

# Mechanistic interplay between information spreading and opinion polarization

Kleber Andrade Oliveira  <sup>a,\*</sup>, Henrique Ferraz de Arruda  <sup>b,c,\*</sup> and Yamir Moreno  <sup>c,d</sup>

<sup>a</sup>Social Dynamics Research Lab, Department of Psychology, University of Limerick, Limerick V94 T9PX, Ireland

<sup>b</sup>ARAID Foundation, Av. de Ranillas 1-D, planta 2a, oficina B, Zaragoza 50018, Spain

<sup>c</sup>Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Mariano Esquillor, s/n, Zaragoza 50018, Spain

<sup>d</sup>Department of Theoretical Physics, Faculty of Sciences, University of Zaragoza, Pedro Cerbuna, 12, Zaragoza 50009, Spain

\*To whom correspondence should be addressed: Email: [kleber.oliveira@ul.ie](mailto:kleber.oliveira@ul.ie) (K. A. O.); Email: [henrique.ferraz@bifi.es](mailto:henrique.ferraz@bifi.es) (H. F. A.)

Edited By Attila Szolnoki

## Abstract

We investigate how information-spreading mechanisms affect opinion dynamics and vice versa via an agent-based simulation on adaptive social networks. First, we characterize the impact of reposting on user behavior with limited memory, a feature that introduces novel system states. Then, we build an experiment mimicking information-limiting environments seen on social media platforms and study how the model parameters can determine the configuration of opinions. In this scenario, different posting behaviors may sustain polarization or reverse it. We further show the adaptability of the model by calibrating it to reproduce the statistical organization of information cascades as seen empirically in a microblogging social media platform. Our model combines mechanisms for platform content recommendation, connection rewiring, and limited-attention user behavior, paving the way for a robust understanding of echo chambers as a specialized phenomenon of opinion polarization.

## Significance Statement

Polarization and information diffusion are often studied separately, yet in online environments they are deeply intertwined. We present an agent-based model that incorporates limited memory, reposting, and recommendation algorithms, capturing how real social media feeds shape both cascades of information and opinion dynamics. Small changes in posting or recommendation rules can flip a system from consensus to polarization. Moreover, the parameters that reproduce empirical cascade statistics from the Brexit and vaccine debates also yield polarized opinion states. This mechanistic link provides a framework for understanding echo chambers and suggests practical guides for intervention, advancing both fundamental research and policy-relevant insights into digital democracy.

## Introduction

The term *echo chamber* was popularized by Jamieson and Cappella (1) to describe a bounded media space that amplifies messages while insulating them from rebuttal. With the rise of social media, echo chambers, often conflated with *filter bubbles* (2), have become central to debates on misinformation, democratic resilience, and collective decision-making (3, 4). These *information-limiting environments* are thought to exacerbate polarization by restricting exposure to diverse viewpoints (5, 6).

Despite their importance, the mechanisms driving echo chambers remain difficult to pin down. Online behavior arises from a multilayered interplay between individual activity and platform-level algorithms. Features such as feed ranking, recommendation, and follow/unfollow dynamics evolve continuously and are rarely transparent. To understand and eventually mitigate polarization, it is crucial to develop *mechanistic models* that disentangle user-driven interactions from algorithmic interventions and allow systematic exploration of their effects.

Opinion dynamics models have long provided insights into how interactions shape collective beliefs (7–11). Beyond static interactions, adaptive rewiring shows how homophily can fragment networks into polarized communities (12–15). Recommendation mechanisms can further sharpen polarization, even when rewiring is unbiased (16, 17). Yet, the role of information diffusion itself, including limited attention and reposting, has received comparatively less mechanistic treatment.

Empirical studies show that individuals do not scale their attention with the abundance of available content (18). Instead, reposting, memory limits, and competition among posts produce heavy-tailed popularity distributions and critical-like dynamics (19–21). While such mechanisms are central to the physics of social media (22), their connection to polarization and echo-chamber formation remains underexplored.

Here, we investigate the interplay between information spreading and opinion polarization using an agent-based model on adaptive networks. We extend the framework of de Arruda et al. (16) by

**Competing Interest:** The authors declare no competing interests.

**Received:** August 28, 2025. **Accepted:** November 23, 2025

© The Author(s) 2025. Published by Oxford University Press on behalf of National Academy of Sciences. This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (<https://creativecommons.org/licenses/by-nc/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is properly cited. For commercial re-use, please contact [reprints@oup.com](mailto:reprints@oup.com) for reprints and translation rights for reprints. All other permissions can be obtained through our RightsLink service via the Permissions link on the article page on our site—for further information please contact [journals.permissions@oup.com](mailto:journals.permissions@oup.com).



equipping agents with limited memory, enabling reposting of past content as in real-world feeds. This extension lets us track information cascades while studying how posting, recommendation, and rewiring interact with opinion dynamics.

We address three questions: (i) How does reposting under limited memory alter the states of consensus and polarization? (ii) Under what conditions do information-limiting environments sustain polarization or allow depolarization? (iii) Can the model reproduce empirical cascade patterns observed in polarized debates, and what does this imply about the mechanisms behind polarization? Through extensive simulations and calibration against datasets on Brexit and vaccine debates (23), we show that the same mechanisms that generate realistic cascade distributions also reproduce polarized opinion states. This provides a mechanistic link between diffusion and polarization.

## The agent-based model

Our model extends the one introduced by de Arruda et al. (16). That model represents social media platform interactions and treats opinion as a continuous variable, bounded between  $-1$  and  $1$ , held by each user. Social media users are represented as nodes in an adaptive directed network. A user points to another if they receive content from them (followership relation); that is, the content is spread following the opposite direction of edges.

Each iteration of the model follows successive steps, which act as filters that determine whether a randomly activated user goes to the next step.

In this work, we extend the model by allowing users to post previously existing content. Each piece of content has a fixed opinion value  $\theta$ , always drawn uniformly at random upon its creation. However, here each user is equipped with a memory list of size  $\alpha$ , filled as they receive content from neighbors (i.e. a social media feed). Each user's memory can be seen as an ordered list of  $\theta$  values from previous posts.

Memory lists are updated in such a way that the last piece of content received is prioritized. We index the memory list with  $\theta_c$ , such that  $c \in \{1, \dots, \alpha\}$ , and the lower the index, the more likely the content will be posted. New pieces of content are introduced from the beginning of the list ( $\theta_1$ ) based on what users receive from their neighbors. In this case, existing pieces of content move to a lower order (i.e.  $\theta_{c+1} \leftarrow \theta_c$ ).

When the list is at full capacity (there are  $\alpha$  pieces of content) and new content is received, the piece at the bottom ( $\theta_\alpha$ ) is removed, and all other pieces move one position lower so that the new one is placed at the top.

When reposting, users choose the first piece of content in their memory lists ( $\theta_1$ ). Once the content is posted, it moves from the beginning to the end of the list, so it will be picked last. Hence, the opinion of the content which was just posted becomes  $\theta_\alpha$  in the memory list, and each other piece moves to a higher order (i.e.  $\theta_c \leftarrow \theta_{c+1}$ ).

We further introduce a new parameter called the innovation probability  $\mu$ , which, at the posting step, controls how many iterations inject new content into the system. That is, when  $\mu = 1$ , the base model from de Arruda et al. (16) is recovered regardless of the memory lists.

Each iteration of our model is described by the following steps:

1. **Activating:** a user  $i$ , chosen uniformly at random, becomes active.
2. **Posting:** the active user  $i$  either creates a new content with probability  $\mu$ , or chooses the content at the top of their

memory list ( $\theta = \theta_1$ ) with probability  $1 - \mu$ . The content is posted according to a probabilistic filter function (Eq. 1) of their opinion  $b_i$  and the content opinion  $\theta$ .

3. **Receiving:** Each follower  $j$  of node  $i$  receives the content according to another probabilistic filter function (Eq. 2) of  $b_j$ ,  $b_i$  and a parameter  $\phi$ . When they do, the content is added to their memory list from the top. The control of this filter through  $\phi$  is a proxy for a social media recommendation algorithm.
4. **Realigning:** Each opinion  $b_j$  of followers who successfully received the content increases or decreases by a constant  $\Delta$ . The opinions are repelled away from the post with probability  $|\theta - b_j|/2$  and attracted otherwise.
5. **Rewiring:** Each follower  $j$  may stop following  $i$  and start following another user at random through a probability proportional to the difference between  $b_i$  and  $b_j$ .

These steps are visually summarized in Fig. 1. In our model extension, we implement rules to define user behavior, which were previously studied in Refs. (16, 24). In particular, we set the probabilistic rule acting as the posting filter as

$$P_p(\theta, b_i) = \cos^2\left(\frac{\pi}{2}|\theta - b_i|\right), \quad (1)$$

where  $\theta$  is the opinion of the content (whether created or picked from the memory list), and  $b_i$  is the opinion of the active node  $i$ . This is referred to in Ref. (24) as *conflicting posting*, which is the same as  $P_t^{\text{pol}}$  from Eq. 2 in Ref. (16).

Correspondingly, the receiving filter is given by

$$P_r(b_i, b_j, \phi) = \cos^2\left(\frac{\pi}{2}|b_i - b_j| + \phi\right). \quad (2)$$

This filter was previously studied as the function  $P_d^l$  from Eq. 5 in Ref. (16). We refer to the parameter  $\phi$  as the *recommendation control*. It determines where the cosine-squared function begins, which enables us to adjust the initial behavior of the recommendation algorithm. That is, it may let information propagate between pairs of users who agree with each other or the opposite. Notice the receiving step does not consider the content opinion  $\theta$ .

## Opinion polarization measures

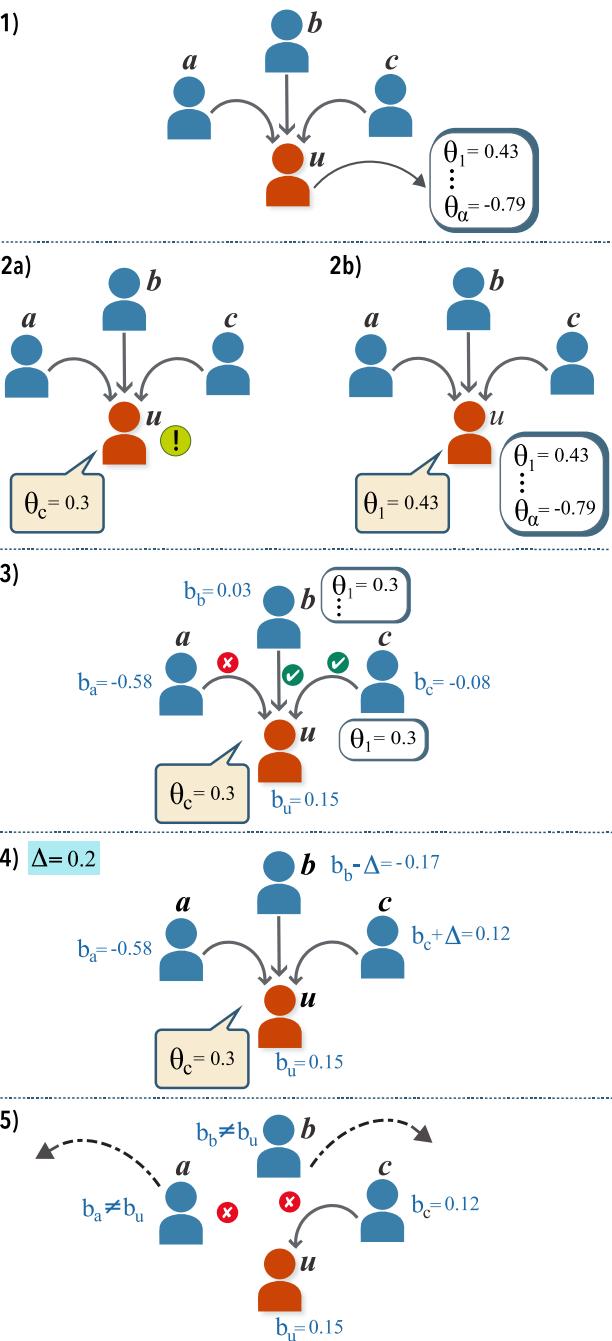
The main system state we are interested in measuring is the distribution of user opinions. In particular, we characterize a state as polarized when the opinion distribution is bimodal.

To measure the bimodality of the opinion distributions, we adopt the bimodality coefficient used by Arruda et al. (16), which ranges from 0 to 1. The bimodality coefficient  $BC(b)$  (25) is defined as

$$BC(b) = \frac{g^2 + 1}{k + \frac{3(n-1)^2}{(n-2)(n-3)}}, \quad (3)$$

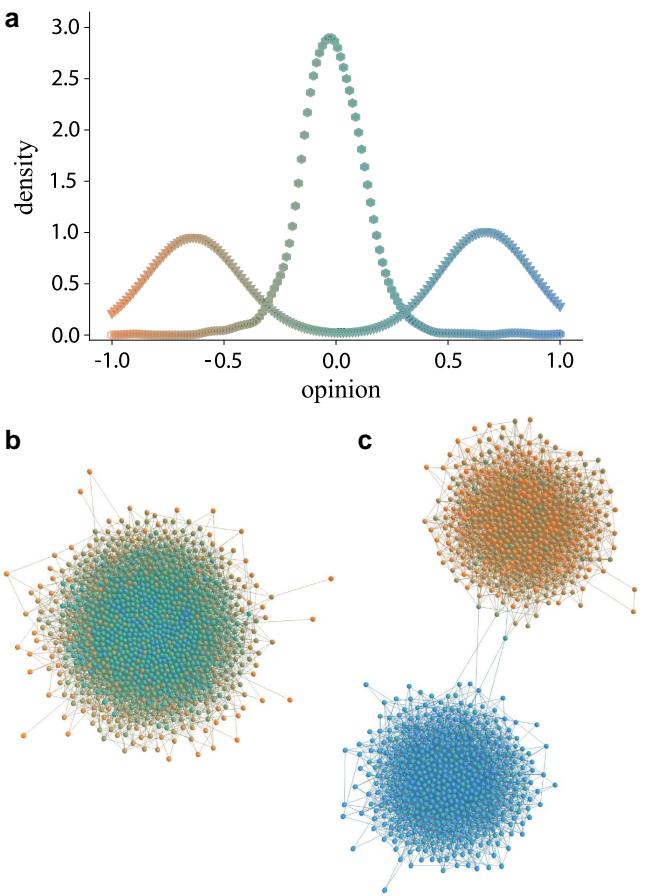
where  $g$  is the sample skewness (26) and  $k$  is the excess kurtosis (26) of the opinion distribution  $b$ . The distribution  $b$  is typically considered bimodal when  $BC(b) > 5/9$ , as empirically shown in (25).

Another way to measure how the opinion distribution changes towards polarization is via the moment ratio diagram, which is composed of bounded versions of the third- and fourth-moment ratios of the distribution called L-moments (27). That is, given the opinion distribution  $b$  and its  $k$ th order statistic  $b_{(k)}$ , its sample  $r$ th L-moment  $\lambda_r$  is defined by



**Fig. 1.** Schematics of the model iterations: the diagram represents each iteration step. In the first (activating), user  $u$  was chosen at random and has three followers (users  $a$ ,  $b$ , and  $c$ ). Each node features a memory list with previous content, but only  $u$ 's is shown. Next, the posting step may happen in two ways. With probability  $\mu$ , user  $u$  posts new content (whose opinion was drawn as  $\theta = 0.3$ ) that was not in his memory list; otherwise, with probability  $1 - \mu$ , the first content of user  $u$ 's list is selected and posted ( $\theta_1 = 0.43$ ). Following on to the receiving step and assuming  $u$  introduced the new content ( $\theta_c = 0.3$ ), the probabilistic filter (defined in Eq. 2) does not allow the new content to be incorporated into  $a$ 's memory list, but only to the other two neighbors. Then in the realigning step, neighbors who incorporated the new content have their opinion either attracted or repelled towards  $\theta$  by a constant factor ( $\Delta = 0.2$  in the example). Finally, those neighbors whose opinions are sufficiently different than  $u$  rewire their followership edge to another random node in the rewiring step.

$$\lambda_r = r^{-1} \binom{n}{r}^{-1} \sum_{b_{(1)} < \dots < b_{(j)} < \dots < b_{(r)}} (-1)^{r-j} \binom{r-1}{j} b_{(j)}. \quad (4)$$



**Fig. 2.** New system states introduced with the parameter  $\mu$ : the same parameterization except for different innovation probability  $\mu$  can lead to consensus in a homogeneous topology or two polarized communities. In a), the opinion distribution is in the other two panels, where hexagon markers represent the network in (b) and triangle markers the one in (c). In b), the innovation probability  $\mu = 1$ . In c),  $\mu = 0.1$ .

The moment ratio diagram is composed of the L-skewness  $\tau_3 := \lambda_3/\lambda_2$ , which is bounded between  $-1$  and  $1$ , and the L-kurtosis  $\tau_4 := \lambda_4/\lambda_2$ , which is bounded between  $-1/4$  and  $1$ .

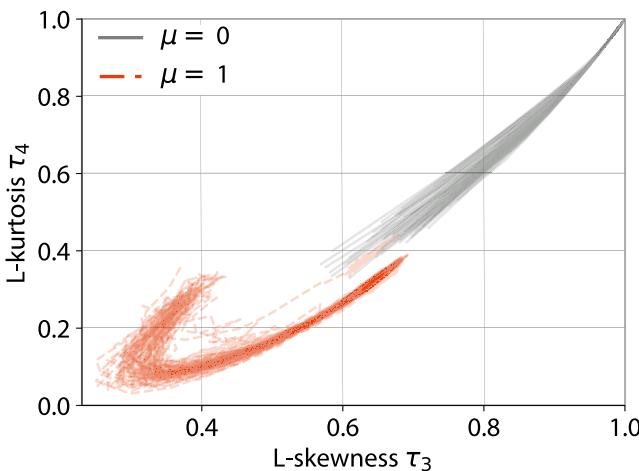
We choose these over the conventional moments (which are unscaled) to compare different stages in the evolution of the opinion distribution under the same scale. However, the L-moments share many of the properties of their conventional counterparts. For example, a distribution with high L-skewness would be interpreted as being skewed. We work out a small example with synthetic distributions to show how these measures behave in the [Supporting Information S1](#).

## Results

### Innovation probability introduces new states

As the first step to characterize our model, we run an experiment on a known parameterization in de Arruda's model to understand the impact of the innovation probability  $\mu$ . To this end, we use an Erdős-Rényi network of size  $N = 10^3$  and mean in-degree  $z = 10.21$ , and pick parameters known to lead the system to a consensus scenario when the original model is recovered ( $\mu = 1$ ).

However, running the simulation with the same parameters but a different innovation probability ( $\mu = 0.1$ ) yields a polarized system. We compare the two configurations for the same number of iterations in Fig. 2. In Fig. 2a, the opinion distributions are



**Fig. 3.** Characterizing the effect of innovation on opinion evolution: the moment ratio diagram for the evolution of the opinion distribution in an Erdős-Rényi network of size  $N = 10^3$  and mean in-degree  $z = 10.21$  for the minimum and maximum innovation probability  $\mu$ . Both parameterizations are tested with 100 different trajectories starting from the same initial opinion distribution (uniformly random between  $-1$  and  $1$ ), omitted in the diagram for clarity. Each trajectory has 100 points, and every two consecutive points are produced from steps of  $10N$  iterations of the model.

plotted against each other. We also show how each network structure is wired in each case, with nodes colored according to their opinions. Figure 2b shows the case for  $\mu = 1$ , which is statistically the same as the initial network, and opinions are homogeneously spread around 0. Figure 2c reveals that, for the polarized state at  $\mu = 0.1$ , the network becomes separated into two densely connected groups of largely opposing opinions.

In order to characterize the impact of the innovation probability  $\mu$  in the opinion distribution, we run a hundred simulations on the same network for  $\mu = 0$  (no innovation) and  $\mu = 1$  (de Arruda's original model). We then use each of these simulations to produce a trajectory of opinion distributions with 100 points. These trajectories are then compared via the moment ratio diagram of Fig. 3, which uses L-moments to describe these distributions. The L-skewness and L-kurtosis are bounded counterparts to the sample skewness and excess kurtosis. Skewness measures the asymmetry of a distribution, indicating whether the distribution of opinions is skewed. Kurtosis indicates whether the data has heavier or lighter tails compared to a normal distribution, with higher kurtosis indicating more extreme values.

We show in Fig. 3 how opinion distributions become very separated, even though they share the same starting point for both  $\mu = 0$  and  $\mu = 1$ . This starting point was omitted to improve visibility, but it is found around  $(\tau_3, \tau_4) = (0.16, 0.16)$ , corresponding to an opinion distribution which is approximately uniform (considering the finite-sized network) between  $-1$  and  $1$ .

The final point of the  $\mu = 1$  trajectories is around  $(0.65, 0.3)$ , while that of the  $\mu = 0$  trajectories is nearly the extreme coordinates  $(1, 1)$ . The extreme coordinates correspond to distributions accumulated into a few specific values, as  $\mu = 0$  prevents the creation of posts with different opinions  $\theta$  once the system is initialized. This is the same as the predominance of very few values of  $\theta$  across the system (see Supporting Information S2).

## Simulating an information-limiting environment

Next, we implement an experiment to investigate opinion dynamics parameterized to hinder the spread of information. This is

done by running simulations on a network with two communities via a stochastic block model, where users in one community are initialized with opinion 1, and those in the other are initialized with opinion  $-1$ . Each community has 500 users with a mean in-degree and out-degree of  $z = 8$ , and only 1% of edges bridge the two communities. Users are forced to remain in their community as we disable the possibility of rewiring. Furthermore, we specify an alternate posting filter (Eq. 5) and compare simulation outcomes between this filter and the original one (Eq. 1). See Materials and methods for more details.

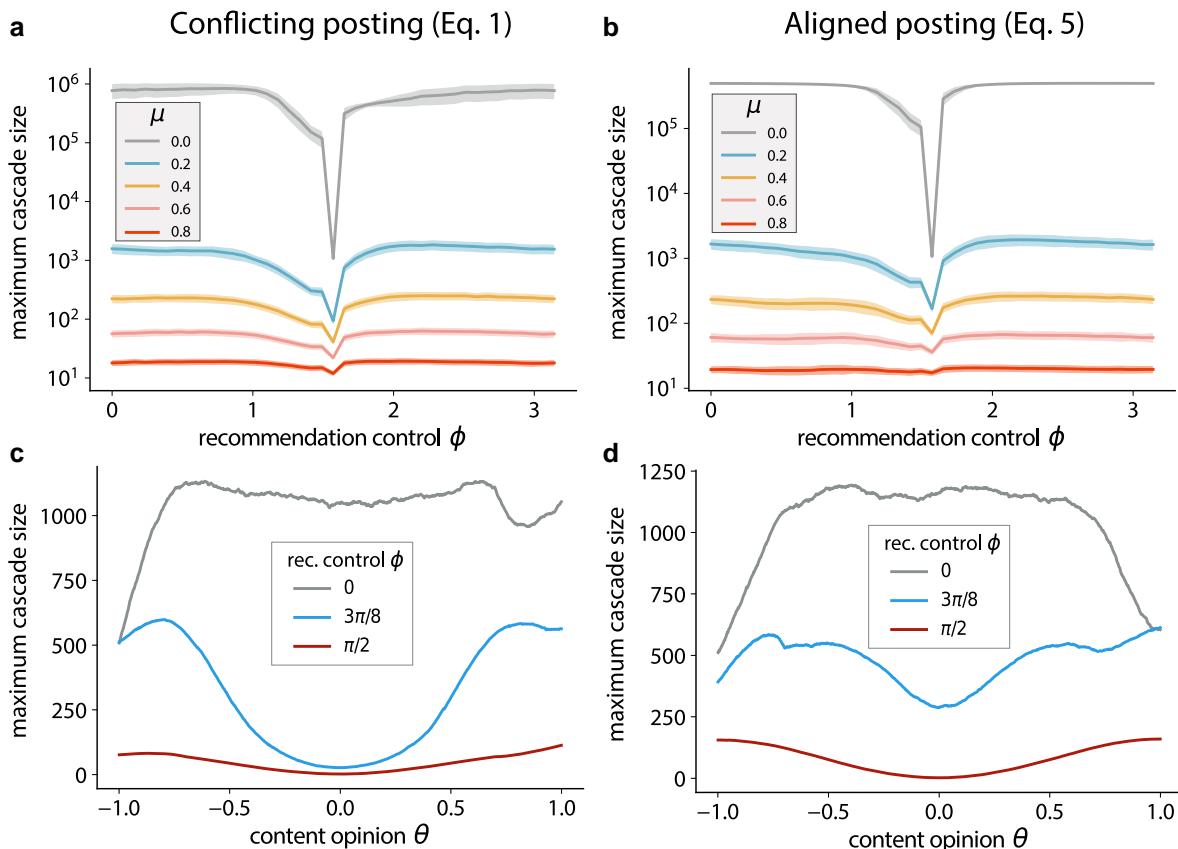
The spread of information is measured by looking at the maximum size reached by cascades, which is the same as the number of users who share the same piece of content. In Fig. 4a and b, we show how cascades can grow according to the relationship between the innovation probability  $\mu$  and the recommendation control  $\phi$ , each for a specific posting filter (Eq. 1 for (a), Eq. 5 for (b)). For each pair of  $\mu$  and  $\phi$ , we observe the maximum cascade size in each of the 500 simulations and average them to obtain the solid line, with shades representing the difference to the corresponding SD.

As expected, the lower the innovation probability  $\mu$  is, the greater the sizes cascades can reach since fewer posts are competing with each other. Both posting filters have qualitatively similar behavior, with the conflicting posting showing a slightly higher SD for  $\mu = 0$ . Our experiment setup is shown to limit the spread of information around  $\phi = \pi/2$ , where the receiving filter (Eq. 2) hinders communication within and between communities. Next, we measure how contents spread in this scenario regarding their opinion  $\theta$ . By fixing  $\mu = 0.2$  and three values of  $\phi$ , Fig. 4c and d shows that the less permissive recommendation control regime ( $\phi = \pi/2$ ) favors extreme content opinion, while the most permissive ( $\phi = 0$ ) punishes it. However, the middle regime ( $\phi = 3\pi/8$ ) displays different behaviors for each posting filter. In Fig. 4c, it noticeably hinders the growth of moderate content (i.e.  $\theta$  around 0). In Fig. 4d, this effect is less pronounced, meaning the dynamics still permit the spread of content from the entire  $\theta$  range.

Then, we demonstrate in Fig. 5 how the increase of content variety through the parameter  $\mu$  affects the final opinion distribution for users locked into the two communities starting with extreme opinions. To do this, we use the bimodality coefficient BC. In Fig. 5a, the conflicting posting filter blocks moderate posts (by not allowing their cascades to grow, as seen in Fig. 4c). However, when the aligned posting filter is employed (Fig. 5b), the system is depolarized as the variety of content increases with  $\mu$ . For  $\phi = \pi/2$ , cascade growth is hindered, and therefore, the system remains polarized throughout the range of  $\mu$  for both filters, well above the line indicating the BC threshold at  $5/9$ . But when the aligned posting filter is used, for  $\phi = 3\pi/8$  and  $\phi = 0$ , we see that even though the BC values of the final opinion distribution show a wide dispersion at small  $\mu$ , they narrow into unimodal opinion distributions as soon as  $\mu = 0.3$  and remain under the dashed threshold  $5/9$  for most of the remaining  $\mu$  range.

## Model calibration against data

We implement a calibration task to see how our model conforms to real-world scenarios. Given the availability of network structure, user opinions, and cascade sizes in the datasets, we opt to focus on the latter as the target of our model calibration. That is, we consider the empirical cascade size distribution to be aggregated over time, while the network structure and user opinions are left as degrees of freedom in the time evolution of our model (even though all simulations start from the same network structure).



**Fig. 4.** Effects of model calibration on information spreading in a synthetic information-limiting environment: In a and b), for two different posting filters, the maximum cascade size produced for a range of recommendation control  $\phi$  and five values of innovation probability  $\mu$ . For each pair of  $\mu$  and  $\phi$ , we run 500 simulations and obtain the maximum cascade size in each of them; the solid line is the average among the 500 maximum cascade sizes, while the shade is the average added or subtracted to the corresponding SD. In c and d), again for each respective posting filter, we compare the maximum cascade size with content opinion  $\theta$  for three values of recommendation control  $\phi$  on a fixed innovation probability  $\mu = 0.2$ . Since each cascade is associated with an opinion content  $\theta$ , which is continuous, we group cascades into each of a thousand bins between  $-1$  and  $1$ . From these, we extract those with the maximum cascade size; the resulting values are represented with lines smoothed via the Savitzky-Golay filter (28).

Thus, we formulate the model calibration as an optimization task to reproduce the closest possible distribution of cascade sizes over our parameter space. In particular, we search for the parameters that control the behavior of the dynamics: the innovation probability  $\mu$ , the recommendation control  $\phi$ , the opinion variation  $\Delta$ , and the number of iterations  $n_{\text{iter}}$ . The precise formulation of the optimization task is given in Materials and methods.

In Fig. 6, we show the results for simulations parameterized according to the lowest values of KS statistic observed, which are at  $\alpha = 27$  for the Brexit dataset and  $\alpha = 29$  for the VaxNoVax dataset. For values of  $\alpha < 30$ , there is no significant difference in the KS results, indicating that  $\alpha$  does not play a major role in the model's dynamic outcomes. More details regarding the optimization are shown in Supporting Information S4.

For both datasets, the optimized models represent most of the cascades found in the real data. Specifically, the values of the KS statistics are on average  $(3.617 \pm 0.595) \times 10^{-3}$  for Brexit and  $(1.385 \pm 0.146) \times 10^{-2}$  for VaxNoVax. Note that a common choice to consider the KS statistic as a good fit is 0.05. In the case of our results, the values obtained are significantly lower. We acknowledge that there is a limitation in the tails of the distributions, which are not well represented. However, since the plot axes follow logarithmic scales, the discrepancy between data and simulation is much more pronounced at the extreme events (see Fig. 6).

To further understand the opinion dynamics and the effect of the cascades on the opinions, we calculate the bimodality

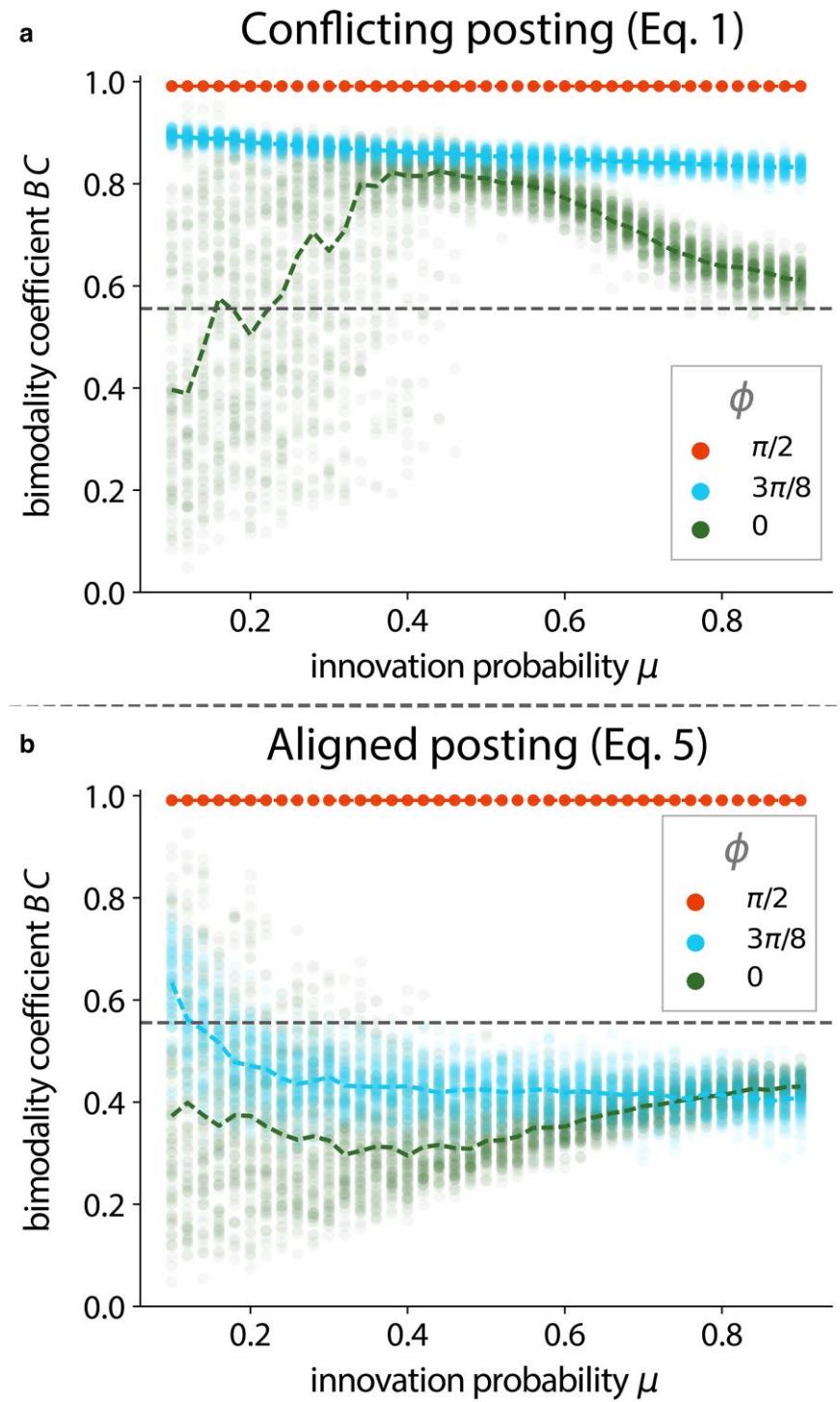
coefficient of the resulting opinions  $BC(b)$ . On average,  $BC(b)$  is  $(0.892 \pm 0.007)$  for Brexit and  $(0.623 \pm 0.010)$  for VaxNoVax. Both values indicate that the opinion distributions are bimodal, and thus, the opinions are polarized. Interestingly, the parameters that generate cascades compatible with the real data also polarize opinions in the synthetic system. Furthermore, the bimodality of VaxNoVax is consistent with the results found in real data in our previous study (16), reported to be between 0.60 and 0.67.

The methodology used considers that the best result is not necessarily when the dynamics converge to a fixed result. However, for the best results obtained here, we found that  $BC$  does not change significantly even when the simulation runs for many more iterations.

## Discussion and conclusion

We introduced an agent-based model that captures the two-way coupling between information diffusion and opinion polarization on adaptive social networks. By incorporating limited memory and reposting behavior, the model reflects key features of social media feeds, enabling us to study how cascades of information interact with the dynamics of opinion formation.

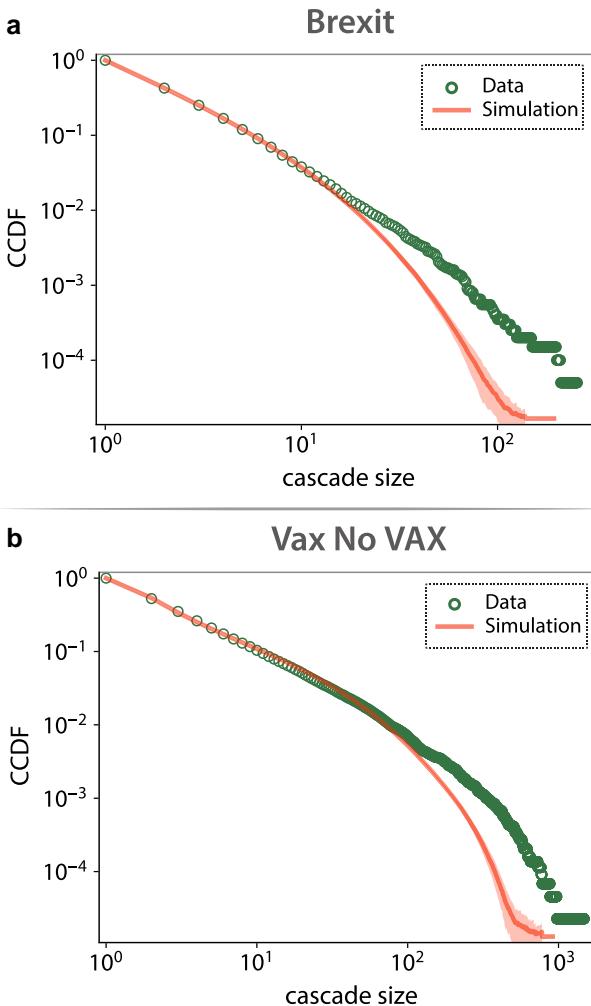
Our findings reveal several insights. First, the innovation probability (i.e. the likelihood of introducing new content) acts as a critical switch between consensus and polarization. Low innovation reduces content diversity, allowing entrenched narratives to



**Fig. 5.** Polarized communities reacting to the increase of content innovation: for each combination of posting filter (Eq. 1 in (a), Eq. 5 in (b)), each recommendation control  $\phi$  (coded by color) and each value of innovation probability  $\mu$ , a thousand simulations are run and represented by scatter points corresponding to the bimodality coefficient  $BC$  associated to the final opinion distribution. Colored dashed lines are the median of the 1,000 points at each value of  $\mu$ . After being initialized with extreme opposite opinions in each module, the experiment is performed without rewiring, and the resulting opinion distributions are obtained after  $10^6$  iterations. The dark dashed line at  $BC = 5/9$  signalizes where opinion distributions are considered bimodal.

dominate and splitting the network into polarized camps. Second, in information-limiting environments polarization is sustained or mitigated, even when users cannot rewire their social ties. Third, by calibrating the model to empirical cascade data from Brexit and vaccine debates, we showed that the parameter regimes that best reproduce real cascade statistics also lead to polarized opinion states.

Conceptually, our work offers a mechanistic explanation of echo chambers as emergent phenomena arising from the combined effects of user memory, reposting, and recommendation. Methodologically, it bridges theoretical opinion dynamics with empirical observations, showing that calibrated agent-based models can reproduce not only qualitative patterns but also statistical properties of real data. Practically, our results suggest that



**Fig. 6.** Calibrating the model against data: complementary cumulative distribution function (CCDF) of cascade sizes obtained from the optimal simulation against those from each dataset. The green line corresponds to the average of 100 simulations, while the shades account for the SD of simulated cascade sizes. In a), the model is calibrated against cascades from the Brexit dataset, and the set of optimal simulation parameters found is  $(\mu, \Delta, \phi, n_{\text{iter}}) = (0.658, 0.373, 1.366, 6.379 \times 10^5)$  for a feed size  $\alpha = 27$ . The KS statistic between the sample distribution from data and those from simulations averages at  $(3.617 \pm 0.595) \times 10^{-3}$ . In b), the model is calibrated against cascades from the VaxNoVax dataset, and the set of optimal parameters is  $(\mu, \Delta, \phi, n_{\text{iter}}) = (0.093, 0.724, 2.863, 6.474 \times 10^5)$  for a feed size  $\alpha = 29$ . The KS statistic between the sample distribution from data and those from simulations averages at  $(1.385 \pm 0.146) \times 10^{-2}$ .

interventions targeting platform parameters, such as increasing content diversity or modifying reposting incentives, may help reduce polarization.

While our study advances understanding of polarized online environments, open challenges remain. Extreme events in cascade growth, not fully captured here, call for incorporating bursty dynamics and heterogeneity in user activity (21, 29, 30). Additionally, richer datasets combining structural, temporal, and opinion measures would enable more comprehensive calibration and validation. One may also want to understand how effective interventions can be designed from the insights of our model, which would involve modeling incentives via evolutionary game theory (31, 32). Furthermore, the study of agent-based models powered by large language models (LLMs) is gaining traction in the field of social simulations (33).

In conclusion, polarization and information diffusion cannot be studied in isolation. By tracking how cascades unfold alongside opinion shifts, our model provides a framework for investigating the mechanisms that sustain or dissolve echo chambers and offers a foundation for designing interventions to foster healthier online discourse.

## Materials and methods

### Datasets

In our study, we consider two datasets published by Minici et al. (23). They were collected from the social platform Twitter under a background of strongly polarized debates. These datasets contain measurements of network structure, cascades of retweets, and opinions of either users or posts. All of these measures are represented in our model, which makes the data good candidates for empirical studies.

The first dataset, named “Brexit,” was measured in the context of the Brexit referendum in the United Kingdom between May and July of 2016. It features a network with 7,589 nodes and 532,460 edges. The number of cascades is 19,963, with minimum and maximum sizes of 2 and 256, respectively. The complementary cumulative distribution function (CCDF) of the cascade sizes is shown in Fig. 6a.

The second dataset, dubbed “VaxNoVax,” was collected during vaccine debates in Italy in 2018. Its associated network has 14,315 nodes and 1,714,180 edges. It contains 43,923 cascades ranging in size from 2 to 1,468 (see Fig. 6b). We also provide the degree distribution of both networks in [Supporting Information S3](#).

### Parameter characterization

To produce Fig. 2 and 3, we set up the simulation on an Erdős-Rényi network of size  $N = 10^3$ , with recommendation control  $\phi = \pi/2$ , opinion variation  $\Delta = 0.1$  and feed size  $\alpha = 1$ . Opinions are initialized as a uniformly random number between  $-1$  and  $1$ , and we execute  $10^5$  iterations for Fig. 2, and 100 steps of  $10N$  iterations for each value of  $\mu$  in Fig. 3.

### Information-limiting environment experiment

Here, we describe how our experiment in an information-limiting environment is designed and set. First, we produce a synthetic network with two communities separated by very few edges. We employ the Poisson degree-corrected stochastic block model as implemented in Ref. (34) to generate a network with size  $N = 10^3$ , with both in-degree and out-degree distributions from Poisson with an average  $z = 8$  and two modules. Only 1% of the edges connect the two modules.

We want to design a polarized scenario where users are locked into two distinct communities, both separated by opinion and network structure. Hence, this experiment is executed with the rewiring rule disabled, and users have opinions initialized as  $1$  in one module and  $-1$  in the other. That is, the system always starts fully polarized in this experiment.

We also highlight an important distinction to the original model setup. Instead of the posting filter from Eq. 1, we opt to use the following:

$$P_p(\theta, b_i) = \begin{cases} \cos^2\left(\frac{\pi}{2}|\theta - b_i|\right), & \text{if } |\theta - b_i| \leq 1 \\ 0, & \text{otherwise.} \end{cases} \quad (5)$$

This change means that, in practice, a user  $i$  will never post a content of opinion  $\theta$  when the difference  $|\theta - b_i|$  is greater than one. Since each user is initialized with an opinion of either  $1$  or  $-1$ , they will not post any content whose  $\theta$  has a different sign to their

opinion unless it has changed since the beginning of the simulation.

Then, we set the opinion variation  $\Delta = 0.1$  and feed size  $\alpha = 1$  and proceed to understand how these features impact the growth of cascades of content posting by systematically testing how parameters react to each other, especially the recommendation control  $\phi$  and innovation probability  $\mu$ .

To produce Fig. 4a and b, we fix each value of  $\mu$  and a range of 41 values of  $\phi$  between 0 and  $\pi$ . For each pair of  $\mu$  and  $\phi$ , we run 500 simulations of  $10^6$  iterations each and annotate what is the maximum cascade size resulting from each. The solid line represents the average of the 500 maximum cascade sizes, and the shades are the corresponding SD added or subtracted from the average.

We obtain Fig. 4c and d by fixing  $\mu = 0.2$ , three values of  $\phi$  and running 100 simulations for  $10^6$  iterations each. Each value of  $\phi$  then generates a large number of cascades: 9,372,911 cascades for  $\phi = 0$ , 8,901,320 cascades for  $\phi = 3\pi/8$  and 5,090,826 cascades for  $\phi = \pi/2$ . These are binned through a thousand values of content opinion  $\theta$ , which we then use to extract a maximum cascade size. Finally, the maximum cascade sizes for each bin of  $\theta$  are represented together through smoothed lines via the Savitzky–Golay filter (28).

Finally, Fig. 5 considers a range of 41 values of  $\mu$ , from 0.1 to 0.9, and the three previously tested values of  $\phi$  (0,  $3\pi/8$ , and  $\pi/2$ ). At each point given by a combination of parameters, we run 1,000 simulations.

## Model calibration task

Our model is calibrated via a search in a multidimensional space made of four parameters, which are  $\mu$ ,  $\phi$ ,  $\Delta$ , and  $n_{\text{iter}}$ . In particular, we minimize the Kolmogorov–Smirnov (KS) distance via a single-objective genetic algorithm from the implementation in Ref. (35) by considering the feed size  $\alpha$  as a fixed parameter when the heuristic is initialized. The values of  $\alpha$  are selected considering the KS statistic obtained for a range up to  $\alpha = 45$ , found in the [Supporting Information S4](#).

Since the datasets are a small sample (36) of the activity in subsets of the Twitter platform retrieved by the first version of the API, it is a challenging task to accurately estimate the parameters a priori due to their dependence on time. We cannot know how the volume of activity represented in the sample maps to the actual one in the system at the time. In essence, it is not trivial to match one simulation iteration to a time scale in the real world.

As such, two of our parameters encapsulate the time frame we are trying to match: the number of iterations  $n_{\text{iter}}$ , but also the opinion variation  $\Delta$  as it modulates how intense changes in user opinion are. Intuitively, a high value of  $\Delta$  means fewer iterations are needed for the system to reach stable configurations. Once there is a time window set by  $n_{\text{iter}}$  and  $\Delta$ , the innovation probability  $\mu$  is fundamental to fix the maximum cascade size.

Finally, the recommendation control  $\phi$ , mentioned back in section “The agent-based model” as related to how users receive posts of different opinions, can drive the system towards various states of opinion polarization (as observed in previous works (16, 24)).

Let  $d$  be the empirical distribution of cascade sizes from data and  $s$  the sample distribution of cascade sizes obtained from the simulation. As each dataset has a fixed distribution of cascade sizes, let the empirical cumulative distribution function of  $d$  be given by  $F_d(x)$ . As for the sample distribution  $s$ , it is parameterized with the innovation probability  $\mu$ , the recommendation control  $\phi$ , the opinion variation  $\Delta$ , and the number of iterations  $n_{\text{iter}}$ .

We are then interested in finding the simulation parameters that minimize the Kolmogorov–Smirnov (KS) statistic (37) between the two samples. That is, the model calibration task is

$$\underset{\mu, \phi, \Delta, n_{\text{iter}}}{\text{minimize}} \sup_x |F_d(x) - F_s(x, \mu, \phi, \Delta, \alpha, n_{\text{iter}})| \quad (6a)$$

$$\text{subject to } 0.01 \leq \mu \leq 0.99 \quad (6b)$$

$$0 \leq \phi \leq \pi \quad (6c)$$

$$0 \leq \Delta \leq 1 \quad (6d)$$

$$1.5 \times 10^5 \leq n_{\text{iter}} \leq 6.5 \times 10^5 \quad (6e)$$

where  $F_s(x, \mu, \Delta, \phi, \alpha, n_{\text{iter}})$  is the empirical cumulative distribution of cascade sizes from simulations parameterized accordingly.

After fixing a value of  $\alpha$ , which we obtain by testing the KS statistic for different ranges of parameters (see [Supporting Information S4](#)), we obtain an optimal set of parameters with the genetic algorithm. From this set of parameters, Fig. 6 is produced by comparing the empirical cascade size tail distribution in the datasets to the one produced by averaging 100 simulations with the set of parameters.

These simulations are executed using the empirical network structure from the dataset, with opinions initialized uniformly at random between the entire interval  $[-1, 1]$ . Each iteration follows the description of section “The agent-based model,” including the posting filter given by Eq. 1.

## Supplementary Material

Supplementary material is available at [PNAS Nexus](#) online.

## Funding

K.A.O. was partially supported by the European Union through ERC grant (ID-COMPRESSION, grant number: 101124175). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council Executive Agency. Neither the European Union nor the granting authority can be held responsible for them. Y.M. was partially supported by the Government of Aragón, Spain, and “ERDF A way of making Europe” through grant E36-23R (FENOL), and by Ministerio de Ciencia, Innovación y Universidades, Agencia Española de Investigación (MICIU/AEI/10.13039/501100011033) grant no. PID2023-149409NB-I00. We acknowledge the use of the computational resources of COSNET Lab at Institute BIFI, which was funded by Banco Santander (grant Santander-UZ 2020/0274) and the Government of Aragón (grant UZ-164255). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

## Author Contributions

H.F.A. and K.A.O.: conceptualization, data curation, software, formal analysis, validation, investigation, visualization, methodology, writing-original draft, writing-review and editing. Y.M.: conceptualization, investigation, visualization, methodology, writing-original draft, writing-review and editing.

## Preprints

This manuscript was posted on a preprint: <https://arxiv.org/abs/2410.17151>.

## Data Availability

The data used in this article was published by Minici et al. (2022) [DOI:10.1145/3511808.3557253] and are available in a GitHub repository at <https://github.com/mminici/Echo-Chamber-Detection>. They were downloaded on 2023 Feb 21 (Brexit dataset) and 2023 Mar 17 (VaxNoVax dataset). A Python library named DOCES (Dynamical Opinion Clusters Exploration Suite) (38) implements the model described in our paper and is available at <https://github.com/hfarruda/doces>.

## References

- 1 Jamieson KH, Cappella JN. *Echo chamber: rush Limbaugh and the conservative media establishment*. Oxford University Press, New York, NY, USA, 2010.
- 2 Arguedas AR, Robertson C, Fletcher R, Nielsen R. 2022. Echo chambers, filter bubbles, and polarisation: a literature review. Technical report, Oxford, England.
- 3 Bak-Coleman JB, et al. 2021. Stewardship of global collective behavior. *Proc Natl Acad Sci U S A*. 118(27):e2025764118.
- 4 Lorenz-Spreen P, Oswald L, Lewandowsky S, Hertwig R. 2023. A systematic review of worldwide causal and correlational evidence on digital media and democracy. *Nat Hum Behav*. 7(1): 74–101.
- 5 Terren L, Borge-Bravo R. 2021. Echo chambers on social media: a systematic review of the literature. *Rev Commun Res*. 9:99–118. <https://www.rcommunicationr.org/index.php/rcr/article/view/16>.
- 6 Kubin E, von Sikorski C. 2021. The role of (social) media in political polarization: a systematic review. *Ann Int Commun Assoc*. 45(3):188–206.
- 7 Deffuant G, Neau D, Amblard F, Weisbuch G. 2000. Mixing beliefs among interacting agents. *Adv Complex Syst*. 3(01n04):87–98.
- 8 Hegselmann R, Krause U. 2002. Opinion dynamics and bounded confidence: models, analysis and simulation. *J Artif Soc Soc Simul*. 5(3).
- 9 Galam S. 2002. Minority opinion spreading in random geometry. *Eur Phys J B-Condens Matter Complex Syst*. 25(4):403–406.
- 10 Noorazar H, Vixie KR, Talebanpour A, Hu Y. 2020. From classical to modern opinion dynamics. *Int J Mod Phys C*. 31(7):2050101.
- 11 Chen G, et al. 2017. Deffuant model on a ring with repelling mechanism and circular opinions. *Phys Rev E*. 95(4):042118.
- 12 Holme P, Newman MEJ. 2006. Nonequilibrium phase transition in the coevolution of networks and opinions. *Phys Rev E*. 74(5): 056108.
- 13 Blex C, Yasseri T. 2022. Positive algorithmic bias cannot stop fragmentation in homophilic networks. *J Math Sociol*. 46(1):80–97.
- 14 Sasahara K, et al. 2021. Social influence and unfollowing accelerate the emergence of echo chambers. *J Comput Soc Sci*. 4(1): 381–402.
- 15 Han W, Feng Y, Qian X, Yang Q, Huang C. 2020. Clusters and the entropy in opinion dynamics on complex networks. *Phys A Stat Mech Appl*. 559:125033.
- 16 Ferraz de Arruda H, et al. 2022. Modelling how social network algorithms can influence opinion polarization. *Inf Sci*. 588:265–278.
- 17 Valensise CM, Cinelli M, Quattrociocchi W. 2023. The drivers of online polarization: fitting models to data. *Inf Sci*. 642(51):119152.
- 18 Weng L, Flammini A, Vespignani A, Menczer F. 2012. Competition among memes in a world with limited attention. *Sci Rep*. 2(1):335.
- 19 Lerman K, Ghosh R. 2010. Information contagion: an empirical study of the spread of news on digg and twitter social networks. *Proc Int AAAI Confer Web Soc Med*. 4(1):90–97.
- 20 Oliveira DFM, Chan KS, Leonel ED. 2018. Scaling invariance in a social network with limited attention and innovation. *Phys Lett A*. 382(47):3376–3380.
- 21 Notarmuzi D, Castellano C, Flammini A, Mazzilli D, Radicchi F. 2022. Universality, criticality and complexity of information propagation in social media. *Nat Commun*. 13(1):1308.
- 22 Luca Ciampaglia G, Flammini A, Menczer F. 2015. The production of information in the attention economy. *Sci Rep*. 5(1):9452.
- 23 Minici M, Cinus F, Monti C, Bonchi F, Manco G. Cascade-based echo chamber detection. In: *Proceedings of the 31st ACM International Conference on Information & Knowledge Management*, CIKM '22. Association for Computing Machinery, New York, NY, USA; 2022. p. 1511–1520.
- 24 Ferraz de Arruda H, Oliveira KA, Moreno Y. 2024. Echo chamber formation sharpened by priority users. *iScience*. 27. <https://doi.org/10.1016/j.isci.2024.111098>.
- 25 Pfister R, Schwarz KA, Janczyk M, Dale R, Freeman J. 2013. Good things peak in pairs: a note on the bimodality coefficient. *Front Psychol*. 4. <https://doi.org/10.3389/fpsyg.2013.00700>.
- 26 Kokoska S, Zwillinger D. *CRC standard probability and statistics tables and formulae*, student edition. 1st ed. CRC Press, Boca Raton, FL, USA, 2000. <https://doi.org/10.1201/b16923>.
- 27 Hosking JRM. 2018. L-moments: analysis and estimation of distributions using linear combinations of order statistics. *J R Stat Soc Series B Stat Methodol*. 52(1):105–124.
- 28 Savitzky A, Golay MJE. 1964. Smoothing and differentiation of data by simplified least squares procedures. *Anal Chem*. 36(8): 1627–1639.
- 29 Crane R, Sornette D. 2008. Robust dynamic classes revealed by measuring the response function of a social system. *Proc Natl Acad Sci U S A*. 105(41):15649–15653.
- 30 Karsai M, Jo H-H, Kaski K. 2017. *Bursty human dynamics*, SpringerBriefs in Complexity. 1st ed. Springer Cham, Cham, Switzerland, <https://doi.org/10.1007/978-3-319-68540-3>.
- 31 Wang S, Chen X, Xiao Z, Szolnoki A, Vasconcelos VV. 2023. Optimization of institutional incentives for cooperation in structured populations. *J R Soc Interface*. 20(199):20220653.
- 32 Cimpeanu T, Santos FC, Moniz Pereira L, Lenaerts T, Anh Han T. 2022. Artificial intelligence development races in heterogeneous settings. *Sci Rep*. 12(1):1723.
- 33 Lu Y, Aleta A, Du C, Shi L, Moreno Y. 2024. Llms and generative agent-based models for complex systems research. *Phys Life Rev*. 51(e2313925121):283–293.
- 34 Peixoto TP. 2014. The graph-tool python library. *figshare*. <https://doi.org/10.6084/m9.figshare.1164194>.
- 35 Blank J, Deb K. 2020. pymoo: multi-objective optimization in Python. *IEEE Access*. 8:89497–89509.
- 36 Morstatter F, Pfeffer J, Liu H, Carley K. 2021. Is the sample good enough? Comparing data from twitter's streaming API with twitter's firehose. *Proc Int AAAI Confer Web Soc Med*. 7(1):400–408.
- 37 Corder GW, Foreman DI. *Nonparametric statistics: a step-by-step approach*. 2nd ed. John Wiley & Sons, Nashville, TN, USA, 2014.
- 38 Ferraz de Arruda H, Andrade Oliveira K, Moreno Y. 2025. Dynamical opinion clusters exploration suite: modeling social media opinion dynamics. *SoftwareX*. 30(4):102136.