ¹⁸L. Horstmeier and C. Kuehn, "Adaptive voter model on simplicial complexes," *Physical Review E* **101**, 022305 (2020).
- ¹⁹P. S. Skardal and A. Arenas, "Abrupt desynchronization and extensive multistability in globally coupled oscillator simplexes," *Physical review letters* **122**, 248301 (2019).
- ²⁰C. Xu, X. Wang, and P. S. Skardal, "Bifurcation analysis and structural stability of simplicial oscillator populations," *Physical Review Research* **2**, 023281 (2020).
- ²¹A. P. Millán, J. J. Torres, and G. Bianconi, "Explosive higher-order Kuramoto dynamics on simplicial complexes," *Physical Review Letters* **124**, 218301 (2020).
- ²²J. Grilli, G. Barabás, M. J. Michalska-Smith, and S. Allesina, "Higher-order interactions stabilize dynamics in competitive network models," *Nature* **548**, 210–213 (2017).
- ²³A. R. Benson, R. Abebe, M. T. Schaub, A. Jadbabaie, and J. Kleinberg, "Simplicial closure and higher-order link prediction," *Proceedings of the National Academy of Sciences* **115**, E11221–E11230 (2018).
- ²⁴I. Iacopini, G. Petri, A. Barrat, and V. Latora, "Simplicial models of social contagion," *Nature communications* **10**, 1–9 (2019).
- ²⁵G. F. de Arruda, M. Tizzani, and Y. Moreno, "Phase transitions and stability of dynamical processes on hypergraphs," *arXiv preprint arXiv:2005.10891* (2020).
- ²⁶F. B. Pedro Cisneros-Velarde, "Multi-group SIS epidemics with simplicial and higher-order interactions," *arXiv preprint arXiv:2005.11404v1* (2020).
- ²⁷G. F. de Arruda, G. Petri, and Y. Moreno, "Social contagion models on hypergraphs," *Physical Review Research* **2**, 023032 (2020).
- ²⁸J. T. Matamalas, S. Gómez, and A. Arenas, "Abrupt phase transition of epidemic spreading in simplicial complexes," *Physical Review Research* **2**, 012049 (2020).
- ²⁹B. Jhun, M. Jo, and B. Kahng, "Simplicial SIS model in scale-free uniform hypergraph," *Journal of Statistical Mechanics: Theory and Experiment* **2019**, 123207 (2019).
- ³⁰M. Boguná and R. Pastor-Satorras, "Epidemic spreading in correlated complex networks," *Physical Review E* **66**, 047104 (2002).
- ³¹J. C. Miller, A. C. Slim, and E. M. Volz, "Edge-based compartmental modelling for infectious disease spread," *Journal of the Royal Society Interface* **9**, 890–906 (2012).
- ³²O. T. Courtney and G. Bianconi, "Generalized network structures: The configuration model and the canonical ensemble of simplicial complexes," *Physical Review E* **93**, 062311 (2016).
- ³³J.-G. Young, G. Petri, F. Vaccarino, and A. Patania, "Construction of and efficient sampling from the simplicial configuration model," *Physical Review E* **96**, 032312 (2017).
- ³⁴A. J. Ward, D. J. Sumpter, I. D. Couzin, P. J. Hart, and J. Krause, "Quorum decision-making facilitates information transfer in fish shoals," *Proceedings of the National Academy of Sciences* **105**, 6948–6953 (2008).
- ³⁵S. C. Pratt, E. B. Mallon, D. J. Sumpter, and N. R. Franks, "Quorum sensing, recruitment, and collective decision-making during colony emigration by the ant *Leptothorax albigenis*," *Behavioral Ecology and Sociobiology* **52**, 117–127 (2002).
- ³⁶B. Guerra and J. Gómez-Gardenes, "Annealed and mean-field formulations of disease dynamics on static and adaptive networks," *Physical Review E* **82**, 035101 (2010).
- ³⁷J. D. Touboul, "The hipster effect: When anti-conformists all look the same," *Discrete & Continuous Dynamical Systems-B* **24**, 4379–4415 (2019).
- ³⁸G. St-Onge, V. Thibault, A. Allard, L. J. Dubé, and L. Hébert-Dufresne, "School closures, event cancellations, and the mesoscopic localization of epidemics in networks with higher-order structure," *arXiv preprint arXiv:2003.05924* (2020).
- ³⁹B. M. Althouse, E. A. Wenger, J. C. Miller, S. V. Scarpino, A. Allard, L. Hébert-Dufresne, and H. Hu, "Stochasticity and heterogeneity in the transmission dynamics of sars-cov-2," *arXiv preprint arXiv:2005.13689* (2020).
- ⁴⁰D. Balcan, B. Gonçalves, H. Hu, J. J. Ramasco, V. Colizza, and A. Vespig-

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0020034

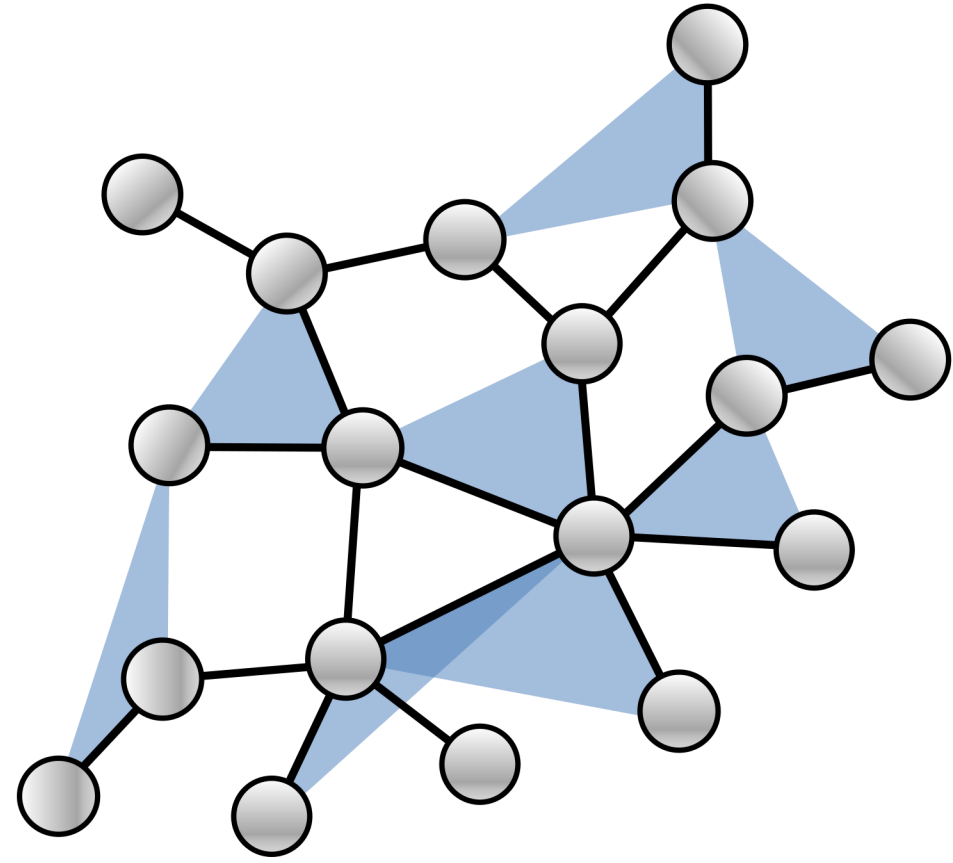
- nani, "Modeling the spatial spread of infectious diseases: The GLobal Epidemic and Mobility computational model," *Journal of computational science* **1**, 132–145 (2010).
- ⁴¹M. Newman, *Networks* (Oxford university press, 2018).
- ⁴²R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics* **74**, 47 (2002).
- ⁴³G. St-Onge, V. Thibeault, A. Allard, L. J. Dubé, and L. Hébert-Dufresne, "Social confinement and mesoscopic localization of epidemics on networks," arXiv preprint arXiv:2003.05924v2 (2020).
- ⁴⁴Á. Bodó, G. Y. Katona, and P. L. Simon, "SIS epidemic propagation on hypergraphs," *Bulletin of mathematical biology* **78**, 713–735 (2016).

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0020034



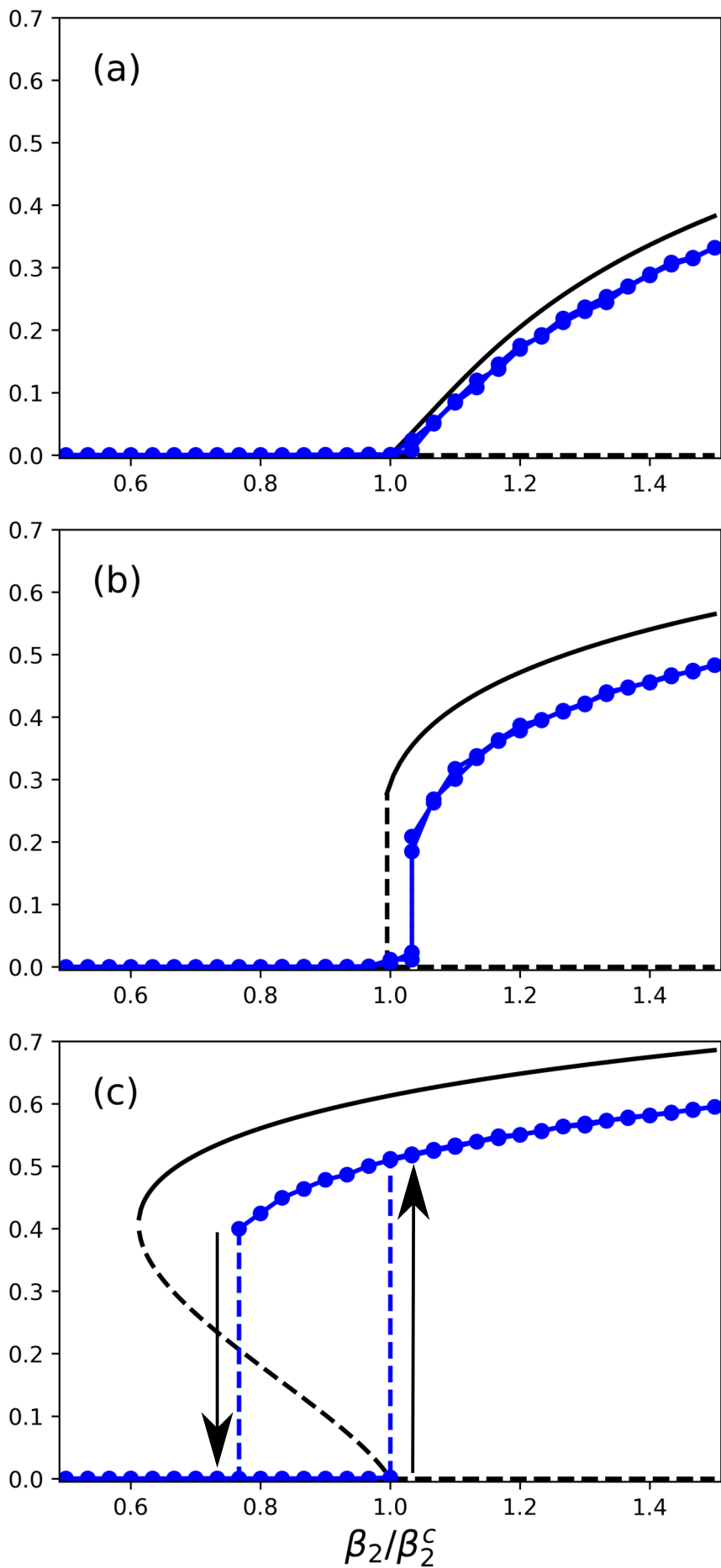


Degree-correlated



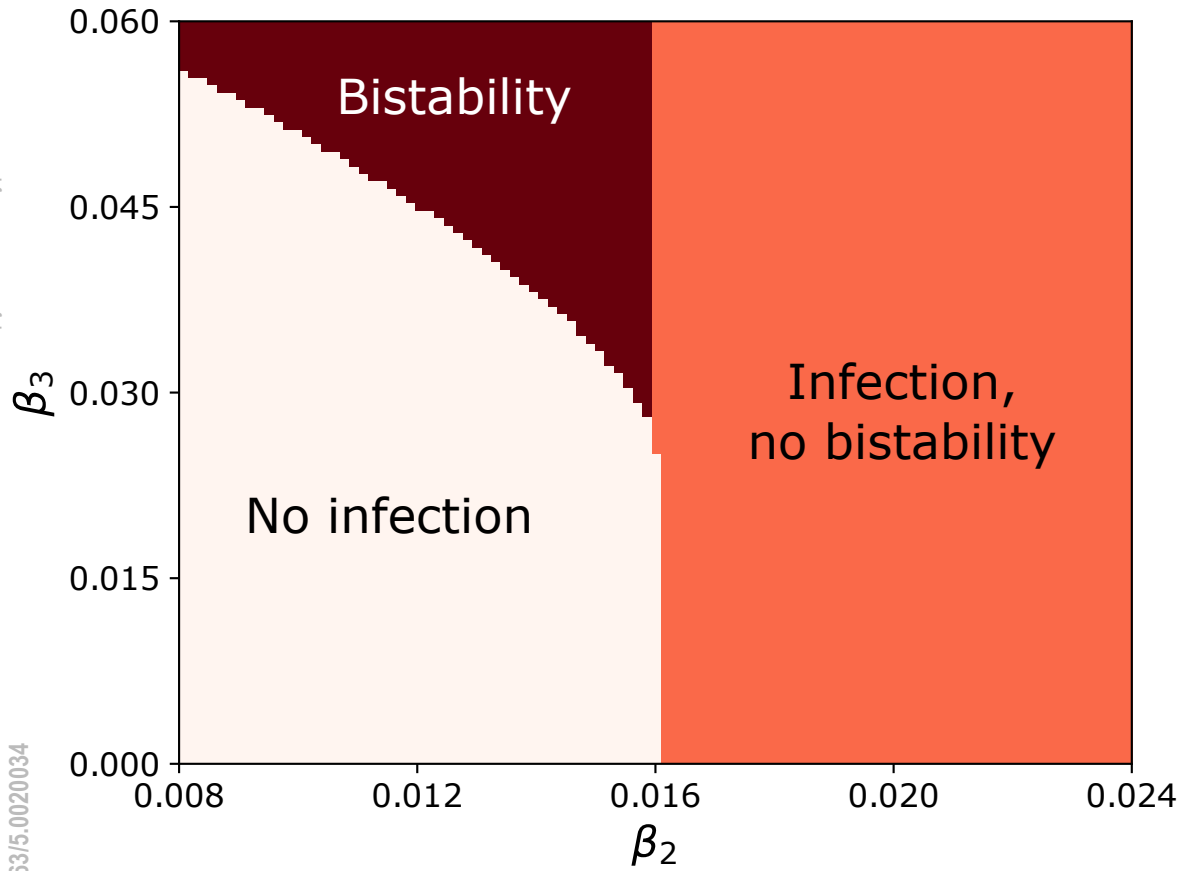
Uncorrelated

This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0020034

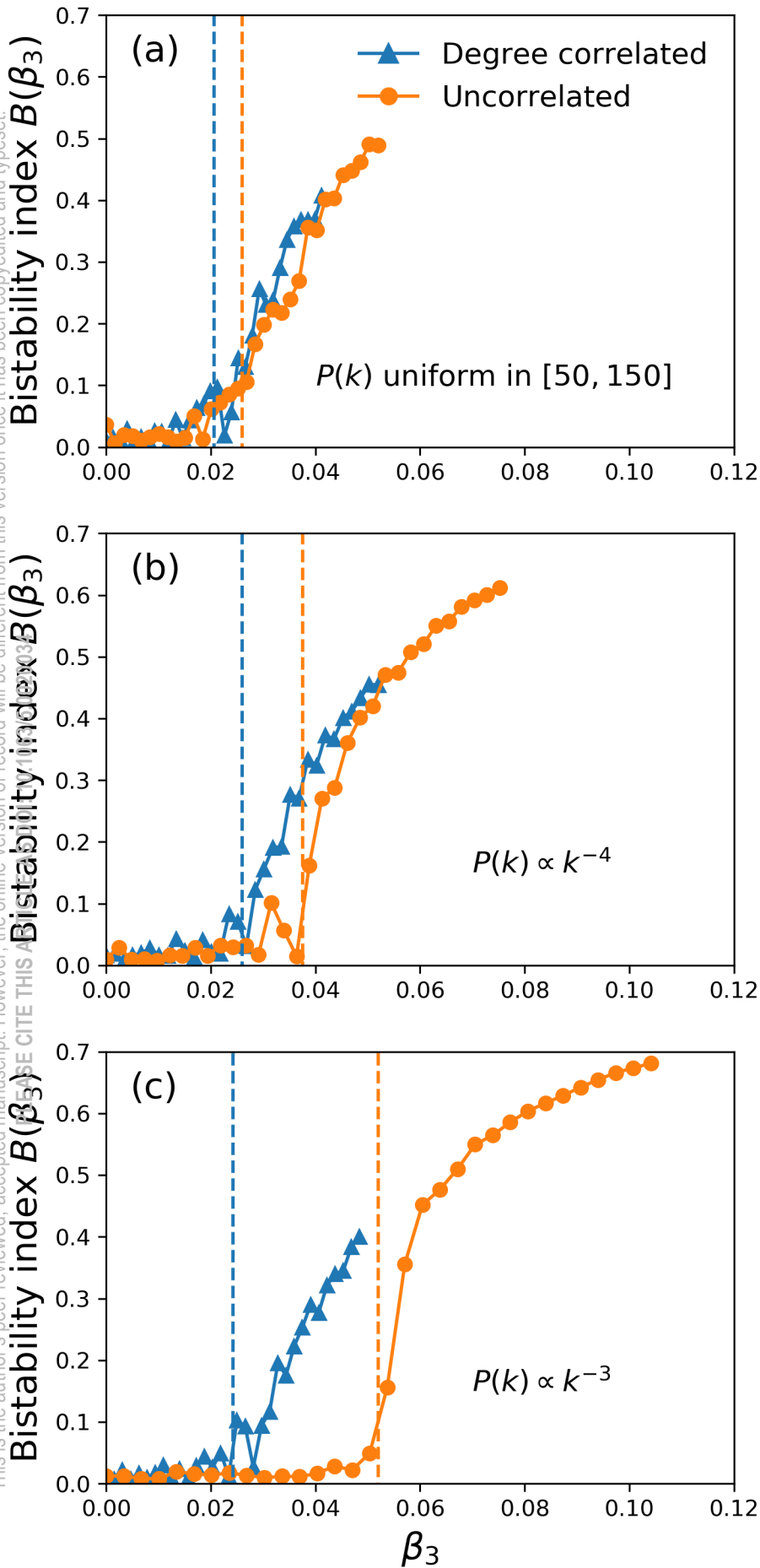


This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0020034

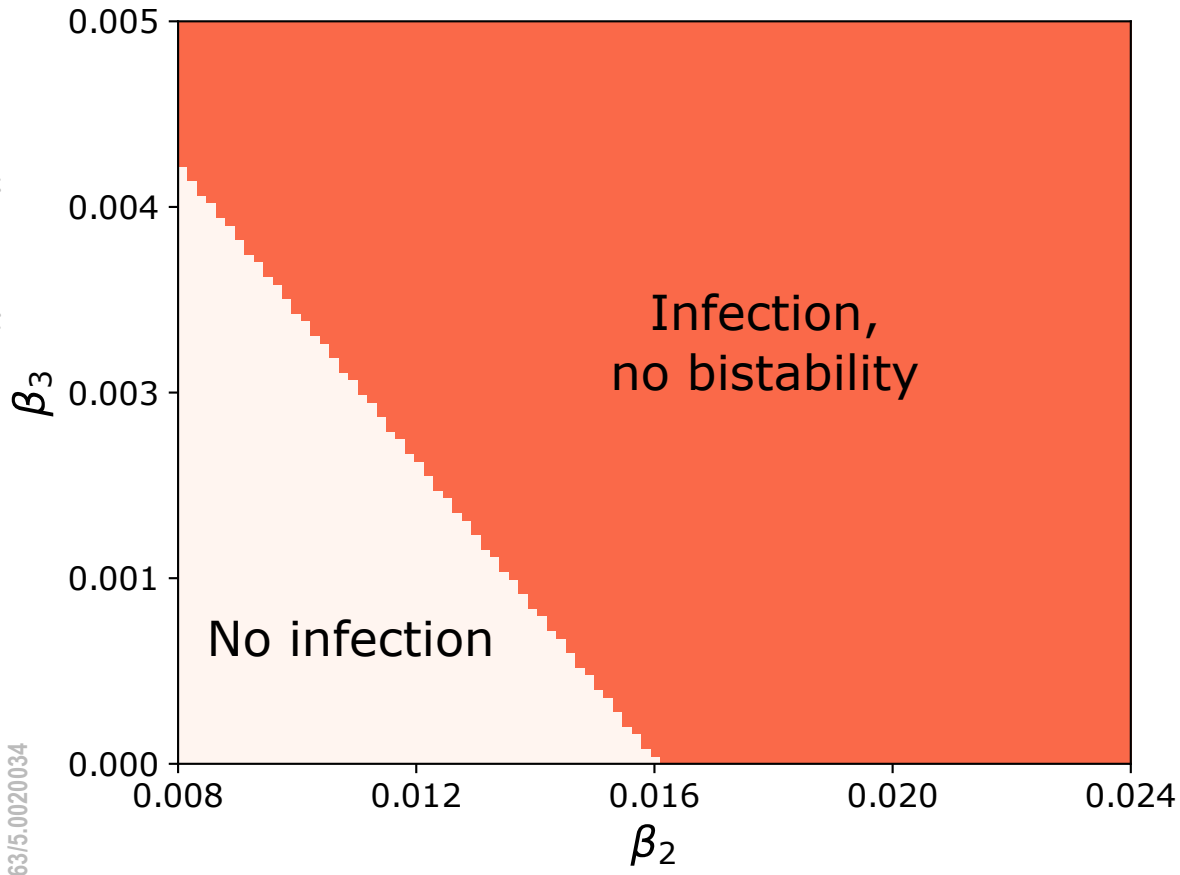


This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.



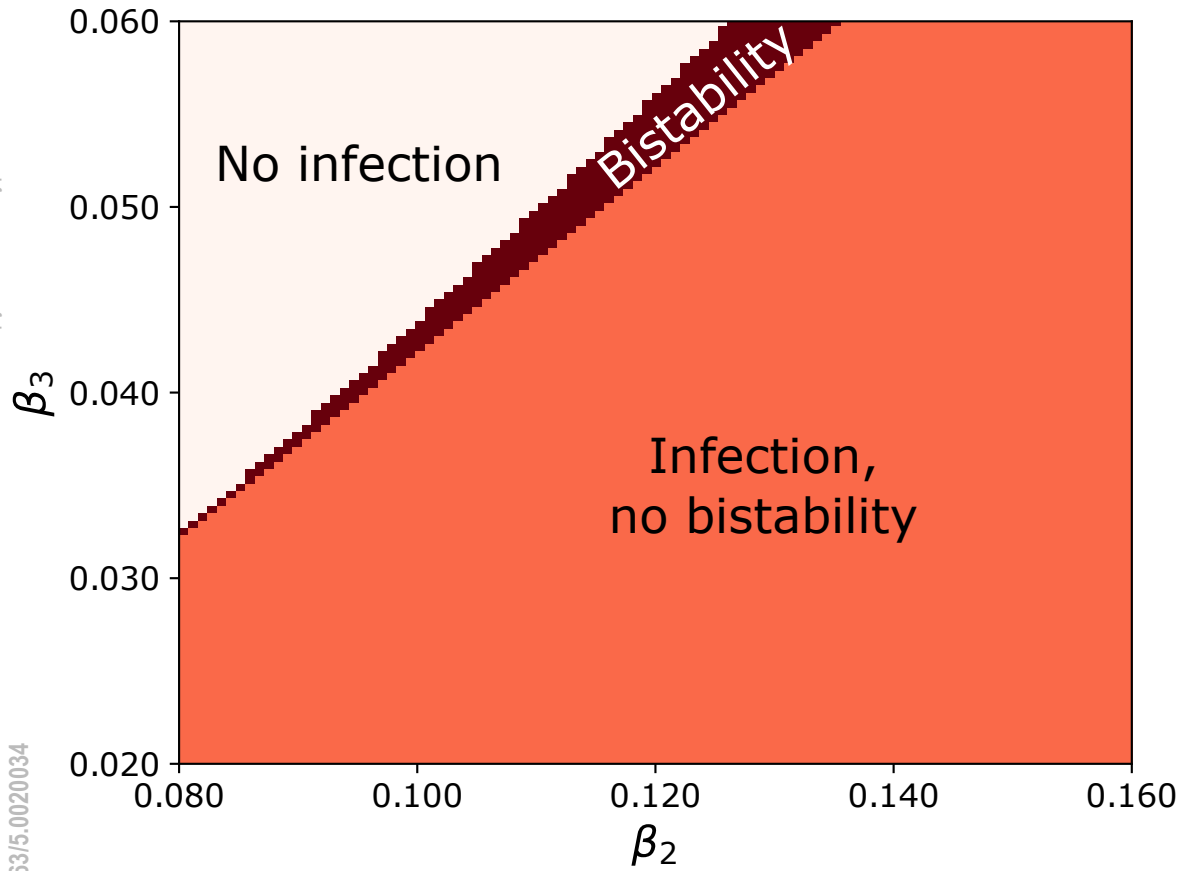
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0020034

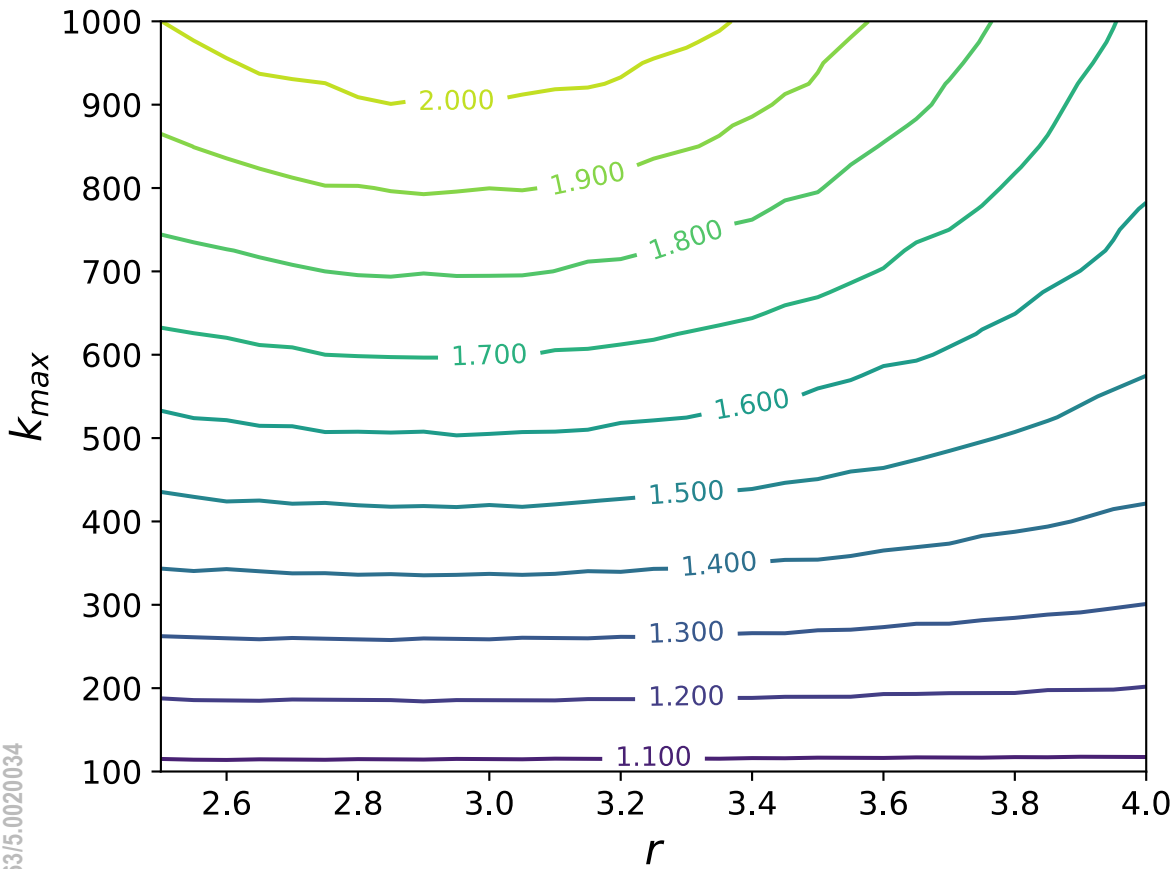


This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.

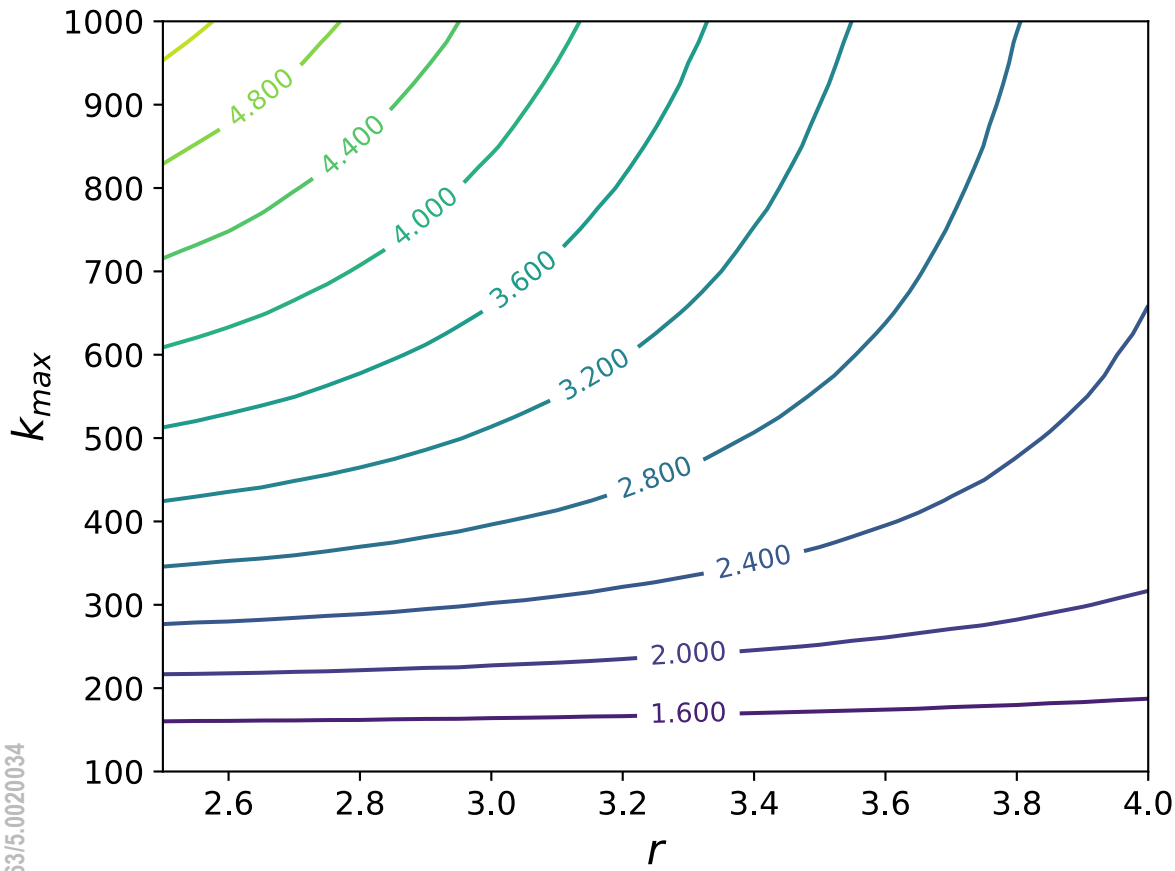
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0020034



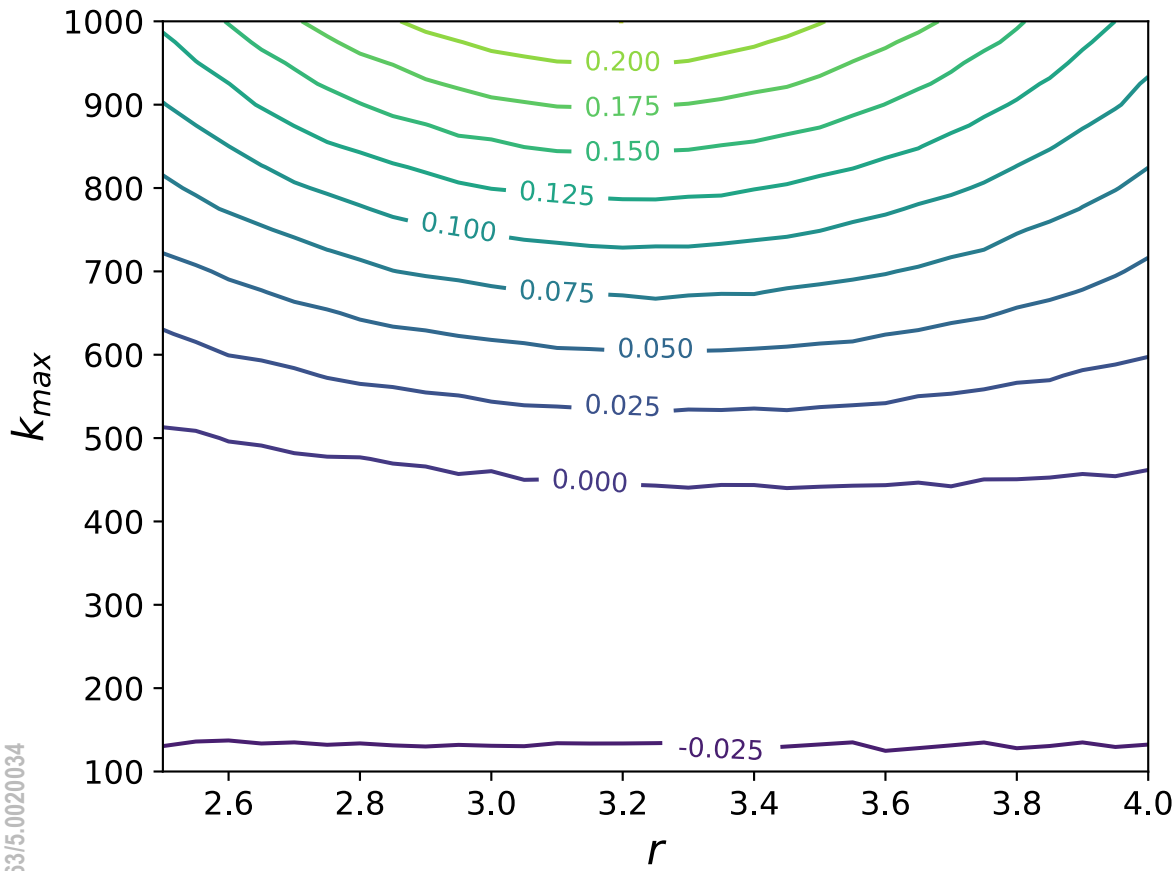
This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0020034



This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0020034



This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0020034



This is the author's peer reviewed, accepted manuscript. However, the online version of record will be different from this version once it has been copyedited and typeset.
PLEASE CITE THIS ARTICLE AS DOI: 10.1063/5.0020034

