



OPEN

Echo chambers and information transmission biases in homophilic and heterophilic networks

Fernando Diaz-Diaz, Maxi San Miguel & Sandro Meloni✉

We study how information transmission biases arise by the interplay between the structural properties of the network and the dynamics of the information in synthetic scale-free homophilic/heterophilic networks. We provide simple mathematical tools to quantify these biases. Both Simple and Complex Contagion models are insufficient to predict significant biases. In contrast, a Hybrid Contagion model—in which both Simple and Complex Contagion occur—gives rise to three different homophily-dependent biases: emissivity and receptivity biases, and echo chambers. Simulations in an empirical network with high homophily confirm our findings. Our results shed light on the mechanisms that cause inequalities in the visibility of information sources, reduced access to information, and lack of communication among distinct groups.

Information transmission in the context of Information and Communication Technologies is a great opportunity to create a better-informed society, but in practice, these technologies are also promoting phenomena such as the viral spreading of fake news^{1–3}, echo chambers^{4–6}, perception biases like false consensus or majority illusions⁷ and social polarization^{5,6,8}. We understand by echo chambers situations in which the transmission of information among individuals belonging to the same opinion group is dominant, while transmission among individuals with different opinions is hindered. The real social impact of echo chambers and their causal link with misinformation cascades are debated topics^{9–11}, but data-driven and computational approaches confirm that the structural properties of social networks are tied to the emergence of echo chambers^{4–6}. In particular, the homophily of the network—that is, the tendency of nodes to be connected to other nodes of the same group—seems to be a key ingredient to generate echo chambers⁵ and perception biases⁷.

Among the phenomena associated with information transmission, echo chambers are presently the subject of intensive research, but they are only an aspect of a broader subject: the “information transmission biases” (in short, IT biases), which include all the possible alterations in the transmission of information that appear when changing the nature of the nodes that generate and receive such information. In addition to echo chambers, examples of IT biases include: (a) an enhanced/inhibited emission of information by a certain group (for example, how female opinions were wronged and overheard based on gender stereotypes¹²), and (b) an enhanced/inhibited reception of information by a certain group. These biases in information transmission have been observed in real-world networks and are influenced by their structural properties¹³. Let us remark that some studies regarding IT biases focus on how distinct types of information have different transmission probability^{14,15}, while here we will focus on the differences in the transmission induced by intrinsic properties of the node that emits and/or receives the information in the network.

Modeling processes of information transmission first requires the choice of a dynamical model. Often, the spreading of information is assumed to follow the same laws as the spreading of diseases. Because of this, epidemic models (also called *Simple Contagion* models)¹⁶ have been used for discussing the transmission of information. However, spreading of information, adoption of innovations, etc. are examples of social contagion phenomena in which individuals often require multiple exposures to a given piece of information to adopt it¹⁷. These social mechanisms are included in models of *Complex Contagion*^{18–22} inspired in the Threshold Model by Granovetter^{23,24} in which adoption requires a threshold number of neighboring agents that have already adopted it. In this sense, Complex Contagion, at variance with Simple Contagion, is a nonlinear process that requires group or many-agent interactions. Still, these group interactions are built from a combination of pairwise interactions, while possible higher-order many-agent interactions would call for a different approach^{25–28}. Several works have addressed the question of the validation of Complex Contagion models against experimental data^{29–31}, as well as the comparison of Simple and Complex Contagion models in this empirical context^{32–34}. However, there

IFISC (UIB-CSIC), Institute for Cross-Disciplinary Physics and Complex Systems, Campus Universitat de les Illes Balears, 07122 Palma de Mallorca, Spain. ✉email: sandro@ifisc.uib-csic.es

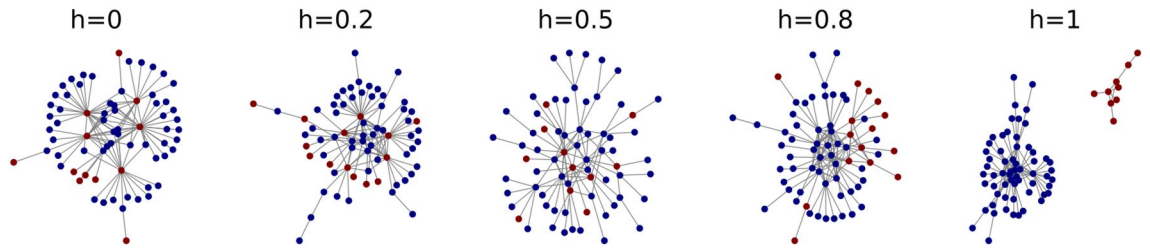


Figure 1. Topology of the networks generated by the BAh model, for different values of the homophily parameter h . Blue nodes indicate a majority group, while red nodes indicate a minority group. Other parameters: number of nodes $N = 70$, BAh parameter $m = 1$, minority fraction $f_a = 0.2$.

is also empirical evidence that many processes of information transmission involve both Simple and Complex Contagion, with some agents adopting in a single interaction and others requiring multiple exposures^{35,36}. As a consequence, different models of *Hybrid Contagion* combining Simple and Complex Contagion have been proposed^{36–40}. In this paper, we will compare predictions of Simple, Complex, and Hybrid Contagion models concerning IT biases.

A second ingredient in the modeling of information transmission is the choice of an underlying social network. In this work, we do not construct networks that would a priori lead to IT biases (for example due to a community structure). Instead, we focus on how echo chambers and other IT biases can emerge from the interplay between the structural properties of the network and the dynamics of the information. An important issue in this problem is disentangling the effects of social influence, included in the dynamical models, and homophily. To this end, we choose here to consider Barabasi–Albert⁴¹ networks with tunable homophily, following the model proposed in⁴² for social networks.

In this paper, we provide mathematical tools that can be used to define IT biases and to find out in which parameter regimes those IT biases exist. We show that the mechanism of *Hybrid Contagion* leads to the three IT biases mentioned before: emissivity bias, receptivity bias, and echo chambers. Importantly, the echo chamber bias, which is not present in neither Simple nor Complex Contagion, arises for a wide range of homophily parameters for Hybrid Contagion. Moreover, simulations in an empirical network confirm the presence of echo chambers and the other biases.

Results

Homophilic and heterophilic networks. To accurately model the structure of social networks, we use the Barabasi–Albert-homophily model (BAh model) proposed by Karimi et al.^{7,42}. This model is a generalization of the classical Barabasi–Albert (BA) model⁴¹ for scale-free networks. In this model, each node has a binary attribute, called the *group* of the node. We distinguish a *majority* and a *minority* group. The fraction of minority nodes in the network, i.e., the probability that a newly introduced node belongs to the minority group, is denoted by $f_a (< 0.5)$. As in the classical BA model, one node i is added to the network and connected to m existing nodes in each time step (throughout the work we set $m = 1$). Unlike it, however, the probability of attachment with an existing node j , Π_{ij} , depends not only on the degree of the existing node, k_j , but also on a *homophily parameter* h :

$$\Pi_{ij} \propto \begin{cases} hk_j & \text{if Group}(i) = \text{Group}(j) \\ (1-h)k_j & \text{if Group}(i) \neq \text{Group}(j) \end{cases} \quad (1)$$

From Eq. (1), it is clear that setting $h > 0.5$ ($h < 0.5$) generates homophilic (heterophilic) networks. For $h = 0.5$, one recovers the standard Barabasi–Albert model; while for $h = 1$, the probability of connecting two nodes of different groups is zero; thus, the network fragments into two components corresponding to the two groups. Figure 1 shows examples of networks generated using this model, for different values of the homophily parameter h . A first insight we can take is that in the heterophilic regime (i.e., $h < 0.5$), the hubs are minority nodes, while in the homophilic regime (i.e., $h > 0.5$) they belong to the majority group. Additionally, for $h < 0.5$ the average degree of the hubs is higher than for $h > 0.5$. This happens because, in heterophilic networks, the abundant majority nodes preferably attach to the rare minority ones, leading to larger degrees. In contrast, in homophilic networks, the probability of choosing one particular majority node to make a connection is smaller, due to their higher abundance. As will be shown below, this disparity in the hubs' sizes influences the information transmission on the different networks.

Simple vs complex contagion. *Definitions and description of the models.* We want to study the differences in information transmission that arise when considering different emitters and receivers of information. We call *source node* the node that initiates the information transmission (i.e., the seed of the transmission process). We also define the group to which it belongs as the *source group*. If the source node belongs to the majority, we talk about *majority source*, otherwise we call it *minority source*.

To quantify the spreading of information between groups, we define four information transmission observables (in short, *IT observables*) denoted by IT_{ab} . Each one represents the probability of successfully transmitting information from a source of a given group a to a target of a given group b , where a and b can be minority (m) or majority (M). For example, the probability of information transmission from a minority source to a minority

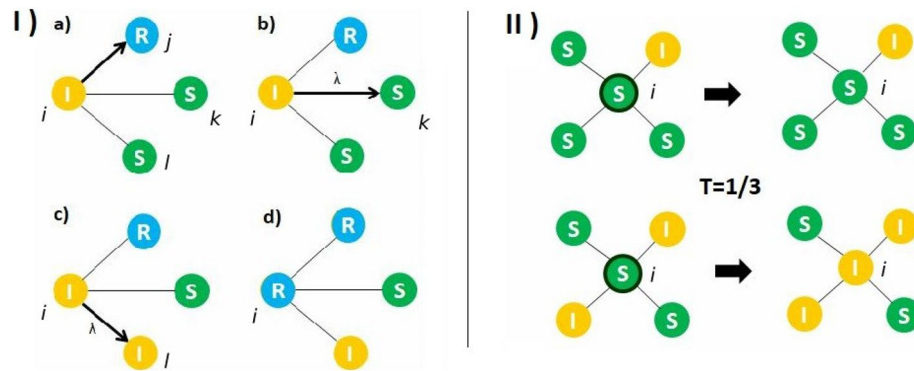


Figure 2. (I) Simple contagion: each informed node (I) tries to inform all of its susceptible (S) neighbors before becoming recovered (R). (a) Node i tries to inform node j , but fails because j is in a recovered state. (b) Then, node i tries to inform node k , with a probability of success λ , failing in this case. (c) Finally, node i succeeds in informing node l . (d) After having tried to inform all of its neighbors, node i becomes recovered (R). (II) Complex Contagion (with threshold $T = \frac{1}{3}$). Upper panels: node i has only a fraction $\frac{1}{4}$ of informed neighbors, so it remains susceptible. Lower panels: node i has a fraction $\frac{1}{2}$ of informed nodes, so the threshold is surpassed and therefore it becomes informed.

target is represented by IT_{mm} , the probability from a minority source to a majority target is represented by IT_{mM} , and so on. In addition, please note that calculating one IT_{ab} is equivalent on average to finding the final density of informed nodes belonging to group b , when the seed of the contagion belongs to group a . The details of this equivalence and other aspects of the simulation procedure can be found in the “Methods” section.

To start modeling information transmission (IT) on homophilic and heterophilic networks, we first consider a Simple Contagion model used by Karimi et al.⁴² This is a modification of the SIR dynamics¹⁶ in which each node can be in one of the following three possible states: susceptible, informed (also called adopter), and recovered. In this context, “recovered” refers to nodes that know the information, but choose not to spread it. When a node becomes informed, it tries to propagate the information to each of its neighbors only once. In each of these trials, recovered nodes cannot become informed again, while susceptible nodes become informed with probability λ . This parameter λ is called the *infectivity* of the contagion process. Once an informed node has tried to inform all of its neighbors, it automatically becomes a recovered node. The main difference with the standard SIR model is that, in the considered model, each link has only one opportunity to transmit the information, while in the standard SIR model each node will continue to transmit the information until it transitions into the recovered state. Sketches of the transitions can be found in the left panels of Fig. 2.

We will compare the results of the Simple Contagion model with processes of Complex Contagion^{18,19}. For Complex Contagion we use the threshold model proposed by Granovetter²³ and later studied by Watts²⁴. In this model, in each time step, one node of the network is randomly selected. If the node is informed, it remains in that state. If the node is susceptible, it changes state if the fraction of informed nodes in its neighborhood exceeds a value T , called the *threshold* of the Complex Contagion. An illustration of this process can be found in the right panel of Fig. 2.

Complex Contagion gives rise to cascade processes that manifest themselves in first-order transitions, i.e., a discontinuous jump in the global maximum of the probability density function (pdf) of the density of informed nodes. In contrast, a Simple Contagion model exhibits a second-order transition, i.e., the maximum of the pdf changes continuously. Notice that the control parameters of both models have different meanings and opposite behavior: in Complex Contagion, a larger threshold T inhibits contagion; while in Simple Contagion, a larger infectivity λ enhances it.

Capturing the impact of homophily in information spreading by using Complex Contagion. Let us use the defined IT observables to compare Simple and Complex Contagion. In Fig. 3, we show how the change of the homophily parameter h affects information transmission. For Simple Contagion, the dependency of our four observables on the homophily parameter is very weak (panel a). The strongest dependence appears for $h = 1$, but it is trivial because the network is fragmented and there is no path connecting different groups. When zooming in (insets of panel a), slight dependencies in both the homophily parameter and the source and target groups can be noticed. These differences were discussed by Karimi et al.⁴². Changing the degrees of the source and target nodes does not cause novel behaviors, so they can be disregarded as control parameters (not shown). Importantly, this simplified model does not account for the observed strong biases in the presence of homophily^{4,7}.

On the other hand, for Complex Contagion, we observe a stronger dependency on the homophily parameter h (Fig. 3, panel b). In fact, we find a completely different behavior in the homophilic and heterophilic regimes. As long as $h < 0.5$, increasing h improves information transmission. This is caused by the larger degree of the heterophilic hubs (Fig. 1), which prevents Complex Contagion^{18,24}. On the other hand, for $h > 0.5$, the effect of homophily is much less pronounced, because the change in the hubs’ degree is much milder in the homophilic regime. Additionally, increasing the threshold decreases all IT observables, although the decrease is more

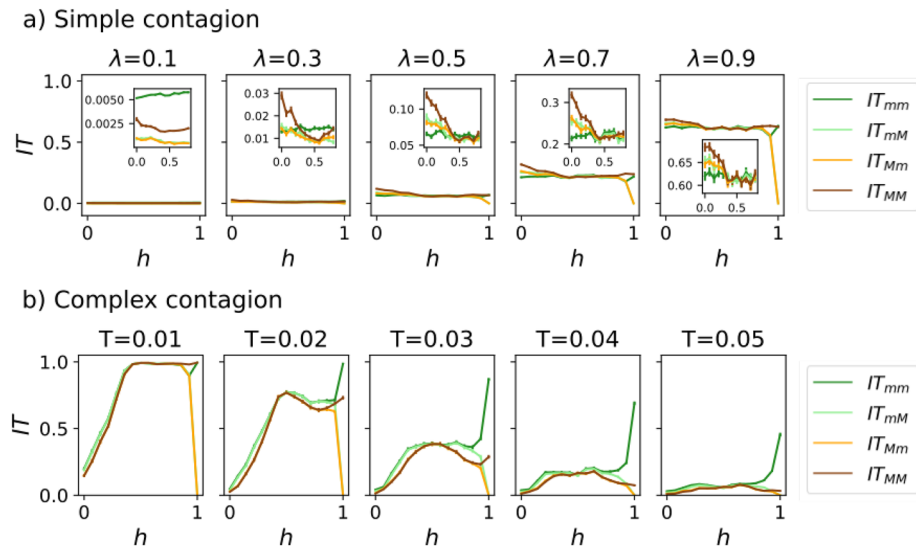


Figure 3. Information transmission observables as a function of the homophily parameter h . The chosen dynamics are Simple Contagion in panel (a) and Complex Contagion in panel (b). Each column shows a different infectivity λ or threshold T . The inset in panel a depicts the same IT observables as a function of the homophily parameter, but with a rescaled y-axis, so that variations among the curves are visible. To improve clarity, we have removed the values corresponding to $h > 0.85$ from the inset. In the legend, the subindexes indicate source and target groups respectively (for example, IT_{mM} indicates that the source group is the minority and the target group is the majority). Parameters used: $N = 1000$ nodes, $m = 1$, minority fraction $f_a = 0.2$, $M = 1000$ realizations (varying both the network structure and the location of the seed).

pronounced when $h > 0.5$. There are also slight dependencies on the source group but not on the target group (except for $h \approx 1$), suggesting some kind of bias in the transmission process.

Critical threshold for Complex Contagion. After analyzing the role of the homophily parameter, a question naturally arises: does h affect the nature of the transition? In other words, will we find a first-order transition with an “all-or-nothing” outcome, so that either every node or only a negligible fraction becomes informed?

To answer this question, we study how the probability density function (pdf) of the density of informed nodes changes when varying the threshold T , for two values of h (Fig. 4). The histograms confirm the existence of a first-order transition in both cases: the pdf shows two maxima at $\rho_f = 0$ and $\rho_f = 1$, corresponding to no contagion and full contagion, respectively.

The histograms allow identifying the critical threshold T_c , which is defined as the point where the global maximum changes from $\rho_f = 0$ to $\rho_f = 1$. Calculating the critical threshold for varying levels of homophily, we obtain the plot from Fig. 4b, where nontrivial dependencies on h arise. In the homophilic regime ($h > 0.5$), T_c is almost constant, while in the heterophilic regime ($h < 0.5$), T_c decays as the network becomes more heterophilic ($h \rightarrow 0$). This means that information propagates less efficiently in heterophilic networks, in agreement with Fig. 3b. Moreover, for $h = 1$, the threshold for minority sources increases significantly. This is a consequence of the fragmentation of the network: the minority group forms a separate network where the hubs have a smaller degree, thus favoring contagion. Finally, the critical threshold is slightly bigger for minority sources, suggesting again a bias in the information propagation. We conclude that the order of the transition is maintained, but the location of the transition point is strongly affected by h .

We have also analyzed the dependency of the critical threshold on the number of nodes. Figure 4b shows that the size of the network changes the height of the curve, but not its shape. This implies that the dependency on the homophily parameter is robust with respect to size. Furthermore, in Fig. 4c we also observe that, regardless of the homophily, T_c tends to zero as $N \rightarrow \infty$. This is a general phenomenon in scale-free networks²⁴, because the higher degree of the hubs inhibits the contagion process.

Hybrid Contagion. Even though for Complex Contagion we found that h strongly affects IT , the dependency on the source and target groups was minimal and lacked the strong biases -like echo chambers⁴- found in real-world networks. This discrepancy between model and data can be attributed to the “all-or-nothing” behavior of the system. Motivated by this, we propose a *Hybrid Contagion* model (HC). In this new version, Simple Contagion represents a viral transmission of information among nodes with sympathy towards the source of the information—for example, spreading of information is easier between individuals belonging to the same group-, whereas Complex Contagion models skeptical individuals that require multiple exposures to become convinced—e.g. when information comes from an unreliable source. This creates a scenario where one group has more difficulties when trying to convince members of the opposite group, in alignment with situations in real-world social networks³⁵.

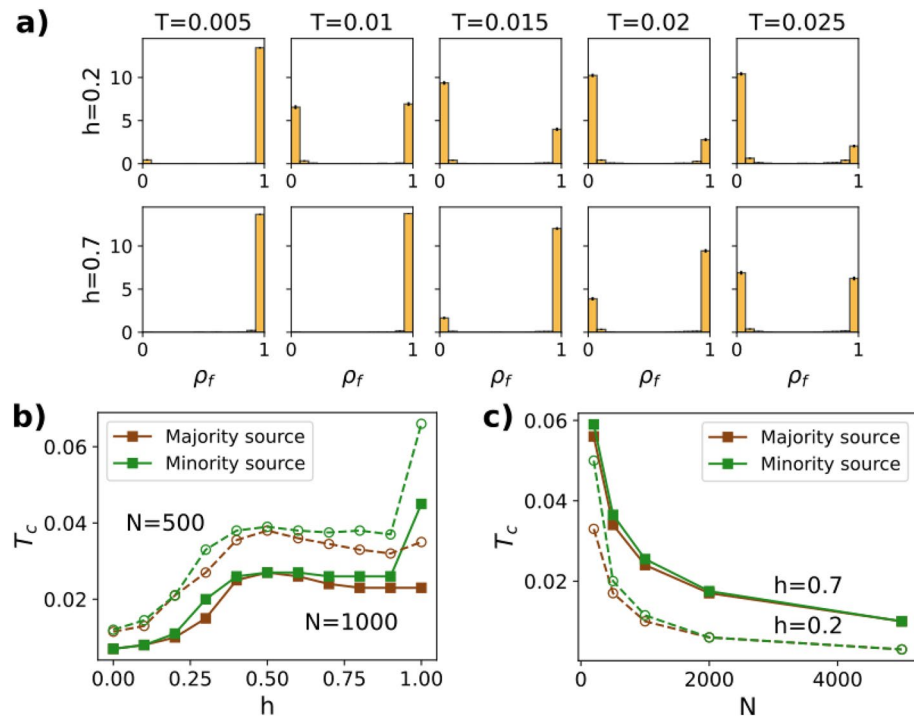


Figure 4. Critical threshold for Complex Contagion dynamics. **(a)** Histograms with the probability distribution of the final density of informed nodes, for a majority source and $h = 0.2$ (upper row) and $h = 0.7$ (lower row). The pdf is normalized so that the total area of each histogram equals one. Used parameters: $N = 1000$ nodes, $m = 1$, $M = 1000$ realizations. **(b)** Critical threshold as a function of the homophily parameter, for minority and majority sources. Simulations were performed for two values of N : $N = 500$ (dashed line) and $N = 1000$ (solid line). **(c)** Critical threshold as a function of the number of nodes, for minority and majority sources and two values of the homophily parameter: $h = 0.2$ (dashed line) and $h = 0.7$ (solid line). In all simulations, the minority size was fixed to $f_a = 0.2$.

Description of the model. Hybrid Contagion is the combination of Simple and Complex spreading into a unified model. In this way, nodes on the network can follow Simple or Complex Contagion depending on their group and the source of information. Nodes following Complex Contagion (called *complex nodes*) become informed when they have a fraction of informed neighbors bigger than the threshold, T . Nodes following Simple Contagion (*simple nodes*) can become informed with just one informed neighbor, with probability λ .

To simplify the model, we will make the following three assumptions. Firstly, we assign to each group a contagion type.; i.e., all majority nodes are simple and all minority nodes are complex, or vice versa. Secondly, we assume that the group that initiates the contagion process is always the group that follows Simple Contagion. This is justified by the fact that, usually, ideas generated inside a community spread more easily among the members of the community with the respect to other members that do not share the same views. Finally, for the same reason, we assume that the simple nodes always become informed whenever they have at least an informed neighbor; this is, we set the infectivity equal to one. $\lambda = 1$. With these simplifications, the only remaining control parameter of this model is the threshold T .

Information transmission with Hybrid Contagion. To investigate the predictions of this new model, we have calculated again the information transmission for all combinations of source and target (Fig. 5). This time, we observe not only a strong dependency on h , but also on the source and target groups (the latter only in the homophilic regime). In fact, the two possible source groups show opposite tendencies: IT from a minority source is enhanced in heterophilic networks, while IT from a majority source is enhanced in homophilic networks. These two opposite effects can be simultaneously explained in terms of one single observation: IT in Hybrid Contagion is favored when the network hubs follow Simple Contagion. The effects of the threshold T are also different depending on the source: increasing T causes a strong drop in the IT with a minority source, but the change is much weaker with a majority source. It is remarkable that, although T is a parameter that only controls the behavior of complex nodes, the information transmission to simple nodes is affected too; in fact, in the heterophilic regime, it is irrelevant whether the target node is simple or complex. This can be attributed to the high bipartivity of the network (most routes connecting two simple nodes cross a complex node).

Additionally, the source and target dependencies imply strong biases in the information transmission. In particular, IT from minority to majority becomes negligible for $h > 0.5$: we have a lack of IT even when the network is connected. On the other hand, information transmission from majority to minority is nonzero. This

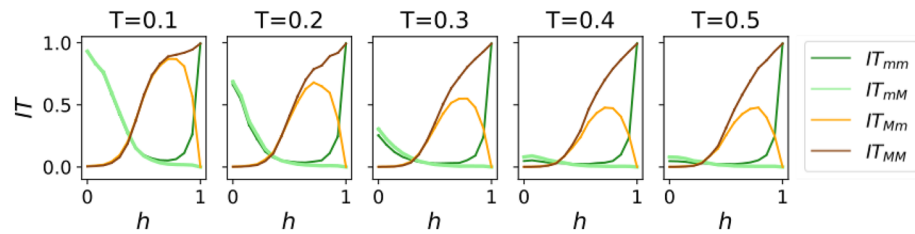


Figure 5. Information transmission observables as a function of the homophily parameter h , for Hybrid Contagion dynamics. Each column shows a different threshold T . The first subindex of each element of the legend indicates the source group, while the second subindex indicates the target group. Parameters used: $N = 1000$ nodes, $m = 1$, minority fraction $f_a = 0.2$, $M = 1000$ realizations.

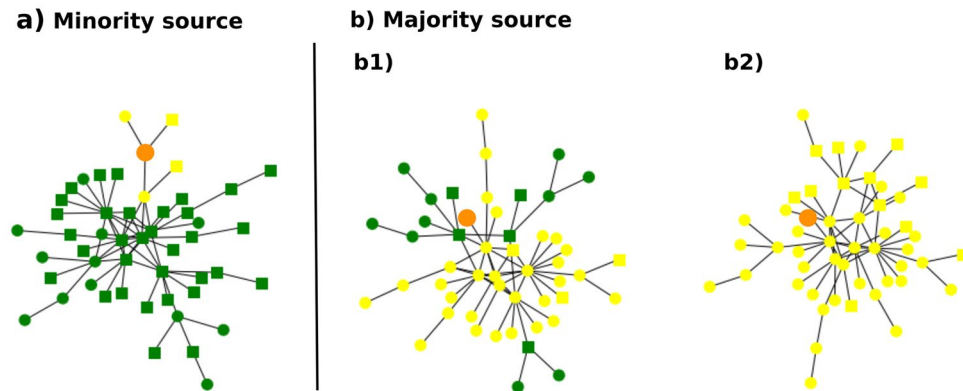


Figure 6. Examples of Hybrid Contagion on BA_h networks. Each network shows the final state of one simulation, where green nodes are susceptible and yellow ones are informed. The orange node is the source of the information. Simple nodes are denoted by circles and complex nodes by squares. (a) When the source node belongs to the minority, Hybrid Contagion is limited. (b) However, when the source node belongs to the majority, a wide range of cascade sizes is obtained for the same parameters. Parameters used: $N = 50$, $h = 0.7$, $f_a = 0.2$, $T = 0.2$.

asymmetry suggests the existence of some phenomenon similar to an echo chamber, which will be the main topic of the following sections.

Critical threshold for Hybrid Contagion. As happened for Complex Contagion, a system following Hybrid Contagion shows a rich dynamics that is hard to understand by looking only at the IT. A simplified version of these rich dynamics is shown in Fig. 6, for a threshold $T = 0.2$. One sees that a minority source produces a small cascade (panel a), consisting mainly of nodes of the same group. On the other hand, a majority source produces cascades with a wide size range, from 70% (panel b1) to 100% (panel b2) of the nodes. In some cases, clusters of simple nodes get shielded from contagion due to the presence of a complex node.

To investigate this behavior in more depth and to study the nature of the transition, we calculate the pdf of the density of informed nodes for different values of the threshold T (Fig. 7, panel a). As opposed to Complex Contagion, where differences between majority and minority source were minimal, now we see two different behaviors depending on the source group. In particular, for a minority source (upper panels of Fig. 7a), we obtain the “all-or-nothing” behavior typical of Complex Contagion. This is not the case when the source belongs to the majority, with a different form of the pdf that reflects the different behaviors of Fig. 6b. Moreover, the pdf does not change in shape for thresholds above $T = 0.2$ (compare, for example, the histograms corresponding to $T = 0.2$ and $T = 0.9$ in the lower panels of Fig. 7a). Nevertheless, both histograms show a discontinuous jump in the global maximum. All these phenomena could be indicators of a hybrid transition^{43,44}, where one of the maxima of the pdf varies continuously but the global maximum changes discontinuously at a certain threshold value T_c .

When we plot the change of the critical threshold T_c as a function of h (Fig. 7b), we observe opposite tendencies depending on the source groups: for minority sources, T_c decreases with h , while for majority sources, it increases. More interestingly, for a majority source, T_c diverges around $h = 0.63$. Above this value, only the phase corresponding to information transmission exists. In short, for a majority source and sufficiently homophilic networks, the information is always able to propagate through the network, regardless of the threshold.

Echo chambers and other IT biases. *Measuring echo chambers and IT biases.* As discussed in the previous section, the Hybrid Contagion model shows a rich behavior that cannot be easily understood by simply

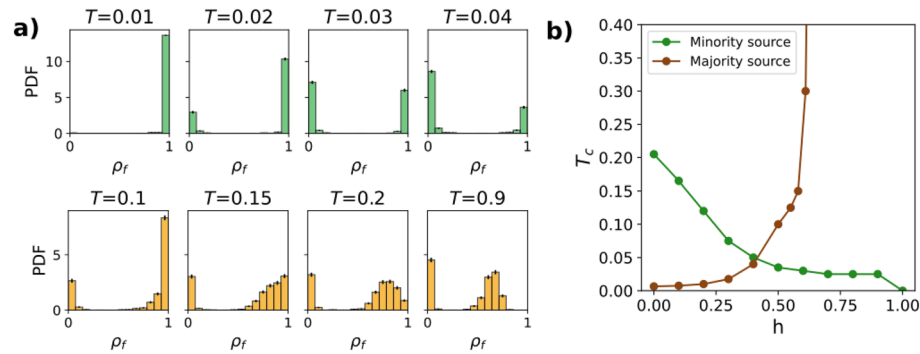


Figure 7. Critical threshold for Hybrid Contagion dynamics. **(a)** Histograms with the probability distribution of the final density of informed nodes, for a minority source (green histograms) and a majority source (orange histograms). The pdf is normalized so that the total area of each histogram equals one. The homophily parameter is $h = 0.6$. **(b)** Critical threshold as a function of the homophily parameter, for minority and majority sources. Other parameters used: $N = 1000$ nodes, $m = 1$, $M = 1000$ realizations, $f_a = 0.2$.

measuring the four IT observables. In particular, Hybrid Contagion exhibits *information transmission biases* (IT biases), which can be defined as dependencies of information transmission on the source and target groups. To be able to classify these biases, as well as to determine their strength, we propose the following set of *bias variables*:

$$\overline{IT} = \frac{1}{4} (+IT_{mm} + IT_{mM} + IT_{Mm} + IT_{MM}) \tag{2a}$$

$$B_E = \frac{1}{2} (-IT_{mm} - IT_{mM} + IT_{Mm} + IT_{MM}) \tag{2b}$$

$$B_R = \frac{1}{2} (-IT_{mm} + IT_{mM} - IT_{Mm} + IT_{MM}) \tag{2c}$$

$$B_{EC} = \frac{1}{2} (+IT_{mm} - IT_{mM} - IT_{Mm} + IT_{MM}) \tag{2d}$$

Notice that, since $0 < IT_{ab} < 1$, we have $0 < \overline{IT} < 1$ and $-1 < B_E, B_R, B_{EC} < 1$.

The previously used IT observables focused on specific groups, and as a consequence, they lacked information about the information propagation on the network as a whole. In contrast, the bias variables are able to distinguish global aspects of the transmission process, e.g. the receptivity of the nodes towards certain types of information. The first one, \overline{IT} , is the *mean information transmission* over the four possible combinations of source and target. It is high whenever the network transmits information effectively regardless of who emits and who receives the information. Secondly, B_E is the *emissivity bias*. If it is bigger than zero, it indicates that the information that starts in the majority propagates more effectively than starting in the minority, regardless of the target. On the contrary, $B_E < 0$ indicates that information starting in the *minority* propagates more effectively. Thirdly, B_R is the *receptivity bias*. $B_R > 0 (< 0)$ indicates that the majority has a bigger (smaller) probability of receiving information than the minority. Finally, we define B_{EC} as the *echo chamber bias*. It estimates whether information propagates only between nodes of the same group or not. If it is close to one, information only propagates inside of each group; while if it is negative, it indicates the existence of an “anti echo chamber”: information only flows between nodes of the opposite groups.

Armed with these new variables, we can easily determine the biases of an information transmission process. One just needs to measure the IT parameters IT_{ab} from a given data set and substitute them into Eq. (2). Equation (given data set and substit) also gives an objective definition of an echo chamber: a social network shows an echo chamber if the bias B_{EC} of an information transmission process is larger than zero. Finally, one should also note that *any* bias in the information transmission between two groups can be expressed as a linear combination of the three biases (plus a possible contribution from \overline{IT}). In other words, any bias can be decomposed into a mixture of B_E, B_R and B_{EC} .

As an example of the usefulness of this new framework, we will analyze the results of Simple and Complex Contagion—already analyzed in Fig. 3—in terms of these variables. As expected, Simple Contagion presents negligible biases B_E, B_R, B_{EC} (not shown). A similar picture arises for Complex Contagion (Fig. 8): only \overline{IT} shows an important dependency on h ; however, its behavior is the same that we saw in Fig. 3. Additionally, the other three bias variables show an almost neutral behavior. Only significantly high values of h lead to an increase in B_{EC} and $|B_R|$ until $h = 1$, where they reach their maximum due to the fragmentation of the network. These last results draw a clear picture: the transmission of information under Complex Contagion is sensitive to the homophily parameter and may be greatly inhibited, but it does not show strong biases.

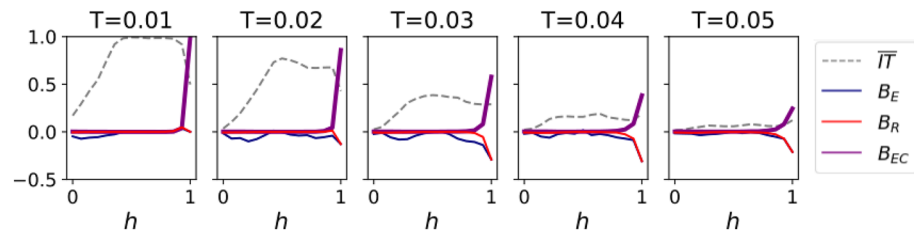


Figure 8. Bias variables \overline{IT} , B_E , B_R and B_{EC} as a function of the homophily parameter h , for Complex Contagion dynamics. The simulation parameters are the same as in Fig. 3b.

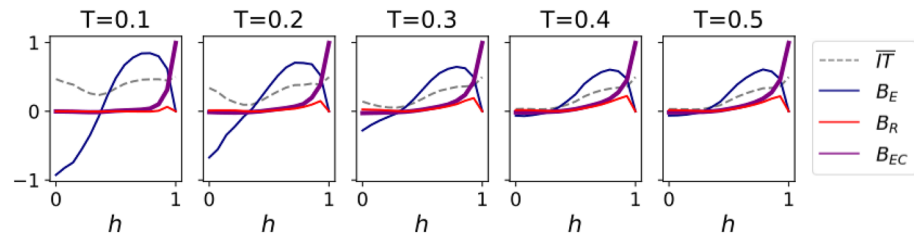


Figure 9. Bias variables \overline{IT} , B_E , B_R and B_{EC} as a function of the homophily parameter h , for Hybrid Contagion dynamics. The simulations and parameters are the same as in Fig. 5.

Echo chambers and other biases in Hybrid Contagion. Unlike Complex Contagion, Hybrid Contagion shows important dependencies on both the source and target groups. The most relevant was the impossibility of transmitting information from minority to majority for $h > 0.5$. To quantify these dependencies, we plot the four bias variables for Hybrid Contagion in Fig. 9.

Importantly, \overline{IT} is smaller than one and practically constant with respect to T and h , although it is minimum in the heterophilic regime. Additionally, the system has a strong emissivity bias, negative in the heterophilic regime and positive in the homophilic one. Increasing the threshold reduces the absolute value of the emissivity in the heterophilic regime, with a limited effect in the homophilic one. The receptivity bias becomes also appreciable for $h > 0.9$ and large thresholds. However, the most relevant observation is that the echo chamber variable is different from zero throughout *all* the homophilic regime ($h > 0.5$), especially for higher thresholds and for high homophily parameters.

Summing up, the combination of Simple and Complex Contagion (Hybrid Contagion) leads not only to strong emissivity biases for a wide range of h , but also to the emergence of echo chambers in the homophilic regime. Nevertheless, the average information transmission remains almost constant and is minimum in the heterophilic regime. This highlights the independence between the quantity and the lack of biases of transferred information: neither the presence of echo chambers and other biases implies that information transmission is hindered, nor does a strong transmission of information guarantee that the information is bias-free.

Echo chambers in real-world networks. To put our results on more solid ground, we test our results on a real collaboration network. We simulate Hybrid Contagion in a network of scientific citations between papers published in the journals of the American Physical Society (Fig. 10). This network was already used to study perception biases in⁷. In particular, we focus on papers devoted to statistical mechanics, where majority and minority groups have been identified with Quantum Statistical Mechanics and Classical Statistical Mechanics, respectively. The network is highly homophilic, with a homophily parameter $h = 0.92$.

Repeating our analysis regarding the IT observables and the bias variables, we obtain results that are fully consistent with the prediction in synthetic networks (Fig. 5). Indeed, the four IT observables show the same behavior as the synthetic curves for high homophily parameters ($IT_{MM} > IT_{mm} > IT_{Mm} > IT_{mM}$ in panel 10a. Surprisingly, we find high values of the echo chamber bias ($B_{EC} \approx 0.5$), even higher than in synthetic networks with the same parameter h (Fig. 10b). A possible explanation is that real-world networks have a higher clustering than our synthetic networks. This could favor intragroup contagion and enhance the echo chamber bias.

The remaining bias variables are also in agreement with our predictions in synthetic networks: the emissivity is positive and information transmission remains high even for large thresholds. In summary, our results contribute to the understanding of the emergence of echo chamber biases in information transmission between different groups.

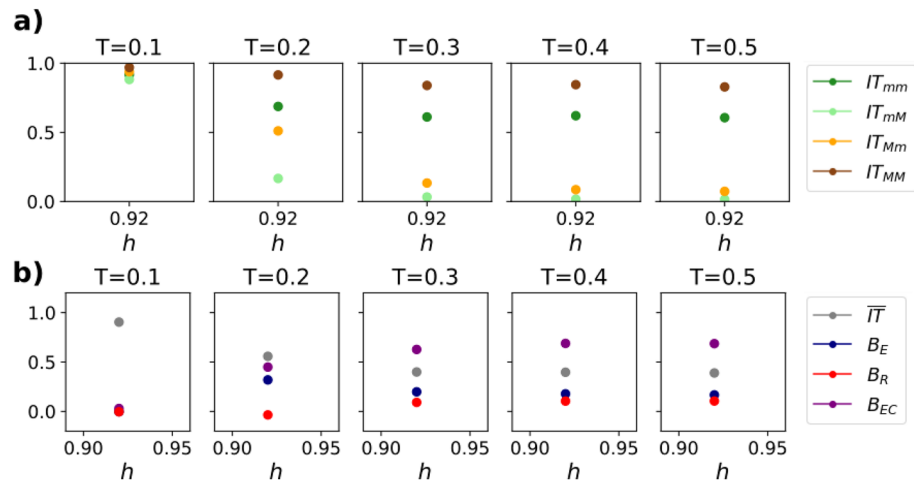


Figure 10. (a) Information transmission observables for Hybrid Contagion dynamics in a network of scientific citations between papers of the APS, for different thresholds T . The legend follows the same conventions as Fig. 5. Parameters of the network: $h = 0.92$, $f_a = 0.3177$. (b) Bias variables corresponding to the IT observables of panel (a).

Discussion

In this work, we have explored how information transmission (IT) in homophilic/heterophilic scale-free networks can be modeled, focusing on alterations of information transmission such as the emergence of echo chambers. To achieve this, we have analyzed three different models and proposed a decomposition of information transmission that allowed a straightforward quantification of the presence of biases. Starting from a structural model able to generate networks with a tunable level of homophily⁴², we analyzed Simple Contagion by employing a slightly modified version of the SIR dynamics; Complex Contagion, and finally a Hybrid model in which the spreading between two nodes changes depending on whether they belong to the same or different groups.

Our main conclusion is that the choice of the dynamical model greatly influences both the average information transmission and the emerging biases. In particular, we find that Simple Contagion leads to negligible biases and a minimal dependency on the homophily parameter h , whereas Complex Contagion shows strong dependencies on h , with high average information transmission in the homophilic regime. These differences are also reflected in the Complex Contagion transition between the informed and uninformed regimes. In particular, the threshold for the transition is higher in the homophilic regime, leading to an enhancement of the information transmission. The strong variability of the mean IT when changing h is not correlated with strong biases in information transmission. In fact, most biases only appear as the network becomes disconnected, except for a slight emissivity bias for a wider range of h .

A richer phenomenology is found in the Hybrid Contagion model, in which information transmission has an important dependency on h and follows opposite trends when changing the source of the information. In particular, information originating in a minority node spreads easily in the heterophilic regime, while in the homophilic regime the transmission is dominated by information originating in the majority. Thus, interchanging the source group affects not only the behavior of the transition—with a divergence of the majority-source critical threshold at $h \approx 0.63$, but also the qualitative behavior of the pdf, with hints of a possible hybrid transition. As opposed to Complex Contagion, the dependencies on the source and target groups cause the appearance of stronger IT biases, not only in the emissivity and receptivity of information but also in the emergence of echo chambers, where information starting in the minority fails to reach the majority groups. The echo chamber bias is different from zero for any $h > 0.5$, which implies that intragroup communication is favored and intergroup communication is hindered in all the homophilic regime. Moreover, even though for $h < 1$ the network is still connected, the echo chamber bias reaches a significant value ($B_{EC} \approx 0.5$). When analyzing a citation network between papers of the APS, we find even stronger biases, highlighting the relevance of our results for real-world scenarios.

On a more general note, our study points out three important factors when analyzing information transmission. Firstly, that homophily and heterophily play a key role in how well information is transmitted and which biases appear. Secondly, the quantity of transmitted information is not necessarily correlated with lack of biases: our analysis showed that models with low average information transition can be free from biases (like Complex Contagion for high threshold parameters), whereas models with high mean information transmission can show strong biases (Hybrid Contagion). Thirdly, biases in information transmission are not limited to echo chambers. Other biases (such as different levels of emissivity) can play a comparable role and affect the transmission of knowledge in our society^{10,12}. In summary, we believe that our decomposition of information transmission into an average value plus three distinct biases can help clarify complicated information transmission patterns in real data.

All the presented models show an important limitation: our group category is binary. In general, splitting society's complexity into just two groups is too reductive, one clear example being the aggregation of political viewpoints into "left" and "right" groups. In this sense, a generalization of the BAh model to a continuum of groups would be helpful to better understand how individuals can gradually transition between groups, and may cause unexpected behaviors in the bias magnitudes. Another variation of these models could also incorporate a coevolving network in which nodes rewire to maximize the number of neighbors with the same opinion. This rewiring is known to lead to network polarization, echo chambers, and ultimately fragmentation^{8,45,46}, but its influence on the other biases is unknown. Finally, the generalization of the model to multidimensional topic spaces could help in the understanding of how ideologies form⁶.

In conclusion, we have shown that contagion models beyond Simple Contagion can exhibit information transmission biases, including echo chambers. We hope that this work provides awareness about information transmission biases and the simple mathematical tools able to quantify them, so that further research can better understand how they emerge and ultimately overcome them.

Methods

Measuring information transmission. As mentioned in the main text, the final density of informed nodes within a group g_t , when taking the seed node within a group g_s , coincides *on average* with the probability of information transmission from a source from g_s to a target from g_t . Here we present proof of this equivalence.

Let N_g be the number of nodes belonging to group g and let I_g the number of informed nodes of group g in the final (absorbing) state. We define the final density of informed nodes within group g as $\rho_{fg} = \frac{I_g}{N_g}$. On the other hand, the *average* probability that information spreads from a given source to a given target, P_{tr} , when choosing source and target nodes randomly within the source and target groups, is:

$$P_{tr} = \Pr(\text{target is informed} \mid \text{contagion starts at source}) = \frac{\text{No. targets that are informed}}{\text{No. of possible targets}} = \frac{I_g}{N_g} = \rho_{fg} \quad (3)$$

Thus, we have shown that measuring the information transmission observables can be reduced to measuring the final density of informed nodes.

In the following, we describe the exact simulation procedure to measure ρ_{fg} :

1. Set the source and target groups.
2. Generate a network with the Barabasi–Albert-homophily model.
3. Select a source node and a target node randomly, with the constraint that they belong to the source and target groups respectively.
4. Mark the source node as informed. Simulate a contagion process with the corresponding model (Simple, Complex, or Hybrid Contagion).
5. Once the absorbing state is reached, measure the number of nodes belonging to group g , N_g , and the number of informed nodes of group g , I_g , to get the final density of informed nodes of group g , ρ_{fg} .
6. Repeat steps 2–5 for several networks and source and target nodes and find the average final density of informed nodes ρ_{fg} .

Empirical network. To test the validity of our results in real networks, we performed a simulation in a network with high homophily: a network of scientific citations of the APS^{7,42}. The nodes in the network correspond to individual scientific papers related to statistical mechanics, and each link corresponds to a citation between two of them. We disregard the directional nature of the links, since the BAh model is designed for undirected networks. To ensure that contagion is possible, we only take the largest component and disregard all the small components. We select the minority and majority groups based on identifiers from the Physics and Astronomy Classification Scheme (PACS). In particular, the minority group corresponds to papers devoted to Classical Statistical Mechanics (CSM) and the majority group to Quantum Statistical Mechanics (QSM).

Taking these considerations into account, we obtain a network with 1281 nodes and 3064 links. From these nodes, 407 belonged to the minority group and 874 to the majority. The minority fraction is thus $f_a = 0.32$ and the homophily parameter is $h = 0.92$. The homophily parameter was estimated using the procedure described in⁴².

Data availability

The empirical citation network between papers published in APS journals has been made publicly available by the authors of^{7,42} and can be found online at <https://github.com/frbkrm/NtwPerceptionBias>.

Received: 17 March 2022; Accepted: 23 May 2022

Published online: 07 June 2022

References

1. Goel, S., Anderson, A., Hofman, J. & Watts, D. J. The structural virality of online diffusion. *Manage. Sci.* <https://doi.org/10.1287/mnsc.2015.2158> (2015).
2. Vosoughi, S., Roy, D. & Aral, S. The spread of true and false news online. *Science* **359**, 1146–1151. <https://doi.org/10.1126/science.aap9559> (2018).
3. Juul, J. L. & Ugander, J. Comparing information diffusion mechanisms by matching on cascade size. *Proc. Natl. Acad. Sci.* **118**, e2100786118. <https://doi.org/10.1073/pnas.2100786118> (2021).

4. Cinelli, M., De Francisci Morales, G., Galeazzi, A., Quattrociocchi, W. & Starnini, M. The echo chamber effect on social media. *Proc. Natl. Acad. Sci.* **118**, e2023301118. <https://doi.org/10.1073/pnas.2023301118> (2021).
5. Baumann, F., Lorenz-Spreen, P., Sokolov, I. M. & Starnini, M. Modeling echo chambers and polarization dynamics in social networks. *Phys. Rev. Lett.* **124**, 048301. <https://doi.org/10.1103/PhysRevLett.124.048301> (2020).
6. Baumann, F., Lorenz-Spreen, P., Sokolov, I. M. & Starnini, M. Emergence of polarized ideological opinions in multidimensional topic spaces. *Phys. Rev. X* **11**, 011012. <https://doi.org/10.1103/PhysRevX.11.011012> (2021).
7. Lee, E. *et al.* Homophily and minority-group size explain perception biases in social networks. *Nat. Hum. Behav.* **3**, 1078–1087. <https://doi.org/10.1038/s41562-019-0677-4> (2019).
8. Tokita, C. K., Guess, A. M. & Tarnita, C. E. Polarized information ecosystems can reorganize social networks via information cascades. *Proc. Natl. Acad. Sci.* **118**, e2102147118. <https://doi.org/10.1073/pnas.2102147118> (2021).
9. Dubois, E. & Blank, G. The echo chamber is overstated: The moderating effect of political interest and diverse media. *Inf. Commun. Soc.* **21**, 729–745. <https://doi.org/10.1080/1369118X.2018.1428656> (2018).
10. Barberá, P., Jost, J. T., Nagler, J., Tucker, J. A. & Bonneau, R. Tweeting from left to right: Is online political communication more than an echo chamber?. *Psychol. Sci.* **26**, 1531–1542. <https://doi.org/10.1177/0956797615594620> (2015).
11. Törnberg, P. Echo chambers and viral misinformation: Modeling fake news as complex contagion. *PLoS One* **13**, e0203958. <https://doi.org/10.1371/journal.pone.0203958> (2018).
12. Fricker, M. & Press, O. U. *Epistemic Injustice: Power and the Ethics of Knowing*. Oxford Scholarship Online (Clarendon Press, 2007).
13. Mesoudi, A., Whiten, A. & Dunbar, R. A bias for social information in human cultural transmission. *Br. J. Psychol.* **97**, 405–423. <https://doi.org/10.1348/000712605X85871> (2006).
14. Altshteyn, I. *Evidence for a warning bias in information transmission in social networks*. Master's thesis, University of California, Los Angeles (2014). ProQuest ID: Altshteyn_ucla_0031N_12675. Merritt ID: ark:/13030/m5n88t2w. Retrieved from <https://escholarship.org/uc/item/904684q4>.
15. Fay, N., Walker, B., Kashima, Y. & Perfors, A. Socially situated transmission: The bias to transmit negative information is moderated by the social context. *Cogn. Sci.* <https://doi.org/10.1111/cogs.13033> (2021).
16. Pastor-Satorras, R., Castellano, C., Van Mieghem, P. & Vespignani, A. Epidemic processes in complex networks. *Rev. Mod. Phys.* **87**, 925–979. <https://doi.org/10.1103/RevModPhys.87.925> (2015).
17. Chang, H.-C.H. & Fu, F. Co-diffusion of social contagions. *New J. Phys.* **20**, 095001. <https://doi.org/10.1088/1367-2630/aadce7> (2018).
18. Centola, D., Eguiluz, V. M. & Macy, M. W. Cascade dynamics of complex propagation. *Phys. A* **374**, 449–456. <https://doi.org/10.1016/j.physa.2006.06.018> (2007).
19. Centola, D. *How Behavior Spreads: The Science of Complex Contagions* (Princeton University Press, 2018).
20. Gleeson, J. P. Cascades on correlated and modular random networks. *Phys. Rev. E* **77**, 046117. <https://doi.org/10.1103/PhysRevE.77.046117> (2008).
21. Hackett, A., Melnik, S. & Gleeson, J. P. Cascades on a class of clustered random networks. *Phys. Rev. E* **83**, 056107. <https://doi.org/10.1103/PhysRevE.83.056107> (2011).
22. Oh, S.-W. & Porter, M. A. Complex contagions with timers. *Chaos Interdiscip. J. Nonlinear Sci.* **28**, 033101. <https://doi.org/10.1063/1.4990038> (2018).
23. Granovetter, M. Threshold models of collective behavior. *Am. J. Sociol.* **83**, 1420–1443 (1978).
24. Watts, D. J. A simple model of global cascades on random networks. *Proc. Natl. Acad. Sci.* **6**, 25 (2002).
25. Iacopini, I., Petri, G., Barrat, A. & Latora, V. Simplicial models of social contagion. *Nat. Commun.* **10**, 2485. <https://doi.org/10.1038/s41467-019-10431-6> (2019).
26. Battiston, F. *et al.* Networks beyond pairwise interactions: Structure and dynamics. *Phys. Rep.* **874**, 1–92. <https://doi.org/10.1016/j.physrep.2020.05.004> (2020).
27. Sahasrabudde, R., Neuhäuser, L. & Lambiotte, R. Modelling non-linear consensus dynamics on hypergraphs. *J. Phys. Complex.* **2**, 025006. <https://doi.org/10.1088/2632-072X/abcea3> (2021).
28. de Arruda, G. F., Petri, G., Rodriguez, P. M. & Moreno, Y. Multistability, intermittency and hybrid transitions in social contagion models on hypergraphs. [arXiv:2112.04273](https://arxiv.org/abs/2112.04273) (2021).
29. Centola, D. The spread of behavior in an online social network experiment. *Science* **329**, 1194–1197. <https://doi.org/10.1126/science.1185231> (2010).
30. Rosenthal, S. B., Twomey, C. R., Hartnett, A. T., Wu, H. S. & Couzin, I. D. Revealing the hidden networks of interaction in mobile animal groups allows prediction of complex behavioral contagion. *Proc. Natl. Acad. Sci.* **112**, 4690–4695. <https://doi.org/10.1073/pnas.1420068112> (2015).
31. Aral, S. & Nicolaides, C. Exercise contagion in a global social network. *Nat. Commun.* **8**, 14753. <https://doi.org/10.1038/ncomm14753> (2017).
32. Lerman, K. Information is not a virus, and other consequences of human cognitive limits. *Future Internet* **8**, 21. <https://doi.org/10.3390/fi8020021> (2016).
33. Monsted, B., Sapiezynski, P., Ferrara, E. & Lehmann, S. Evidence of complex contagion of information in social media: An experiment using Twitter bots. *PLoS One* **12**, e0184148. <https://doi.org/10.1371/journal.pone.0184148> (2017).
34. Guilbeault, D. & Centola, D. Topological measures for identifying and predicting the spread of complex contagions. *Nat. Commun.* **12**, 4430. <https://doi.org/10.1038/s41467-021-24704-6> (2021).
35. State, B. & Adamic, L. The diffusion of support in an online social movement: Evidence from the adoption of equal-sign profile pictures. In *Proceedings of the 18th ACM Conference on Computer Supported Cooperative Work & Social Computing*, 1741–1750. <https://doi.org/10.1145/2675133.2675290> (ACM, Vancouver BC Canada, 2015).
36. Karsai, M., Iñiguez, G., Kikas, R., Kaski, K. & Kertész, J. Local cascades induced global contagion: How heterogeneous thresholds, exogenous effects, and unconcerned behaviour govern online adoption spreading. *Sci. Rep.* **6**, 21718. <https://doi.org/10.1038/srep21718> (2016).
37. Dodds, P. S. & Watts, D. J. Universal behavior in a generalized model of contagion. *Phys. Rev. Lett.* **92**, 218701. <https://doi.org/10.1103/PhysRevLett.92.218701> (2004).
38. Czaplicka, A., Toral, R. & San Miguel, M. Competition of simple and complex adoption on interdependent networks. *Phys. Rev. E* **94**, 062301. <https://doi.org/10.1103/PhysRevE.94.062301> (2016).
39. Min, B. & San Miguel, M. Competing contagion processes: Complex contagion triggered by simple contagion. *Sci. Rep.* **8**, 10422. <https://doi.org/10.1038/s41598-018-28615-3> (2018).
40. Kook, J., Choi, J. & Min, B. Double transitions and hysteresis in heterogeneous contagion processes. *Phys. Rev. E* **104**, 044306. <https://doi.org/10.1103/PhysRevE.104.044306> (2021).
41. Barabasi, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 4 (1999).
42. Karimi, F., Génois, M., Wagner, C., Singer, P. & Strohmaier, M. Homophily influences ranking of minorities in social networks. *Sci. Rep.* **8**, 11077. <https://doi.org/10.1038/s41598-018-29405-7> (2018).
43. Baxter, G. J., Dorogovtsev, S. N., Goltsev, A. V. & Mendes, J. F. F. Heterogeneous k-core versus bootstrap percolation on complex networks. *Phys. Rev. E* **83**, 051134. <https://doi.org/10.1103/PhysRevE.83.051134> (2011).
44. Lee, D., Choi, S., Stippinger, M., Kertész, J. & Kahng, B. Hybrid phase transition into an absorbing state: Percolation and avalanches. *Phys. Rev. E* **93**, 042109. <https://doi.org/10.1103/PhysRevE.93.042109> (2016).

45. Vazquez, F., Eguiluz, V. M. & Miguel, M. S. Generic absorbing transition in coevolution dynamics. *Phys. Rev. Lett.* **100**, 108702. <https://doi.org/10.1103/PhysRevLett.100.108702> (2008).
46. Saeedian, M., San Miguel, M. & Toral, R. Absorbing-state transition in a coevolution model with node and link states in an adaptive network: Network fragmentation transition at criticality. *New J. Phys.* **22**, 113001. <https://doi.org/10.1088/1367-2630/abbfd0> (2020).

Acknowledgements

FDD, MSM and SM acknowledge support from the Spanish Agency of Research (AEI) through Maria de Maeztu Program for units of Excellence in R & D (Grant MDM-2017-0711 funded by MCIN/AEI/10.13039/501100011033). In particular, FDD thanks financial support MDM-2017-0711-20-2 and FSE invierte en tu futuro. SM acknowledges funding from the project PACSS RTI2018-093732-B-C22 of the MCIN/AEI/10.13039/501100011033/ and by EU through FEDER funds (A way to make Europe).

Author contributions

F.D.D., M.S.M. and S.M. conceived the models. F.D.D. performed the numerical simulations and analyzed the data. M.S.M. and S.M. supervised the work. All authors contributed in the writing and revision of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022