

Opinion disparity in hypergraphs with community structure

Nicholas W. Landry ^{1,2,3,*} and Juan G. Restrepo ^{3,†}

¹*Vermont Complex Systems Center, University of Vermont, Burlington, Vermont 05405, USA*

²*Department of Mathematics and Statistics, University of Vermont, Burlington, Vermont 05405, USA*

³*Department of Applied Mathematics, University of Colorado at Boulder, Boulder, Colorado 80309, USA*



(Received 28 February 2023; accepted 11 September 2023; published 27 September 2023)

The division of a social group into subgroups with opposing opinions, which we refer to as *opinion disparity*, is a prevalent phenomenon in society. This phenomenon has been modeled by including mechanisms such as opinion homophily, bounded confidence interactions, and social reinforcement mechanisms. In this paper, we study a complementary mechanism for the formation of opinion disparity based on higher-order interactions, i.e., simultaneous interactions between multiple agents. We present an extension of the planted partition model for uniform hypergraphs as a simple model of community structure, and we consider the hypergraph Susceptible-Infected-Susceptible (SIS) model on a hypergraph with two communities where the binary ideology can spread via links (pairwise interactions) and triangles (three-way interactions). We approximate this contagion process with a mean-field model and find that for strong enough community structure, the two communities can hold very different average opinions. We determine the regimes of structural and infectious parameters for which this opinion disparity can exist, and we find that the existence of these disparities is much more sensitive to the triangle community structure than to the link community structure. We show that the existence and type of opinion disparities are extremely sensitive to differences in the sizes of the two communities.

DOI: [10.1103/PhysRevE.108.034311](https://doi.org/10.1103/PhysRevE.108.034311)

I. INTRODUCTION

Modeling the spreading and dynamics of opinions on social networks is a problem of major interest in modern network science. Many classical models of opinion dynamics, such as the voter model [1], the majority rule model [2,3], and the Sznajd model [4], inevitably result in consensus [5], i.e., the same opinion shared by all agents. In reality, however, consensus is rarely reached; instead, strong differences in opinion between different groups, i.e., *polarization*, are commonly observed. This is evident in politics [6], social media [7], and ideology [8]. Many opinion models giving rise to polarization are based on opinion homophily, i.e., the tendency of individuals to associate with similar-minded individuals [9–13], bounded confidence interactions [14–16], unfollowing on social media platforms [17], and social reinforcement and feedback mechanisms [18,19]. Other models for the formation of polarization include the effects of cooperation and partisanship [20,21], echo chambers on social media [7], radicalization dynamics [22], cognitive biases [23], moderates seen as outsiders [24], and media influence [25].

A closely related phenomenon to polarization occurs when social groups have different average opinions, which we refer to as *opinion disparity* to avoid confusion with the extensive literature and existing definitions of polarization. In this paper, we explore a mechanism for the formation of opinion disparity driven by the presence of simultaneous interactions

between multiple agents, known as *higher-order interactions*. Reference [26] noted that higher-order interactions in the Susceptible-Infected-Susceptible (SIS) model can lead to bistability, hysteresis, and explosive transitions in random simplicial complexes, and Ref. [27] extended this result to more complex hypergraphs. Since higher-order interactions lead to bistability in all-to-all and random hypergraphs, one expects that hypergraphs with strong community structure retain—to some extent—the ability to sustain different opinions in different communities. Confirming this, Ref. [28] found that hypergraph community structure combined with higher-order interactions can result in multistability, intermittency, and complex dynamics for a generalized SIS model on hypergraphs. Reference [29] studied the majority rule model on hypergraphs with community structure and found that for sufficiently disconnected communities, the average states of each community are well-separated. Extending these previous results, here we explore the interplay of community structure and higher-order interactions systematically by first developing a generative model for hypergraphs with tunable community structure, and then studying a social contagion model on hypergraphs with community structure to determine which combinations of hypergraph and contagion parameters support communities with different opinions. By analyzing a mean-field model of this social contagion process, we find that opinion disparity is possible for strong enough community structure in the higher-order interactions, and it can appear suddenly as the strength of the community structure is increased. Furthermore, we find that the presence of opinion disparity can be extremely sensitive to the relative size of the communities.

*nicholas.landry@uvm.edu

†juanga@colorado.edu

The structure of our paper is as follows: we define the contagion model and generative hypergraph models that we use in Sec. II, we use these definitions in Sec. III to develop and analyze mean-field models of contagion spread on hypergraphs with community structure, and we discuss our results in Sec. IV.

II. MODELS

In this section, we present both a generative hypergraph model that incorporates community structure as well as the opinion model that we use in the rest of the paper.

A. Notation

A hypergraph, $H = (V, E)$, is a mathematical object that encodes relationships between arbitrary numbers of entities, where V is the set of $N = |V|$ vertices and E is the set of hyperedges, where a hyperedge $e \in E$ is a subset of the vertices. An m -uniform hypergraph is a hypergraph where $|e| = m$ for every hyperedge e in the set of hyperedges E . The m th-order degree of a node i , $k_i^{(m)}$, is the number of m -hyperedges of which node i is a member. In the following, we will use $k = k^{(2)}$ to signify the second-order degree and $q = k^{(3)}$ to signify the third-order degree. We will indicate their mean values by $\langle k \rangle$ and $\langle q \rangle$, respectively. We will also refer to hyperedges of sizes 2 as *links* and to hyperedges of size 3 as *triangles*. When an m -uniform hypergraph is specified and there is no possibility of confusion, we will denote $k = k^{(m)}$ to simplify notation.

B. The stochastic block model for uniform hypergraphs

The stochastic block model (SBM) is a simple random network model incorporating community structure [31]. Given a network with G communities, each node i is assigned a community label $g_i \in \{1, 2, \dots, G\}$. The probability that two nodes i and j are connected with an edge is assumed to depend only on their community labels g_i and g_j . Extending this notion to m -uniform hypergraphs gives rise to the m -uniform hypergraph stochastic block model (m-HSBM) [32,33]. Given an m -uniform hypergraph, suppose that each node i has a community label g_i and that the number of communities is given by G . Then, given nodes i_1, \dots, i_m with community labels g_{i_1}, \dots, g_{i_m} , we define the probability that a hyperedge connects these nodes as $P_{g_{i_1}, \dots, g_{i_m}}$. Because every vertex order in a hyperedge is equivalent, P is symmetric, i.e., $P_{i_1, \dots, i_m} = P_{\sigma_{i_1}, \dots, \sigma_{i_m}}$, where σ is any permutation of the indices.

In our numerical experiments, we will generate hypergraphs from a version of this model. Sampling from this model by iterating through every potential edge and accepting it with a specified probability is inefficient for sparse hypergraphs. In Appendix A, we present a more efficient algorithm for sampling sparse hypergraphs from this model.

C. Planted partition model for uniform hypergraphs

Extending the *planted partition model* for pairwise networks [34] and the hypergraph stochastic block model [32,33,35,36], here we introduce the *m-uniform Hypergraph Planted Partition Model* (HPPM). Using a single parameter,

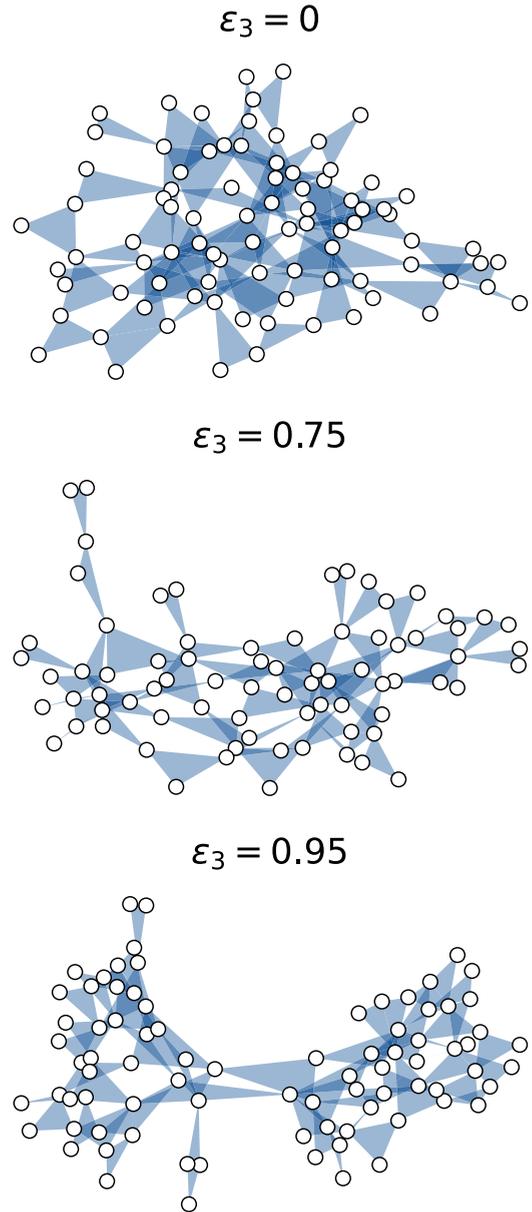


FIG. 1. An illustration of a 3-uniform hypergraph sampled from the HPPM model with varying strengths of community structure using XGI [30].

this generative model interpolates from an m -uniform hypergraph with completely random connections to an m -uniform hypergraph with two disconnected communities, while maintaining a constant expected degree for each node (see Fig. 1).

Consider a set V of N vertices divided into two communities of sizes ρN and $(1 - \rho)N$, with $0 < \rho < 1$. For a given set e of m nodes, we let the probability that a hyperedge connects the edges be

$$P(e \in E) = \begin{cases} p_{\text{in}}, & \text{all nodes in } e \text{ belong to} \\ & \text{the same community,} \\ p_{\text{out}} & \text{otherwise.} \end{cases} \quad (1)$$

Note that we are following the ‘‘all-or-nothing’’ (AON) definition in Ref. [37], which considers a hyperedge to connect

two different communities if it contains at least one member of each community. One can relax this assumption by accounting for the fraction of nodes in each community [37,38], but we do not consider this case here.

To create hypergraphs with a tunable amount of community structure, we change the relative size of p_{in} and p_{out} while ensuring that the mean degree $\langle k^{(m)} \rangle$, which henceforth will be indicated with $\langle k \rangle$, is constant. When there is no community structure, both probabilities are equal to the hypergraph Erdős-Rényi probability p_{ER} ,

$$p_{\text{out}} = p_{\text{in}} = p_{\text{ER}} = \langle k \rangle N / \left[m \binom{N}{m} \right] \approx \langle k \rangle \frac{(m-1)!}{N^m}, \quad (2)$$

where here and in the following we assume $N \gg m$ and approximate $\binom{N}{m}$ by $N^m/m!$. To determine p_{in} and p_{out} when there is community structure, we note that the total expected number of hyperedges M is given by the sum of intra- and intercommunity hyperedges

$$M = \binom{\rho N}{m} p_{\text{in}} + \binom{(1-\rho)N}{m} p_{\text{in}} + s \left[\binom{N}{m} - \binom{\rho N}{m} - \binom{(1-\rho)N}{m} \right] p_{\text{out}}. \quad (3)$$

Using $\langle k \rangle = mM/N$, approximating $\binom{N}{m}$ by $N^m/m!$, and using $\langle k \rangle (m-1)!/N^{m-1} = p_{\text{ER}}$, we find

$$p_{\text{ER}} = q_{\rho,m} p_{\text{in}} + [1 - q_{\rho,m}] p_{\text{out}}, \quad (4)$$

where

$$q_{\rho,m} = \rho^m + (1-\rho)^m. \quad (5)$$

Now we introduce the *imbalance parameter* $\epsilon_{\rho,m}$ [34] that interpolates between a hypergraph without community structure ($p_{\text{out}} = p_{\text{ER}}$) for $\epsilon_{\rho,m} = 0$ and a hypergraph with completely disconnected communities ($p_{\text{out}} = 0$) for $\epsilon_{\rho,m} = 1$. Accordingly, we parametrize p_{in} and p_{out} as

$$p_{\text{out}} = p_{\text{ER}}(1 - \epsilon_{\rho,m}), \quad (6)$$

$$p_{\text{in}} = p_{\text{ER}}(1 + r_{\rho,m} \epsilon_{\rho,m}). \quad (7)$$

Inserting this into Eq. (4), we find

$$r_{\rho,m} = \frac{1 - q_{\rho,m}}{q_{\rho,m}}, \quad (8)$$

which is the ratio between inter- and intracommunity edges. This sheds light on why the imbalance parameter in (7) is multiplied by $r_{\rho,m}$: since there are more potential inter-community hyperedges than intracommunity hyperedges, this factor is needed so that the mean degree remains constant as the relative size of p_{in} and p_{out} is changed.

For the special case of equal-sized communities, $\rho = 1/2$, the probabilities simplify to

$$p_{\text{in}} = p_{\text{ER}} + (2^{m-1} - 1) \epsilon_m p_{\text{ER}}, \quad (9)$$

$$p_{\text{out}} = p_{\text{ER}} - \epsilon_m p_{\text{ER}}. \quad (10)$$

Note that in this case, the factor $r_{1/2,m} = 2^{m-1} - 1$ can be understood by the fact that, given a node in one community, there are $2^{m-1} - 1$ more potential intercommunity links than

intracommunity links. In the following, we define $\epsilon_m \equiv \epsilon_{1/2,m}$ for simplicity. Also in this case, in terms of the expected intra- and intercommunity mean degrees

$$\langle k_{\text{in}} \rangle = \binom{N/2-1}{m-1} p_{\text{in}}, \quad (11)$$

$$\langle k_{\text{out}} \rangle = \left[\binom{N}{m-1} - \binom{N/2-1}{m-1} \right] p_{\text{out}}, \quad (12)$$

the imbalance parameter is given by

$$\epsilon_m = \frac{\langle k_{\text{in}} \rangle - \langle k_{\text{out}} \rangle}{(2^{m-1} - 1) \langle k \rangle}. \quad (13)$$

For a given value ρ and $\langle k \rangle$, changing the probabilities of intra- and intercommunity links according to Eqs. (6) and (7) produces hypergraphs with varying amounts of community structure. Figure 1 shows three hypergraphs obtained by the HPPM with $m = 3$, $\rho = 1/2$, $\langle q \rangle = 2$, and values of ϵ_3 equal to 0 (top), 0.75 (middle), and 0.95 (bottom).

Before moving on, we note that the HPPM differs from the generative model used in Ref. [29] in that the total expected number of hyperedges in the hypergraph is held constant as the amount of community structure is varied, allowing us to isolate the effect of community structure on the dynamics.

D. Opinion model and opinion disparity

Following Refs. [26,27], we study the hypergraph *collective contagion model*. This model allows nodes to hold binary opinions: susceptible (0) and infected (1). An infected node spontaneously transitions to the susceptible state at a rate $\gamma > 0$ independently of the states of neighboring nodes. A susceptible node may transition to the infected state at a rate β_m if that node is a member of a group of size m and all of the other members in this group are infected. It is assumed that the node may be infected independently by each such group to which it belongs. We note that the two opinion states are asymmetric in this model: in the absence of influence or infection by other groups, an individual will almost surely heal given enough time; however, a healthy individual will not become infected without the influence of other individuals. Therefore, the opinion corresponding to the infected state may only be sustained if the opinion is continually shared with neighbors. This is akin to a peer pressure effect where a behavior will likely die out without the influence of groups to sustain it.

In Refs. [26,27], it was found that this model admits bistable regions where a state with no infection and a state with a macroscopic fraction of infected nodes are simultaneously stable. Therefore, viewing hypergraphs with strong community structure as perturbations of disconnected hypergraphs, we expect that, depending on the parameters, we will find stable solutions where the fractions of infected nodes in each community are very different.

The observation above motivates the definition of opinion disparity we will use in this paper. Given a hypergraph with community structure, we define the opinion disparity, ψ_{ij} , between communities i and j as the maximum difference in average opinions between the two communities that can be sustained at equilibrium. For example, a hypergraph where all the nodes in one community have opinion 1 and all the

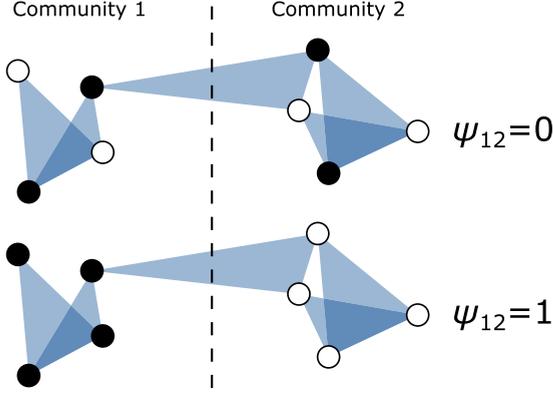


FIG. 2. Example of two communities with no opinion disparity (top) and maximum opinion disparity (bottom). Infected and susceptible nodes are shown as black and white, respectively.

nodes in the other community have opinion 0 in the steady state has the maximum possible opinion disparity, $\psi_{ij} = 1$. On the other hand, when the fraction of nodes with opinion 1 is the same in each community, the opinion disparity is zero. Note that this definition of opinion disparity assumes the existence of predefined communities, and that the community membership of each node is known. The opinion disparity differs from polarization in that polarization is often measured as a property, such as bimodality, of the opinion distribution [39,40]. In this formalism, bimodal opinion distributions can have zero opinion disparity, depending on the community membership of the nodes (see Fig. 2). Therefore, our definition is more appropriate to quantify the extent to which observed community structure in a hypergraph is correlated with nodal opinions. Caution is advised when community structure cannot be reasonably expected because the value of the opinion disparity is dependent on the nodal community labels. In this paper, we do not focus on the inference of these labels, but instead refer the reader to articles discussing higher-order community detection, such as Refs. [33,37,41,42].

III. MEAN-FIELD ANALYSIS

In this section, we derive and analyze mean-field equations describing the dynamics of the opinion model introduced in Sec. IID for hypergraphs with community structure. For simplicity, we only consider hypergraphs with hyperedges of sizes 2 (links) and 3 (triangles). First, we consider the case of equal-sized, i.e., balanced, communities.

A. Balanced communities

Consider a hypergraph of size N with hyperedges of sizes 2 (links) and 3 (triangles) assigned following the HPPM in Sec. IIC with parameters $\rho = 1/2$, $\epsilon_2 \equiv \epsilon_{1/2,2}$, and $\epsilon_3 \equiv \epsilon_{1/2,3}$, respectively. We denote the fraction of infected individuals in community i as x_i and, assuming N is large, extend the mean-field model of Ref. [26] to include community structure.

We derive a rate equation for the fraction of infected nodes in the first community, x_1 , of the form

$$\frac{dx_1}{dt} = -\gamma x_1 + (1 - x_1)[\beta_2 N_2 + \beta_3 N_3], \quad (14)$$

where $-\gamma x_1$ is the rate of spontaneous healing, $(1 - x_1)$ is the fraction of nodes that are susceptible, and N_2, N_3 are the expected number of infected links and triangles, respectively, to which a randomly chosen node in community 1 belongs. The corresponding equation for x_2 can be obtained by permuting x_1 and x_2 .

The number of infected contacts via pairwise interactions, N_2 , can be obtained by adding the expected number of infected neighbors in communities 1 and 2. Note that, according to the planted partition model for $m = 2$, the probability that two nodes are connected to each other is $p_{\text{in}} = (1 + \epsilon_2)\langle k \rangle / N$ if they belong to the same community (in this case, community 1) and $p_{\text{out}} = (1 - \epsilon_2)\langle k \rangle / N$ if they belong to different communities [cf. Eqs. (9) and (10)]. Multiplying these by the expected number of infected nodes in community i , $x_i(N/2)$, and adding, we find

$$N_2 = \frac{1}{2}\langle k \rangle [(1 + \epsilon_2)x_1 + (1 - \epsilon_2)x_2].$$

To find the number of infected triangles to which a node in community 1 belongs, we proceed similarly. According to the planted partition model for $m = 3$, the probability that three nodes form a 3-hyperedge is $p_{\text{in}} = 2(1 + 3\epsilon)\langle q \rangle / N^2$ if they are in the same community and $p_{\text{out}} = 2(1 - \epsilon)\langle q \rangle / N^2$ if they are in different communities. Adding together the cases in which the two other nodes are both in community 1, in communities 1 and 2, in communities 2 and 1, or both in community 2, and taking into account that the number of unique triangles in each community including a given node is $\binom{N/2}{2} \approx N^2/8$, we find

$$N_3 = \frac{1}{4}\langle q \rangle [(1 + 3\epsilon)x_1^2 + 2(1 - \epsilon)x_1x_2 + (1 - \epsilon)x_2^2].$$

Putting all the terms together, we find

$$\begin{aligned} \frac{dx_1}{dt} = & -\gamma x_1 + \frac{\beta_2}{2}\langle k \rangle (1 - x_1)[(x_1 + x_2) + \epsilon_2(x_1 - x_2)] \\ & + \frac{\beta_3}{4}\langle q \rangle (1 - x_1)[(x_1 + x_2)^2 + \epsilon_3(3x_1^2 - 2x_1x_2 - x_2^2)]. \end{aligned}$$

Setting $\gamma = 1$ (after, if necessary, rescaling time), defining $\tilde{\beta}_2 = \beta_2\langle k \rangle$, $\tilde{\beta}_3 = \beta_3\langle q \rangle$, and including the analogous equation for dx_2/dt , we obtain the system of equations

$$\begin{aligned} \frac{dx_1}{dt} = & -x_1 + \frac{\tilde{\beta}_2}{2}(1 - x_1)[(x_1 + x_2) + \epsilon_2(x_1 - x_2)] \\ & + \frac{\tilde{\beta}_3}{4}(1 - x_1)[(x_1 + x_2)^2 + \epsilon_3(3x_1^2 - 2x_1x_2 - x_2^2)], \end{aligned} \quad (15)$$

$$\begin{aligned} \frac{dx_2}{dt} = & -x_2 + \frac{\tilde{\beta}_2}{2}(1 - x_2)[(x_1 + x_2) + \epsilon_2(x_2 - x_1)] \\ & + \frac{\tilde{\beta}_3}{4}(1 - x_2)[(x_1 + x_2)^2 + \epsilon_3(3x_2^2 - 2x_1x_2 - x_1^2)]. \end{aligned} \quad (16)$$

This system of equations characterizes the dynamics of the average opinions in communities 1 and 2, x_1 and x_2 . We refer

to stable fixed points with $x_1 \neq x_2$ as *asymmetric* fixed points, and to fixed points with $x_1 = x_2$ as *symmetric* fixed points. According to our definition of opinion disparity in Sec. IID, the existence of stable asymmetric fixed points corresponds to nonzero opinion disparity, given by

$$\psi_{12} = \max\{|x_1 - x_2| \text{ such that } (x_1, x_2) \text{ is a stable fixed point of (15) and (16)}\}. \quad (17)$$

Since we will consider only two communities, for simplicity we will denote the opinion disparity by $\psi \equiv \psi_{12}$. Before discussing the regimes where opinion disparity exists, it is convenient to study the dynamics without opinion disparity, i.e., along the invariant line $x_1 = x_2 = x$. Along that line, the dynamics are given by

$$\frac{dx}{dt} = -x + \tilde{\beta}_2(1-x)x + \tilde{\beta}_3(1-x)x^2, \quad (18)$$

which describes the system without community structure studied in Ref. [26]. Equation (18) leads to the symmetric fixed points (corresponding to Eq. (4) in [26])

$$x = 0, \quad (19)$$

$$x = \frac{\tilde{\beta}_3 - \tilde{\beta}_2 \pm \sqrt{(\tilde{\beta}_2 + \tilde{\beta}_3)^2 - 4\tilde{\beta}_3}}{2\tilde{\beta}_3}. \quad (20)$$

Considering now the full dynamics, linearization of Eqs. (15) and (16) about the fixed point (0,0) shows that this fixed point becomes unstable for values of $\tilde{\beta}_2$ larger than the epidemic threshold $\tilde{\beta}_2^c = 1$. For $\tilde{\beta}_3 > 1$, the two nonzero symmetric fixed points (20) appear at a saddle-node bifurcation at $\tilde{\beta}_2 = 2\sqrt{\tilde{\beta}_3} - \tilde{\beta}_3$, and the smaller of these fixed points merges with $x = 0$ in a subcritical bifurcation as $\tilde{\beta}_2$ approaches $\tilde{\beta}_2^c = 1$. Interestingly, these critical values do not depend on ϵ_2 and ϵ_3 . Therefore, this analysis indicates that when the communities have the same size, community structure does not modify the Erdős-Rényi model epidemic threshold or onset of bistability, at least regarding the symmetric fixed points. Earlier studies corroborate this result for pairwise networks [43,44]. However, we note that more accurate approximations that include pair correlations (e.g., Ref. [45]) have shown that the epidemic threshold in hypergraph SIS models can depend on the strength of higher-order interactions. Therefore, we anticipate that there might be corrections to our results due to correlations not included in the mean-field analysis.

Having studied the symmetric fixed points, now we turn to the more interesting asymmetric fixed points that support opinion disparity. As an illustrative example, consider the case $\epsilon_2 = 0.5$, $\epsilon_3 = 0.75$, $\tilde{\beta}_2 = 0.2$, and $\tilde{\beta}_3 = 4$ shown in Fig. 3(a). In this case, the only fixed points are the three symmetric fixed points specified in Eqs. (19) and (20), and therefore there is no opinion disparity. Changing ϵ_3 to 0.85 results in the same symmetric fixed points and two additional asymmetric unstable fixed points [Fig. 3(b)] created via a pitchfork bifurcation. Increasing ϵ_3 to 0.95, two additional stable (solid black circles) and two unstable (empty circles) asymmetric fixed points are created through saddle-node bifurcations, resulting in positive opinion disparity [Fig. 3(c)]. Thus, we see that

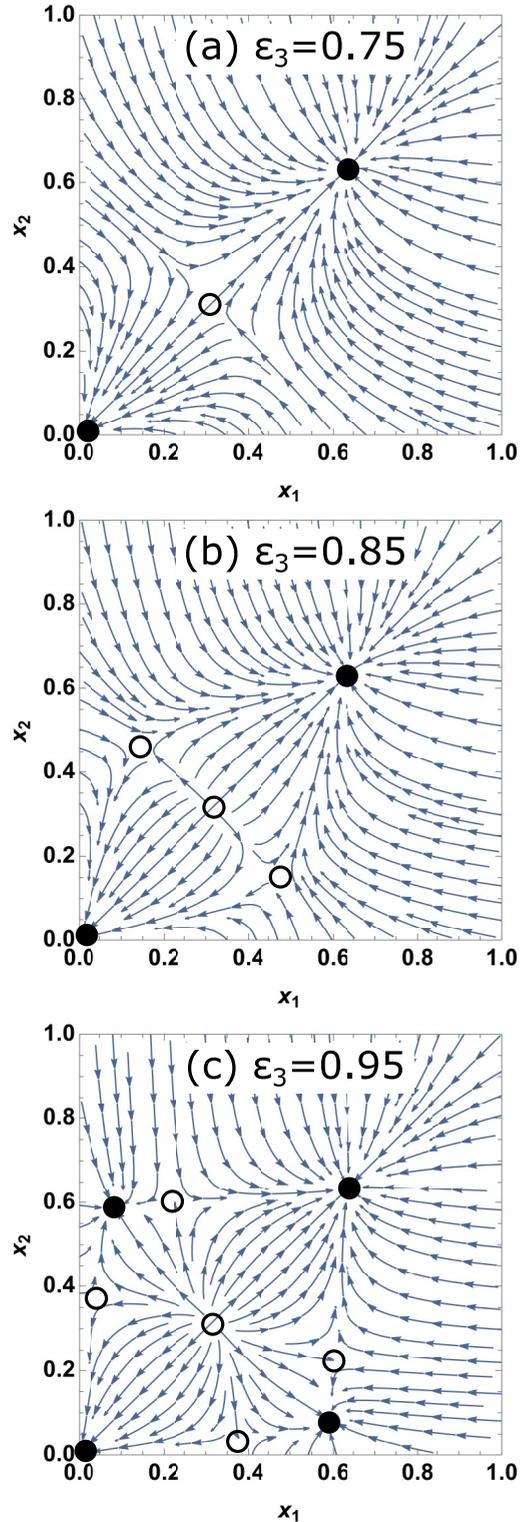


FIG. 3. A phase plot of Eqs. (15) and (16) with $\epsilon_2 = 0.5$, $\tilde{\beta}_2 = 0.2$, and $\tilde{\beta}_3 = 4$ for three different values of $\tilde{\epsilon}_3$. These plots illustrate the cases in which there are no asymmetric fixed points ($\epsilon_3 = 0.75$), only unstable asymmetric fixed points ($\epsilon_3 = 0.85$), and stable asymmetric fixed points ($\epsilon_3 = 0.95$). The filled circles correspond to stable fixed points, and the open circles correspond to unstable fixed points.

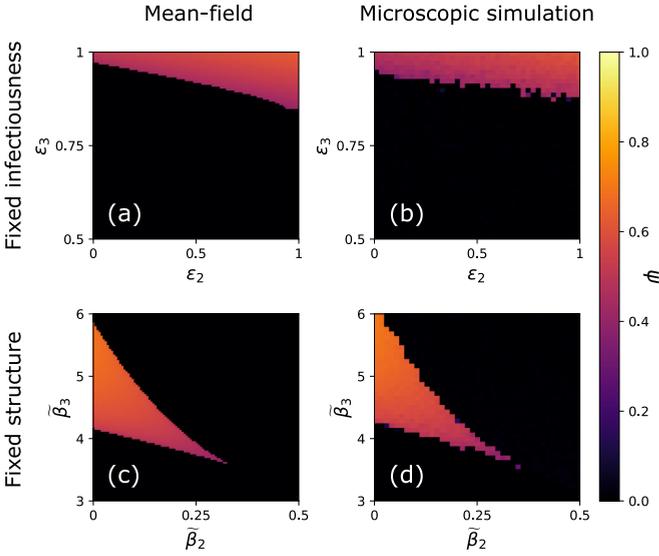


FIG. 4. The opinion disparity with respect to community structure (first row) and infectiousness (second row). Note the different ranges of ϵ_2 and ϵ_3 in panels (a) and (b) and $\tilde{\beta}_2$ and $\tilde{\beta}_3$ in panels (c) and (d). The opinion disparity is computed in two ways: from the stable fixed points of Eqs. (15) and (16) (first column) and from stochastic simulations of the contagion dynamics (second column).

there is a sudden onset of opinion disparity as ϵ_3 is increased above a certain threshold. This behavior is not present for the hypergraph contagion model on random null models without community structure or higher-order interactions.

In the remainder of this section, we will study how the opinion disparity depends on the community structure and infectiousness parameters ($\epsilon_2, \epsilon_3, \tilde{\beta}_2, \tilde{\beta}_3$) for the case of equal-sized communities. Given $\epsilon_2, \epsilon_3, \tilde{\beta}_2$, and $\tilde{\beta}_3$, we compute the opinion disparity numerically by setting the derivatives in Eqs. (15) and (16) equal to zero, using interval root finding methods to find all fixed points of these 2D coupled equations in $[0, 1]^2$, and lastly, selecting the stable fixed point with the largest value of $|x_1 - x_2|$. The stability of each fixed point is determined from the eigenvalues of the Jacobian associated with Eqs. (15) and (16), shown in Appendix C.

In Fig. 4(a), we plot the opinion disparity ψ calculated as described above as a function of ϵ_2 and ϵ_3 with $\tilde{\beta}_2 = 0.2$ and $\tilde{\beta}_3 = 4$. We see that there is a pronounced region for large ϵ_3 where opinion disparity occurs. The existence of opin-

ion disparity is much more sensitive to community structure of the triangles (ϵ_3) than that of the links (ϵ_2). To validate our results, in Fig. 4(b) we plot the opinion disparity found from numerical simulation of the infection model on a hypergraph with links and triangles assigned following the HPPM with $m = 2, \langle k \rangle = 20$ and $m = 3, \langle q \rangle = 20$, respectively, and $N = 10^4$. Further details on the numerical simulations may be found in Appendix B. The agreement between mean-field theory and the stochastic simulations is reasonable. One possible explanation for discrepancies between the two is that although opinion disparity may occur in the mean-field equations according to the definition in Eq. (17), in practice the fixed point might be too weakly stable to allow sustained opinion disparity to occur under the presence of finite-size effects [cf. Fig. 8(a)].

Now we look at the effect of the infection rates for a fixed community structure. We expect opinion disparity to be related to the bistable regime of a single community, which occurs when $\tilde{\beta}_3 > 1, 2\sqrt{\tilde{\beta}_3} - \tilde{\beta}_3 < \tilde{\beta}_2 < 1$. If the communities were completely disconnected, this bistable region would correspond to stable opinion disparity. In practice, connections between the communities result in a much smaller opinion disparity region as shown in Figs. 4(c) and 4(d). In Fig. 4(c), we plot ψ for $\tilde{\beta}_2$ on the interval $[0, 0.5]$ and $\tilde{\beta}_3$ on the interval $[3, 6]$ for $\epsilon_2 = 0.5$ and $\epsilon_3 = 0.95$ found numerically from the mean-field equations, and in Fig. 4(d) we show the corresponding microscopic simulations using the same HPPM described above. In Fig. 5, we explore this further by showing how the opinion disparity region shrinks as ϵ_2 (a) and ϵ_3 (b) are decreased from 1 to 0 and 0.95, respectively, when starting from two completely disconnected communities.

In summary, the analysis of the balanced communities shows that opinion disparity can occur for large values of the triangle community structure parameter ϵ_3 , and that relatively few connections between the communities can significantly reduce the size of the region in $(\tilde{\beta}_2, \tilde{\beta}_3)$ space where opinion disparity occurs. In the next section, we will study the case of imbalanced (i.e., different sized) communities. As we will see, even a small difference in size can cause dramatic changes in the opinion disparity.

B. Imbalanced communities

We now explore the effect of unequal community sizes. Considering the HPPM model of Sec. II C and performing the same calculations as in Sec. III A for an arbitrary ρ , we obtain

$$\begin{aligned} \frac{dx_1}{dt} = & -x_1 + \tilde{\beta}_2(1-x_1)[\rho(1+r_2\epsilon_2)x_1 + (1-\rho)(1-\epsilon_2)x_2] \\ & + \tilde{\beta}_3(1-x_1)[\rho^2(1+r_3\epsilon_3)x_1^2 + 2\rho(1-\rho)(1-\epsilon_3)x_1x_2 + (1-\rho)^2(1-\epsilon_3)x_2^2], \end{aligned} \quad (21)$$

$$\begin{aligned} \frac{dx_2}{dt} = & -x_2 + \tilde{\beta}_2(1-x_2)[(1-\rho)(1+r_2\epsilon_2)x_2 + \rho(1-\epsilon_2)x_1] \\ & + \tilde{\beta}_3(1-x_2)[(1-\rho)^2(1+r_3\epsilon_3)x_2^2 + 2(1-\rho)\rho(1-\epsilon_3)x_2x_1 + \rho^2(1-\epsilon_3)x_1^2], \end{aligned} \quad (22)$$

where we let $\epsilon_2 = \epsilon_{\rho,2}, \epsilon_3 = \epsilon_{\rho,3}, r_2 = r_{\rho,2}$, and $r_3 = r_{\rho,3}$. As before, we can linearize this system and compute the eigenvalues of the Jacobian at $x_1 = x_2 = 0$ to determine stability.

The eigenvalues are given by

$$\lambda = -1 + \frac{1}{2}\tilde{\beta}_2(1+r_2\epsilon_2 \pm \sqrt{\Delta}),$$

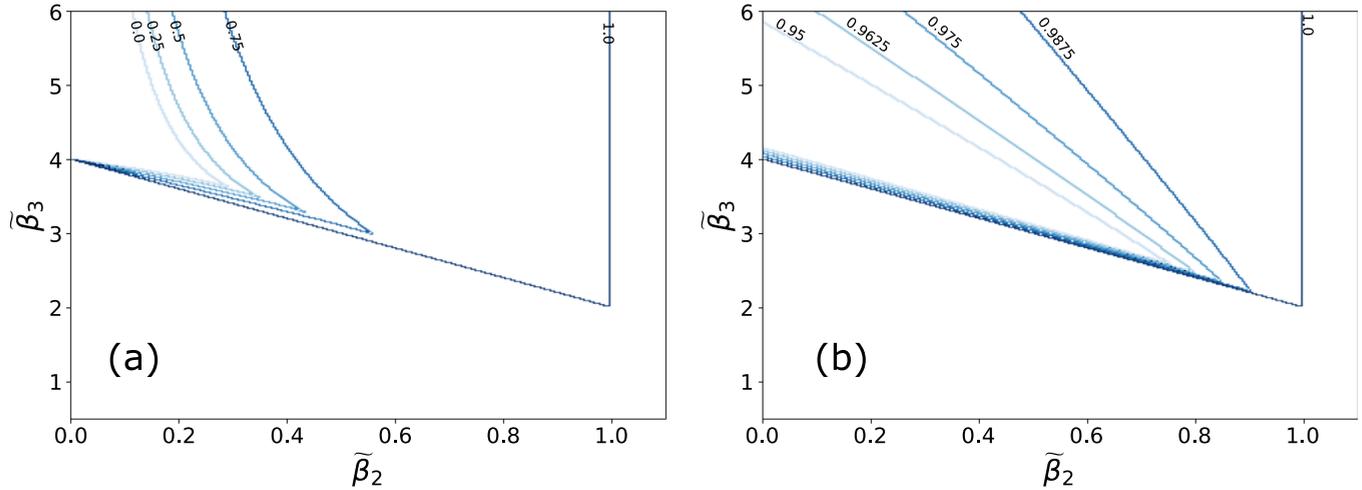


FIG. 5. A phase diagram illustrating how changes with (a) ϵ_2 and (b) ϵ_3 affect the regions over which opinion disparity can occur. In this diagram, the upper left region enclosed by each line corresponds to the $\tilde{\beta}_2, \tilde{\beta}_3$ region in which opinion disparity can occur for given values of ϵ_2 and ϵ_3 , and the remaining area is where opinion disparity does not occur. The boundary lines in panel (a) correspond to $\epsilon_2 = 0, 0.25, 0.5, 0.75, 1.0$ with $\epsilon_3 = 1$, and the boundary lines in panel (b) correspond to $\epsilon_3 = 0.95, 0.9675, 0.975, 0.9875, 1.0$ with $\epsilon_2 = 1$.

where

$$\Delta = 1 + 2\epsilon_2[4(\rho - 1)\rho + (1 - 2\rho)^2 r_2] + \epsilon_2^2[(1 - 2\rho)^2 r_2^2 - 4(\rho - 1)\rho].$$

Setting the maximal eigenvalue equal to zero and substituting the definition of r_2 determines the epidemic threshold

$$\tilde{\beta}_2^c = \frac{2}{1 + \sqrt{1 + \frac{4b}{a}\epsilon_2 - \frac{4c}{a^2}\epsilon_2^2} + (1/a - 1)\epsilon_2}, \quad (23)$$

where $a = 1 - 2\rho + 2\rho^2$, $b = \rho(\rho - 1)$, and $c = 3\rho^4 - 6\rho^3 + 4\rho^2 - \rho$ are functions of ρ . When $\rho = 1/2$ we obtain $a = 1/2$, $b = -1/4$, and $c = -1/16$, and we recover the epidemic threshold $\tilde{\beta}_2^c = 1$. For all other values of ρ , however, the epidemic threshold depends on ϵ_2 . To illustrate this, in Fig. 6 we plot $\tilde{\beta}_2^c$ as a function of ρ for $\epsilon_2 = 0, 0.25, 0.5, 0.75$, and 1. In agreement with our previous calculations, the epidemic threshold is independent of ϵ_2 for balanced com-

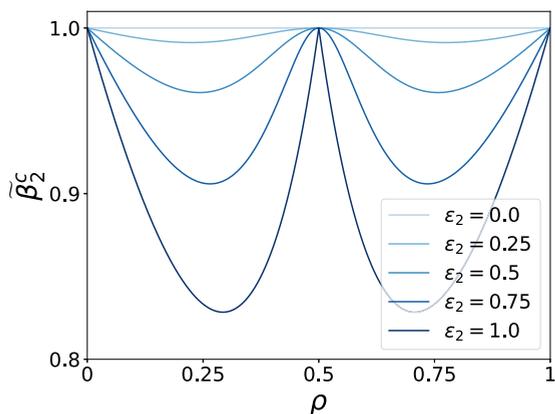


FIG. 6. A plot of the epidemic threshold $\tilde{\beta}_2^c$ with respect to ρ predicted with Eq. (23). The line corresponding to $\epsilon_2 = 0$ is the epidemic threshold for the Erdős-Rényi case.

munities. This plot also shows that different relative sizes are optimal in the spread of a network SIS contagion in the sense that they result in a lower epidemic threshold. Finally, we note that the epidemic threshold is independent of ϵ_3 for all values of ρ .

Changing the relative sizes of the communities not only changes the epidemic threshold, but it also affects the presence and strength of opinion disparity. To quantify opinion disparity in more detail, we define the one-sided opinion disparities to be

$$\psi_1 = \tilde{x}_1 - \tilde{x}_2, \quad (24)$$

$$\psi_2 = \hat{x}_2 - \hat{x}_1, \quad (25)$$

where $(\tilde{x}_1, \tilde{x}_2)$ is the stable fixed point of Eqs. (21) and (22) with the largest value of $x_1 - x_2$, and (\hat{x}_1, \hat{x}_2) is the stable fixed point of Eqs. (21) and (22) with the largest value of $x_2 - x_1$. This allows us to more easily observe which fixed points disappear and appear as we vary ρ .

In Fig. 7 we illustrate how opinion disparity changes as the relative sizes of the communities are changed by slightly tuning the parameter ρ away from $\rho = 1/2$. In the top panel, we plot the quantity $\psi_1 - \psi_2$ as a function of ρ and ϵ_2 . This quantity is positive (light colors) when the state with larger disparity has a larger value of x_1 , and negative (dark colors) when it has a larger value of x_2 . The amount and type of opinion disparity depend very sensitively on both the size imbalance parameter ρ and the amount of pairwise community structure ϵ_2 . In panels (1) and (2) we show the one-sided opinion disparities ψ_1 and ψ_2 as a function of ρ for two fixed values of ϵ_2 , corresponding to the two horizontal lines shown in the top panel. These figures show how asymmetric fixed points appear and disappear as the relative size of the communities changes, giving rise to the pattern in the top panel. The phase plots corresponding to points (a)–(d) marked with stars in the top panel of Fig. 7 are displayed in Figs. 8(a)–8(d), overlaid with the numerically calculated

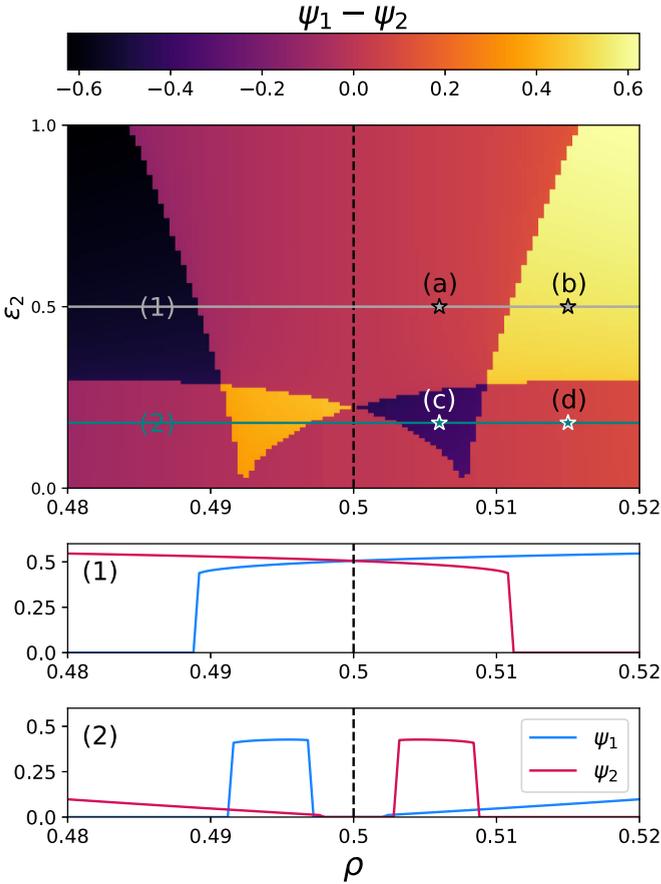


FIG. 7. In this plot, $\langle k \rangle = \langle q \rangle = 20$, $\epsilon_3 = 0.95$, $\gamma = 1$, $\beta_2 = 0.2/\langle k \rangle$, and $\beta_3 = 4/\langle q \rangle$. The top panel illustrates the difference $\psi_1 - \psi_2$ as a function of ρ and ϵ . The two solid horizontal lines (1) and (2) correspond to the panels below, which plot $\psi_1 - \psi_2$ for $\epsilon_2 = 0.18$ and $\epsilon_2 = 0.5$, respectively. The points on each cross-section are at $\rho = 0.506$ and 0.515 . $\psi_1 - \psi_2$ is an odd function about the dashed line defined by $\rho = 1/2$. For large enough ϵ_2 as seen in cross-section (1), when community 1 is larger than community 2 (i.e., $\rho > 0.5$), the opinion disparity is greater when community 1 adopts opinion 1 than when it adopts opinion 0, indicated by $\psi_1 > \psi_2$. In panel (2), we see that there are regimes for small enough ϵ_2 where opinion disparity *only* exists for imbalanced community structure.

fixed points and several reference trajectories. The fixed points are determined by first randomly infecting a fraction ρ_1 of communities 1 and ρ_2 of community 2 from a grid of initial conditions $(\rho_1, \rho_2) \in \{0, 0.05, \dots, 1\} \times \{0, 0.05, \dots, 1\}$ and simulating the system for $t = 0, \dots, 300$. Each black dot represents a weighted average of the last 10 time steps for each initial condition (more details are contained in Appendix B). These numerical simulations show that the stable mean-field fixed points are generally a good representation of the stable states of the stochastic model. However, it is possible that, even though a fixed point is stable in the mean-field description, finite-size effects can drive the stochastic model towards another fixed point. This is illustrated in Fig. 8(c), which shows five sample trajectories starting from the initial condition $(x_1, x_2) = (0, 1)$. After hovering around the asymmetric

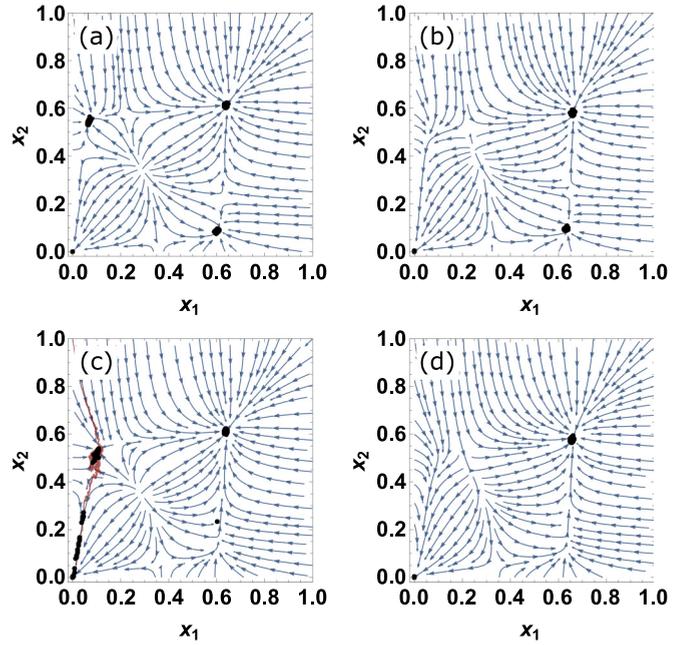


FIG. 8. The phase plots corresponding to points (a)–(d) in Fig. 7 with the fixed points (x_1, x_2) computed from numerical simulations overlaid on top. Each point is generated by computing the time-weighted average of the fraction of communities 1 and 2 infected for different initial conditions and hypergraph realizations, described more in Appendix B. Five sample trajectories with the initial condition of $(x_1, x_2) = (0, 1)$ are shown in red in panel (c) to illustrate that fluctuations due to finite-size effects can cause the state of the system to leave the asymmetric fixed point.

fixed point, some of them drift towards the stable fixed point $(0, 0)$.

The phase plots in Fig. 8 illustrate how the fixed points and their stability change as the imbalance parameter is varied. For small enough ϵ_2 we see the appearance of regions where opinion disparity cannot occur for equal community sizes, but can occur for a small range of imbalanced community structures [e.g., see panel (c)]. In addition, for a large enough difference in community sizes, a majority of “1” opinions is untenable for the smaller community. The plot of $\psi_1 - \psi_2$ with respect to ρ and ϵ_3 is qualitatively similar, so we omit it here. Increasing ρ from the balanced case ($\rho = 0.5$) increases the size of the region where $\psi_2 > 0$ up to a critical value, after which it decreases. Likewise, the size of the region where $\psi_1 > 0$ decreases with increasing ρ . The same phenomenon occurs for ϵ_3 with respect to ρ .

In Fig. 8, we see close agreement between the phase diagram corresponding to Eqs. (21) and (22) and the fixed points computed from numerical simulations.

Notably, regions with strong opposite opinion disparity types are adjacent: there are continuous trajectories in the (ρ, ϵ_2) phase space such that there is a discontinuous change of sign in the quantity $\psi_1 - \psi_2$ along them. In practice this would mean that the establishment or destruction of a few intercommunity ties, or a small change in community memberships, could result in a sudden change in opinion disparity.

IV. DISCUSSION

In this paper, we studied the effect of community structure and higher-order interactions on the existence of opinion disparity in a simple collective social contagion model. To do so, we introduced a simple generative hypergraph model for hypergraphs with tunable community structure. We analyzed the dynamics of the contagion model using a mean-field approach, and we found that opinion disparity can be sustained for a relatively large range of community structure and infection rate parameters. In addition, we found that very small changes in the community sizes can lead to dramatic changes in the opinion disparity.

Here we discuss the limitations of our approach and possible directions for further research. First, our formalism is a model of discrete opinions, but in many situations opinions are better described as a continuous spectrum. In addition, opinions may not be one-dimensional, but rather multidimensional. As mentioned in Sec. II D, the contagion model we chose is asymmetric, which leads to behavior that one may not see if the two opinions were equivalent. Thus, our model is more appropriate to describe opinions or behavior that need to be sustained by group pressure, instead of describing a choice between two symmetric opinions. In addition, we assumed that the dynamics of the opinion formation process does not affect the underlying structure of the hypergraph. The formation of echo chambers, however, may be the result of not only the dynamics of opinion formation but also of the hypergraph adapting to dissolve interactions between individuals with dissimilar opinions [46]. For simplicity, we considered hypergraphs with only two communities and interactions of sizes 2 and 3. However, our models and methods could be extended to model any number of communities and hyperedge sizes. In addition, we only considered variants of the stochastic block model with homogeneous degree, but empirical higher-order systems often contain degree heterogeneity. Including this heterogeneity in our hypergraph models could help bridge the gap between theory and empirical data sets. To this end, it would be useful to simulate the contagion model that we have described on empirical data sets to see if opinion disparity occurs, similar to the approach of Ref. [28]. Lastly, another drawback of our approach is that we need the community labels of the nodes to compute the opinion disparity, but ground truth labels are not always available. To alleviate this, one could complement our approach with community detection algorithms.

Despite these limitations, our work shows that the community structure of higher-order interactions may be one of the ingredients contributing to opinion disparity. In contrast to models of opinion formation which require opinion homophily for opinion disparity to exist, our model only requires the presence of higher-order interactions and sufficiently disconnected communities for opinion disparity to occur. Thus, our results should be interpreted as the exploration of a complementary mechanism of opinion disparity formation. Our results also show that communities of differing sizes can modify the epidemic threshold and create regimes where opinion disparity can occur despite being impossible for equally sized communities.

The code and data sets supporting this work are openly available on [GitHub](#) [47].

ACKNOWLEDGMENTS

N.L. acknowledges financial support from the National Science Foundation Grant No. 2121905, ‘‘HNDS-I: Using Hypergraphs to Study Spreading Processes in Complex Social Networks,’’ and from the National Institutes of Health 1P20 GM125498-01 Centers of Biomedical Research Excellence Award. J.G.R. acknowledges support from National Science Foundation Grant No. DMS-2205967. N.L. would like to acknowledge helpful feedback from Mari Kawakatsu on the conception of this project.

APPENDIX A: AN EFFICIENT ALGORITHM FOR SAMPLING m -HSBM HYPERGRAPHS

Simply sampling naively from the list of all possible hyperedges of size m and accepting them with probability $p \ll 1$ has complexity $O(N^m)$, which can be computationally prohibitive for large m . We modify the algorithm presented in Ref. [48] to m -uniform hypergraphs to sample from the m -HSBM with time complexity $O(m(N + |E|))$.

In Sec. II C it is implicitly assumed that no multiedges or self-loops may occur. However, given an index in the list of all possible unique hyperedges, it can be expensive to recover the hyperedge corresponding to that index. If instead we allow these artifacts to occur, then it is $O(m)$ to recover the hyperedge of interest in contrast to iterating through all unique combinations, which is $O(mN)$. Because of this modification, in practice, we divide every probability derived in Sec. II C by $m!$ to account for the increase in possible hyperedges because of multiedges and hyperedges containing self-loops. Because allowing multiple instances of the same node changes the number of unique neighbors, we remove these artifacts when they occur.

Consider an m -uniform hypergraph with N nodes and a community label g_i for each node i . We denote by \mathbf{g} the vector of all node labels, i.e., $\mathbf{g} = [g_1, \dots, g_N]$. The number of unique community labels is G , and, as in Sec. II B, P_{g_1, \dots, g_m} specifies the probability that nodes i_1, \dots, i_m with community labels g_{i_1}, \dots, g_{i_m} form a hyperedge. The function $\theta(g)$ returns a vector of all the nodes that have community label g , and $|\theta(g)|$ is the number of nodes with community label g . We iterate through each entry b_1, \dots, b_m of the tensor P , and for fixed group assignments of the nodes in a hyperedge, the probability of generating a hyperedge is constant. We generate all the hyperedges associated with these ordered community assignments. The hyperedges are elements of the set formed by the Cartesian product of the indices in each partition, and the maximum index is given by the product $\prod_{b \in \mathbf{b}} |\theta(b)|$.

Our algorithm is an extension of the algorithm in Ref. [48], and the main idea is this: instead of iterating through all possible edges and accepting an edge with probability $P_{g_1, \dots, g_m} = p$, which is expensive when $p \ll 1$, we simply skip the edges that would be rejected by sampling from a geometric distribution. While the current index is less than the maximum index, we increment the index with steps $s \sim \text{Geometric}_1(p)$, the distribution of the number of Bernoulli trials needed for a success. For a given index, we convert to a list of m node labels with Algorithm 1. Because we simulate the community connection probability tensor patch-by-patch, for a given patch we specify the community to which each node belongs as an

Algorithm 1. Return an m -hyperedge from a specified index, given community partitions (IndexToEdge).

```

Input:  $i, \mathbf{b}, \theta$ 
Output:  $e$ 
 $e = \emptyset$ 
 $m = |\mathbf{b}|$ 
 $r = m$ 
while  $r > 0$  do
     $j = \left[ i / \left( \prod_{p=r+1}^m |\theta(b_r)| \right) \right] \bmod |\theta(b_r)|$ 
     $v = \theta(b_r)_j$ 
     $e \leftarrow e \cup v$ 
     $r \leftarrow r - 1$ 
end
return  $e$ 

```

ordered list. The algorithm for sampling from the m -HSBM is given in Algorithm 2.

Generating an Erdős-Rényi hypergraph is a special case of Algorithm 2 where there is a single community [$G = 1$, $g_i = 1$, $i = 1, \dots, N$, $|\theta(\mathbf{g})| = N$], and an efficient algorithm is provided in Ref. [47].

APPENDIX B: NUMERICAL SIMULATIONS

We use the same Gillespie algorithm described in Ref. [49] to simulate the hypergraph SIS model efficiently. We obtain the community labels of the nodes from the labels specified in our m -HPPM model. We use these labels to return the number of nodes infected in each community with respect to time based on the community labels of the nodes. There are three stable states that the simulation can reach in equilibrium: first, the fractions of communities 1 and 2 can remain well-separated indicating the possibility of opinion disparity;

Algorithm 2. Generating the m -uniform stochastic block model for hypergraphs (m -HSBM).

```

Input:  $\mathbf{g}, m, P$ 
Output:  $E$ 
 $N = |\mathbf{g}|$ 
 $G = |\text{unique } b \in \mathbf{g}|$ 
 $B = 1, \dots, G$ 
 $\theta : b \mapsto (i \mid g_i = b, i = 1, \dots, N)$ 
 $E = \emptyset$ 
for  $\mathbf{b} = (b_1, \dots, b_m) \in B \times \dots \times B$  do
     $M = \prod_{b \in \mathbf{b}} |\theta(b)|$ 
     $p = P_{b_1, \dots, b_m}$ 
     $i \sim \text{Geometric}_1(p)$ 
    while  $i < M$  do
         $e = \text{IndexToEdge}(i, \mathbf{b}, \theta)$ 
        if  $|e| = m$  then
             $E \leftarrow E \cup e$ 
        end
         $i \leftarrow i + \text{Geometric}_1(p)$ 
    end
end
return  $E$ 

```

second, the epidemic equilibrium where the average fraction infected is identical in communities 1 and 2; and third, where the simulation dies out and there are no infected individuals in either community 1 or community 2. We assume that if we start in the $(\mathbf{x}_1 = \mathbf{1}, \mathbf{x}_2 = \mathbf{0})$ state and the structural and dynamical parameters admit a polarized stable state, then x_1 and x_2 will remain well-separated. We specify that the initial state of every node in community 1 is infected and every node in community 2 is susceptible, and we run the simulation until a maximum time of $t_{\max} = 100$ is reached or every node is in the susceptible state. There is a nonzero probability, however, that the second or third case will occur for a weakly stable asymmetric fixed point. We heuristically tuned the simulation time to minimize the effect of finite-size effects. We take the absolute value of the difference between these two time series to obtain the opinion disparity as a function of time and perform time-weighted averaging of the resulting time series from the last 10% of the time series (described in more detail in Ref. [49]). For each value of ϵ_2 and ϵ_3 , we generate a single realization from the m -HPPM, and for each set of infectious parameter values, we run a single simulation to preserve the sharp transitions between regions that admit opinion disparity and those that do not.

When generating the fixed points overlaying the phase diagrams in Fig. 8, we simulate the contagion dynamics for many different initial states. We generate the initial nodal states by iterating over a grid of initial infection densities $(\rho_1, \rho_2) \in \{0, 0.05, \dots, 1\} \times \{0, 0.05, \dots, 1\}$. For each (ρ_1, ρ_2) combination, we sample uniformly at random $\rho_1 \rho N$ nodes from community 1 and $\rho_1(1 - \rho)N$ nodes from community 2. We set the state of these nodes to be infected and simulate the contagion process with this initial condition up to $t_{\max} = 300$. We generate a single m -HPPM hypergraph with $\langle k \rangle = \langle q \rangle = 50$, and for each initial condition, we simulate the contagion process. We perform time-weighted averaging on the states of the nodes in communities 1 and 2 as described above.

APPENDIX C: CALCULATING THE STABILITY OF FIXED POINTS

Here, we present the Jacobian, $J(x_1, x_2)$, for the hypergraph SIS model on both the planted partition model and the imbalanced planted partition model.

The Jacobian for the system of equations (15) and (16) governing contagion spread on the planted partition model is given by

$$\begin{aligned}
 J_{1,1} = & -1 + \frac{\tilde{\beta}_2}{2}(1 - x_1)[1 + \epsilon_2] \\
 & - \frac{\tilde{\beta}_2}{2}[x_1 + x_2 + \epsilon_2(x_1 - x_2)] \\
 & + \frac{\tilde{\beta}_3}{2}(1 - x_1)[x_1 + x_2 + \epsilon_3(3x_1 - x_2)] \\
 & - \frac{\tilde{\beta}_3}{4}[(x_1 + x_2)^2 + \epsilon_3(3x_1^2 - 2x_1x_2 - x_2^2)], \quad (C1)
 \end{aligned}$$

$$\begin{aligned}
 J_{1,2} = & \frac{\tilde{\beta}_2}{2}(1 - x_1)[1 - \epsilon_2] \\
 & + \frac{\tilde{\beta}_3}{2}(1 - x_1)[x_1 + x_2 - \epsilon_3(x_1 + x_2)]. \quad (C2)
 \end{aligned}$$

$J_{2,2}$ and $J_{2,1}$ can be obtained by noting the symmetry of Eqs. (15) and (16) and substituting $x_1 \leftrightarrow x_2$ into the expressions for $J_{1,1}$ and $J_{1,2}$.

For the system of equations (21) and (22) governing the contagion dynamics of the imbalanced planted partition model, the Jacobian is

$$J_{1,1} = -1 + \beta_2 \langle k \rangle \rho (1 - x_1) (1 + r_{\rho,2} \epsilon_2) - \beta_2 \langle k \rangle [(1 - \rho)(1 - \epsilon_2)x_2 + \rho(1 + r_{\rho,2} \epsilon_2)x_1] + 2\beta_3 \langle q \rangle (1 - x_1) [\rho^2(1 + r_{\rho,3} \epsilon_3)x_1 + (1 - \rho)\rho(1 - \epsilon_3)x_2]$$

$$- \beta_3 \langle q \rangle [\rho^2(1 + r_{\rho,3} \epsilon_3)x_1^2 + 2\rho(1 - \rho)(1 - \epsilon_3)x_1x_2 + (1 - \rho)^2(1 - \epsilon_3)x_2^2], \quad (C3)$$

$$J_{1,2} = \beta_2 \langle k \rangle (1 - \rho)(1 - \epsilon_2)(1 - x_1) + 2\beta_3 \langle q \rangle (1 - \rho)\rho(1 - \epsilon_3)(1 - x_1)x_1 + 2\beta_3 \langle q \rangle (1 - \rho)^2(1 - \epsilon_3)(1 - x_1)x_2. \quad (C4)$$

$J_{2,1}$ and $J_{2,2}$ can be calculated by noting that Eq. (22) may be obtained from Eq. (21) by substituting $x_1 \leftrightarrow x_2$ and $\rho \leftrightarrow (1 - \rho)$.

-
- [1] P. Clifford and A. Sudbury, A model for spatial conflict, *Biometrika* **60**, 581 (1973).
- [2] S. Galam, Minority opinion spreading in random geometry, *Eur. Phys. J. B* **25**, 403 (2002).
- [3] P. Chen and S. Redner, Majority rule dynamics in finite dimensions, *Phys. Rev. E* **71**, 036101 (2005).
- [4] K. Sznajd-weron and J. Sznajd, Opinion evolution in closed community, *Int. J. Mod. Phys. C* **11**, 1157 (2000).
- [5] H. Hassani, R. Razavi-Far, M. Saif, F. Chiclana, O. Krejcar, and E. Herrera-Viedma, Classical dynamic consensus and opinion dynamics models: A survey of recent trends and methodologies, *Inf. Fusion* **88**, 22 (2022).
- [6] C. Andris, D. Lee, M. J. Hamilton, M. Martino, C. E. Gunning, and J. A. Selden, The rise of partisanship and super-cooperators in the U.S. House of Representatives, *PLoS ONE* **10**, e0123507 (2015).
- [7] M. Cinelli, G. D. F. Morales, A. Galeazzi, W. Quattrociocchi, and M. Starnini, The echo chamber effect on social media, *Proc. Natl. Acad. Sci. USA* **118**, e2023301118 (2021).
- [8] *Political Polarization in the American Public* (Pew Research Center, 2014).
- [9] R. S. Baron, S. I. Hoppe, C. F. Kao, B. Brunzman, B. Linneweh, and D. Rogers, Social corroboration and opinion extremity, *J. Exp. Social Psych.* **32**, 537 (1996).
- [10] M. Macy, J. Kitts, A. Flache, and S. Benard, Polarization in dynamic networks: A Hopfield model of emergent structure, in *Dynamic Social Network Modeling and Analysis* (The National Academies, Washington, DC, 2003), Vol. 162.
- [11] E. Gilbert, T. Bergstrom, and K. Karahalios, Blogs are echo chambers: Blogs are echo chambers, in *Proceedings of the 2009 42nd Hawaii International Conference on System Sciences* (IEEE, Piscataway, NJ, 2009), pp. 1–10.
- [12] P. Dandekar, A. Goel, and D. T. Lee, Biased assimilation, homophily, and the dynamics of polarization, *Proc. Natl. Acad. Sci. USA* **110**, 5791 (2013).
- [13] F. Gargiulo and Y. Gandica, The role of homophily in the emergence of opinion controversies, *J. Artificial Soc. Social Sim.* **20**, 8 (2017).
- [14] G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch, Mixing beliefs among interacting agents, *Adv. Complex Syst.* **03**, 87 (2000).
- [15] R. Hegselmann and U. Krause, Opinion dynamics driven by various ways of averaging, *Computat. Econ.* **25**, 381 (2005).
- [16] M. Del Vicario, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi, Modeling confirmation bias and polarization, *Sci. Rep.* **7**, 40391 (2017).
- [17] K. Sasahara, W. Chen, H. Peng, G. L. Ciampaglia, A. Flammini, and F. Menczer, Social influence and unfollowing accelerate the emergence of echo chambers, *J. Computat. Social Sci.* **4**, 381 (2021).
- [18] S. Banisch and E. Olbrich, Opinion polarization by learning from social feedback, *J. Math. Sociol.* **43**, 76 (2019).
- [19] N. E. Leonard, K. Lipsitz, A. Bizyaeva, A. Franci, and Y. Lelkes, The nonlinear feedback dynamics of asymmetric political polarization, *Proc. Natl. Acad. Sci. USA* **118**, e2102149118 (2021).
- [20] M. Kawakatsu, Y. Lelkes, S. A. Levin, and C. E. Tarnita, Interindividual cooperation mediated by partisanship complicates Madison's cure for "mischiefs of faction," *Proc. Natl. Acad. Sci. USA* **118**, e2102148118 (2021).
- [21] V. C. Yang, D. M. Abrams, G. Kernell, and A. E. Motter, Why are U.S. parties so polarized? A "satisficing" dynamical model, *SIAM Rev.* **62**, 646 (2020).
- [22] F. Baumann, P. Lorenz-Spreen, I. M. Sokolov, and M. Starnini, Modeling Echo Chambers and Polarization Dynamics in Social Networks, *Phys. Rev. Lett.* **124**, 048301 (2020).
- [23] X. Wang, A. D. Sirianni, S. Tang, Z. Zheng, and F. Fu, Public Discourse and Social Network Echo Chambers Driven by Socio-Cognitive Biases, *Phys. Rev. X* **10**, 041042 (2020).
- [24] V. C. Yang, T. van der Does, and H. Olsson, Falling through the cracks: Modeling the formation of social category boundaries, *PLoS ONE* **16**, e0247562 (2021).
- [25] H. Z. Brooks and M. A. Porter, A model for the influence of media on the ideology of content in online social networks, *Phys. Rev. Res.* **2**, 023041 (2020).
- [26] I. Iacopini, G. Petri, A. Barrat, and V. Latora, Simplicial models of social contagion, *Nat. Commun.* **10**, 2485 (2019).
- [27] N. W. Landry and J. G. Restrepo, The effect of heterogeneity on hypergraph contagion models, *Chaos* **30**, 103117 (2020).
- [28] G. Ferraz de Arruda, G. Petri, P. M. Rodriguez, and Y. Moreno, Multistability, intermittency, and hybrid transitions in social contagion models on hypergraphs, *Nat. Commun.* **14**, 1375 (2023).
- [29] J. Noonan and R. Lambiotte, Dynamics of majority rule on hypergraphs, *Phys. Rev. E* **104**, 024316 (2021).

- [30] N. W. Landry, M. Lucas, I. Iacopini, G. Petri, A. Schwarze, A. Patania, and L. Torres, XGI: A Python package for higher-order interaction networks, *J. Open Source Softw.* **8**, 5162 (2023).
- [31] P. W. Holland, K. B. Laskey, and S. Leinhardt, Stochastic block-models: First steps, *Social Netw.* **5**, 109 (1983).
- [32] C. Kim, A. S. Bandeira, and M. X. Goemans, Stochastic block model for hypergraphs: Statistical limits and a semidefinite programming approach, [arXiv:1807.02884](https://arxiv.org/abs/1807.02884).
- [33] S. Cole and Y. Zhu, Exact recovery in the hypergraph stochastic block model: A spectral algorithm, *Lin. Alg. Appl.* **593**, 45 (2020).
- [34] M. Jerrum and G. B. Sorkin, The Metropolis algorithm for graph bisection, *Discrete Appl. Math.* **82**, 155 (1998).
- [35] A. Hickok, Y. Kureh, H. Z. Brooks, M. Feng, and M. A. Porter, A bounded-confidence model of opinion dynamics on hypergraphs, *SIAM J. Appl. Dyn. Syst.* **21**, 1 (2022).
- [36] D. Ghoshdastidar and A. Dukkipati, Consistency of spectral hypergraph partitioning under planted partition model, *Ann. Statist.* **45**, 289 (2017).
- [37] P. S. Chodrow, N. Veldt, and A. R. Benson, Generative hypergraph clustering: From blockmodels to modularity, *Sci. Adv.* **7**, eabh1303 (2021).
- [38] P. S. Skardal, S. Adhikari, and J. G. Restrepo, Multistability in coupled oscillator systems with higher-order interactions and community structure, *Chaos* **33**, 023140 (2023).
- [39] J. A. Adams, G. White, and R. P. Araujo, Mathematical measures of societal polarisation, *PLoS ONE* **17**, e0275283 (2022).
- [40] P. DiMaggio, J. Evans, and B. Bryson, Have American's social attitudes become more polarized? *Am. J. Sociol.* **102**, 690 (1996).
- [41] P. Chodrow, N. Eikmeier, and J. Haddock, Nonbacktracking spectral clustering of nonuniform hypergraphs, *SIAM J. Math. Data Sci.* **5**, 251 (2023).
- [42] B. Kamiński, V. Poulin, P. Prałat, P. Szufel, and F. Théberge, Clustering via hypergraph modularity, *PLoS ONE* **14**, e0224307 (2019).
- [43] E. M. Volz, J. C. Miller, A. Galvani, and L. A. Meyers, Effects of heterogeneous and clustered contact patterns on infectious disease dynamics, *PLoS Comput. Biol.* **7**, e1002042 (2011).
- [44] C. Stegehuis and T. Peron, Network processes on clique-networks with high average degree: The limited effect of higher-order structure, *J. Phys.: Complexity* **2**, 045011 (2021).
- [45] G. Burgio, A. Arenas, S. Gómez, and J. T. Matamalas, Network clique cover approximation to analyze complex contagions through group interactions, *Commun. Phys.* **4**, 111 (2021).
- [46] U. Kan, M. Feng, and M. A. Porter, An adaptive bounded-confidence model of opinion dynamics on networks, *J. Complex Netw.* **11**, cnac055 (2023).
- [47] N. Landry, `nwlandry/opinion-disparity-in-hypergraphs: v1.0`, <https://doi.org/10.5281/zenodo.8279500>.
- [48] J. C. Miller and A. Hagberg, Efficient generation of networks with given expected degrees, in *Algorithms and Models for the Web Graph*, edited by A. Frieze, P. Horn, and P. Prałat, Lecture Notes in Computer Science (Springer, Berlin, 2011), pp. 115–126.
- [49] N. W. Landry and J. G. Restrepo, Hypergraph assortativity: A dynamical systems perspective, *Chaos* **32**, 053113 (2022).