Information Sciences 588 (2022) 265–278

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Modelling how social network algorithms can influence opinion polarization

Henrique Ferraz de Arruda [a,b,*], Felipe Maciel Cardoso [a], Guilherme Ferraz de Arruda [c], Alexis R. Hernández [d], Luciano da Fontoura Costa [b], Yamir Moreno [a,c,e]

[a] Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Zaragoza, Spain
[b] São Carlos Institute of Physics, University of São Paulo, São Carlos, SP, Brazil
[c] ISI Foundation, Turin, Italy
[d] Institute of Physics, Rio de Janeiro Federal University, Rio de Janeiro, RJ, Brazil
[e] Department of Theoretical Physics, Faculty of Sciences, University of Zaragoza, Zaragoza, Spain

ARTICLE INFO

Article history:
Received 6 February 2021
Received in revised form 16 December 2021
Accepted 18 December 2021
Available online 25 December 2021

Keywords:
Network science
Social network
Opinion polarization
Echo chamber

ABSTRACT

The study of the dynamics of opinion formation and transmission in social networks has attracted lots of attention. Here, we propose a model that simulates communication in an online social network, in which randomly created posts represent external information. We consider users and friendship relations to be encoded as nodes and edges of a network. The dynamic of information diffusion is divided into two processes, referred to as *post transmission* and *post distribution*, representing the users' behavior and the social network algorithm, respectively. Individuals also interact with the post content by slightly adjusting their own opinion and sometimes redefining friendships. Our results show that the dynamic converge to various scenarios, which go from consensus formation to polarization. Importantly, friendship rewiring helps promote echo chamber formation, which can also arise for particular networks with well-defined community structures. Altogether, our results indicate that the social network algorithm is crucial to mitigate or promote polarization.

1. Introduction

With the advent of the internet, many different online social networks have been created. In order to understand the impact of these networks on the users' opinions, different models have been proposed [7,26,34,45,49,50]. The simulated dynamics include the voter model [15], the majority rule model [17], and the bounded confidence model [28], among others [7]. Part of these studies considers that the dynamic is executed on a network structure, in which the nodes and edges represent people and their friendship, respectively. Several distinct characteristics of social networks have been studied in order to understand opinion dynamics. For example, [43] considered that when two or more people have the same opinion, it is more likely for them to convince others. In general, these studies consider static network structures. However, other studies have taken into account that people can change their friendship [19,41]. In this case, edge rewirings are allowed, with might give rise to groups of connected people with similar opinions, called echo chambers.

One essential characteristic of these dynamics is how to represent the opinion. In many cases, the opinions are expressed only for two possible states [15,17]. In other cases, a varied number of categories [1], or vectors [4] can also describe the

* Corresponding author.
E-mail address: henrique.dearruda@isi.it (H. Ferraz de Arruda).

https://doi.org/10.1016/j.ins.2021.12.069
0020-0255/

opinions. Another option is to express opinions as a continuous variable [3,12,28,39], which can express problems regarding negotiations. In this case, opinions are not categorical, and the individuals can have intermediate opinions. Promising results have been obtained from this type of dynamic. For instance, in [3], results obtained from simulations were found to be similar to the scenario observed in online social networks.

Although in real social networks, individuals typically have lots of friends, in [1,5] the authors considered that a person is not capable of interacting with lots of other individuals. For this reason, they adopted network models with low average degrees. In order to constrain the interactions between individuals, we considered two complementary mechanisms. The first represents the user's action of posting pieces of information, henceforth called *post transmission*. In contrast to other approaches, individuals can post something different from their own opinions. We also simulate the individuals' information transference, a mechanism that chooses if the data shall be delivered to other users. We named this mechanism as *post distribution*. This step mimics how social network algorithms could manage posts. Moreover, since the nature of the posts received could change opinions positively or negatively, we consider two kinds of interaction: attraction and repulsion, respectively. Finally, we also allow individuals to rewire connections for the cases in which the post repulses the individual's opinion.

In order to model post transmission, we compare functions that represent different scenarios. The first function consists of users who post information they like or dislike, which can be understood as a reaction in a social network. We also analyzed users that only post information they like. In the third scenario, we considered users that did not pay attention to their posts. In the case of post distribution, we also test a wide range of options as this can play an essential role in social networks when it comes to keep the attention of the users.

We analyzed the results obtained by employing two distinct and complementary measurements. The first consists of analyzing the resulting opinion distributions. We quantify how polarized and balanced the opinion distributions are. However, more information can be obtained if the network structure is considered. As well as in [3,10,11], we measured the relationship between the node opinion and the average of the friends' opinions. From this analysis, different resulting structures can be observed, which include the presence of echo chambers. Our model can give rise to echo chambers for a specific scenario, even without friendship rewirings. Additionally, the bimodality of the opinion distribution does not guarantee that the dynamic would converge to echo chambers. We also compared our dynamic with networks obtained from Twitter. Interestingly, we found similarities between the results obtained with the proposed dynamics, including the level of bimodality and echo chambers.

The paper is organized as follows. Section 2 discusses the related works and their comparison with our approach. In Section 3, we present our proposed dynamics, as well as the measurements employed to analyze the results. Section 4 describes the results obtained and the associated discussion. Finally, in Section 5, we conclude the paper and present future works resulting from this study.

2. Related works

In this section, we present a short review of related studies and contrast them with our model. At the end of this section, we summarize the most relevant differences in order to highlight the contributions of the paper.

A confirmation bias can be understood as the seeking for expected information according to previous beliefs [33]. As a consequence, people can reinforce previous opinions and behaviors. This phenomenon is also central to govern the behavior of social networks' users and can promote the dissemination of fake news [27]. Other crucial elements that have been catching the attention of scholars are how the individuals are exposed to the information, which can be selective [30] or incidental [47]. In [47], the authors studied and compared these aspects with a focus on politics by using a survey collected in the United States. They found that the incidental exposition of information can lead the more active partisans to search for political content similar to their opinion. This behavior can increase the sharing of political information. Furthermore, a data-driven study shows that the selective exposure, when combined with the attention of the individuals, can affect the manner social network users consume news [9].

Other related features found in online social networks are polarization and echo chamber formation. The first is produced by the movement of many individuals to go from a moderate opinion to a more extreme point of view [42]. Echo chambers are groups in which individuals are in contact with opinions similar to their own beliefs. Many studies consider the relationship between the social network algorithm and polarization and the formation of echo chambers. For instance, [37] found evidence that information bias can be related to the formation of echo chambers. These concepts have been analyzed in data regarding climate policy in the US [22], to mention a relevant example. Interestingly, it has been observed that the opinion of few individuals, when spread on an echo chamber, could seem to be the opinion of many individuals [22].

Many other studies regarding echo chambers have been developed [1,5,48]. Some of them account for how algorithms indicate content in social networks [48], which play an important role in their dynamics. Depending on the configuration of our model, a sort of confirmation bias can be incorporated. In some cases, this feature could yield the formation of echo chambers. Different methods have been proposed to identify and quantify echo chambers [1,3,10,11,31]. In [1] techniques used to identify network communities have been employed. Furthermore, information regarding the individual's opinion and their neighbors can also be used [3,10,11]. Here, we use the measurement presented in [11] and propose two auxiliary measurements.

To understand the mechanisms that give rise to the previous phenomena in social networks, different models have been proposed [7]. In the model described in [6], under the assumption of homophilic interactions between individuals, the authors show that there is a tendency for fragmentation or echo chamber formation for flexible networks. Another approach extensively studied is the Sznajd model [43]. In an adapted version, in which rewirings are considered, echo chambers can emerge [5]. Furthermore, the underdog effect [46] has also been incorporated into this model [1] and could promote the formation of echo chambers. This underdog effect can be understood as the inclination of individuals to support the most disliked option. Finally, another possibility is the Axelrod model, which accounts for the process of cultural diffusion and has been employed in studies related to opinion polarization [38].

A frequently used type of model for representing the continuous opinions is the bounded confidence model [12]. In this approach, pairs of individuals can change their opinions if they are connected and if the difference between their opinions is smaller than a defined parameter. Modifications of this model were proposed to study the confirmation bias on the polarization of social networks [13]. In more detail, [13] introduced the use of rewiring and unbounded versions of this model. Interestingly, both types of variation could give rise to polarization. Ref. [39] also proposed a variation of the bounded confidence model in order to understand specific mechanisms of social media. The DeGroot-Friedkin model [16] has also been adapted to incorporate the confirmation bias [29]. In [29], the authors investigated this characteristic with the competitive information spreading.

In these opinion dynamics, the influence of one individual on others is typically modeled by considering that the individual tries to spread his/her own opinion. Here, since the posts do not express the individuals' opinions, their friends are not directly influenced by their opinions. Specifically, the social network algorithm considers the individual's opinion to distribute the post, but its friends never receive the opinion itself, just the post. In contrast to the bounded confidence model, here, the transmission and distribution functions control if the individual receives the information. Furthermore, we incorporate the possibility of rewiring, also considering a probability function and not a threshold.

Given these many findings regarding opinion dynamics, we put together many of these characteristics in a single dynamic model. The summary of contributions of this paper are as follows:

- We propose a framework that incorporates both the behavior of users and possible algorithmic bias of the social network;
- In this dynamic, the individuals are influenced by the posts and not by the opinions of others.
- We propose two complementary measures to analyze the bimodality of opinions (bimodality index and the distribution balance);
- Depending on the range of parameters, the dynamic converged to a varied range of outcomes, which include echo-chamber formation and consensus of opinions;
- In contrast to the majority of the models, we also studied the dynamical transient. For specific configurations, we found that opinions can exhibit a bimodal distribution but converge to a consensus;
- We identified scenarios in which the opinion distribution is bimodal but without echo-chamber formation. This case can be related to debates in social networks since users with different opinions keep communicating;
- For particular scenarios, even without the inclusion of friendship rewiring, our dynamic can converge to echo chambers;
- A comparison with a real social network shows similarities with the simulated outcomes.

3. Opinion dynamics

This section describes the proposed dynamics and depicts the experiment design and how the results are analyzed.

3.1. Proposed framework

Our model represents a social network, in which the users (individuals) produce *posts*. Then, the social network algorithm (post distribution) selects the neighbors (friends) that will receive the post. Finally, the opinions of the selected neighbors might change, and the users that strongly disagree can change their friendship (rewiring).

Specifically, each node, i, represents an individual that has an opinion, $b_i \in \mathbb{R}$, with $-1 \leqslant b_i \leqslant 1$. The network edges represent the individual's friendships. The dynamic starts with the opinions randomly initialized following a uniform probability distribution.

At each iteration, a node i is randomly selected, and a post $\mathscr{P}$ is created. In order to define the opinion associated with the post, a number, θ, is randomly generated with uniform probability ($-1 \leqslant \theta \leqslant 1$). Next, a transmission probability, $P_t(x)$, is used to define if the individual i will post $\mathscr{P}$. This probability function is computed according to the difference between the post and the individual's opinion ($x = |\theta - b_i|$). The functions employed in this paper are described in Section 3.3. Additionally, if the individual i produces a post, a function, $P_d(y)$, is calculated to define the probabilities of the neighbors of the node i to receive the information. In this case, the function is calculated for all edges, (i, j), connected to i, where $y = |b_i - b_j|$. This action is associated with how the social network algorithm acts. The probability functions used here are shown in Section 3.3. After this action, another probability could be associated with the individuals to define if they are active in the social network. However, here we consider this probability as one. In other words, we considered that the users see at all posts received.

The opinions of the individuals that receive a post can result in attraction or repulsion. More specifically, for each individual, j, the probability of being attracted is

$$\xi_j(\theta, b_j) = 1 - \frac{|\theta - b_j|}{2}. \tag{1}$$

As observed in [21,32], people update their opinions on topics after interacting or in a discussion and can become more polarized while doing so. In our model, if the individual is attracted, its opinion b_j turns out to be closer to b_i by a Δ amount, otherwise is repulsed and turns out to be farthest to b_i by a Δ amount. It has also been observed that, when confronted with opposing views, people in social media can become more extreme in their opinions [2].

The last action of our dynamic allows an individual to unfollow a given friend and connect to another. This step is henceforth called *rewire*. The unfollowing in social networks have been extensively studied [24]. For instance, the study developed by [24] indicates that Twitter users are less likely to unfollow friends who have acknowledged them. Inspired by this study, here, if an individual is repulsed by a neighbor, the rewire can happen according to a given function, $P_{\text{rewire}}(x)$, for $x = |b_i - b_j|$, see Section 3.3.

The previous steps are repeated n iterations, in which n should be big enough to let the dynamics reach a steady state. In order to automatically execute many times the same program with all of the parameters presented in this section, we use the software *GNU Parallel* [44].

3.2. One step example

In order to organize all the procedures presented in the previous section, we describe a complete example of one iteration of our dynamic, as shown in Fig. 1. The dynamic starts with a network node, i, selected at random, the green node in Fig. 1(a). In this case, the opinion of the chosen individual is $b_i = -0.9$. Next, a number $(-1 \leqslant \theta \leqslant 1)$ is randomly generated according to a uniform distribution, representing a post created by the green individual. In this example, the post is given by $\theta = -0.4$.

In the following, the probability of *post transmission* is calculated according to the difference between b_i and θ (see Fig. 1 (b)). The obtained transmission probability is $P_t^{\text{pol}}(x) = 0.5$. We illustrate our example by employing the polarized function P_t^{pol}, more details regarding the possibilities of transmission probabilities are described in Section 3.3.1. In this example, the post is transmitted (see Fig. 1(c)).

After transmitting the post, the dynamics determine which friends of i should receive the post. For this purpose, the distribution probability is calculated by considering the difference between the opinions of i and its friends. Fig. 1(d) illustrates how the probabilities are calculated (for more details regarding the possibilities of *post distribution*, see Section 3.3.2). According to these probabilities, the social network algorithm chooses if the post is seen by the other users (blue). In our example, Fig. 1(e), a single friend does not receive the post (see the red x).

After receiving the post, the individuals change their opinions according to it. The attraction probability is calculated for all individuals who receive the post, as shown in Fig. 1(f). Note that the attraction probability considers the information received (θ) and the opinion of the individual. We choose this strategy because the individual only has contact with the post but not with the belief of the other individual. Furthermore, we choose a probability that decreases linearly for the sake of simplicity. As a complement, the non-attracted individuals are repulsed. Fig. 1(g) illustrates the opinion changes, represented in bold. The opinions shown in black (b_3) and red (b_1 and b_4) represent the individuals attracted and repulsed by the agent i, respectively.

As the last action, possible changes of friendship are calculated according to the rewiring probability. This probability is computed for all repulsed users according to the difference y, as shown in Fig. 1(h). More information regarding the choice of this probability is provided in Section 3.3.3. In this example, a single edge is rewired (see Fig. 1(i)). The new connection is chosen with the same probability of reconnection for all remaining network nodes.

3.3. Adopted configurations

In this subsection, we describe the chosen probability functions and the rationale for our choices.

3.3.1. Post transmission functions

In the case of post transmission, we considered three distinct possibilities of $P_t(x)$. The first possibility is defined as

$$P_t^{\text{pol}}(x) = \cos^2\left(x\frac{\pi}{2}\right), \tag{2}$$

where $x = |\theta - b_i|$. In this case, the individual tends to post both the most similar and most different with respect to its own belief. This polarized function simulates the scenarios in which the users post pieces of information he/she agrees or disagrees with. In the latter, this possibility represents the cases in which a user's post reflects an opinion against the content. Furthermore, the highest probabilities of posting divergent opinions are reached only if the individuals' opinions, b_i, are close to the extremes, -1 or 1, resulting from the maximum possible value of x. For instance, if $b_i = 0$, the maximum x value is 1, consequently, $P(1) = 0$.

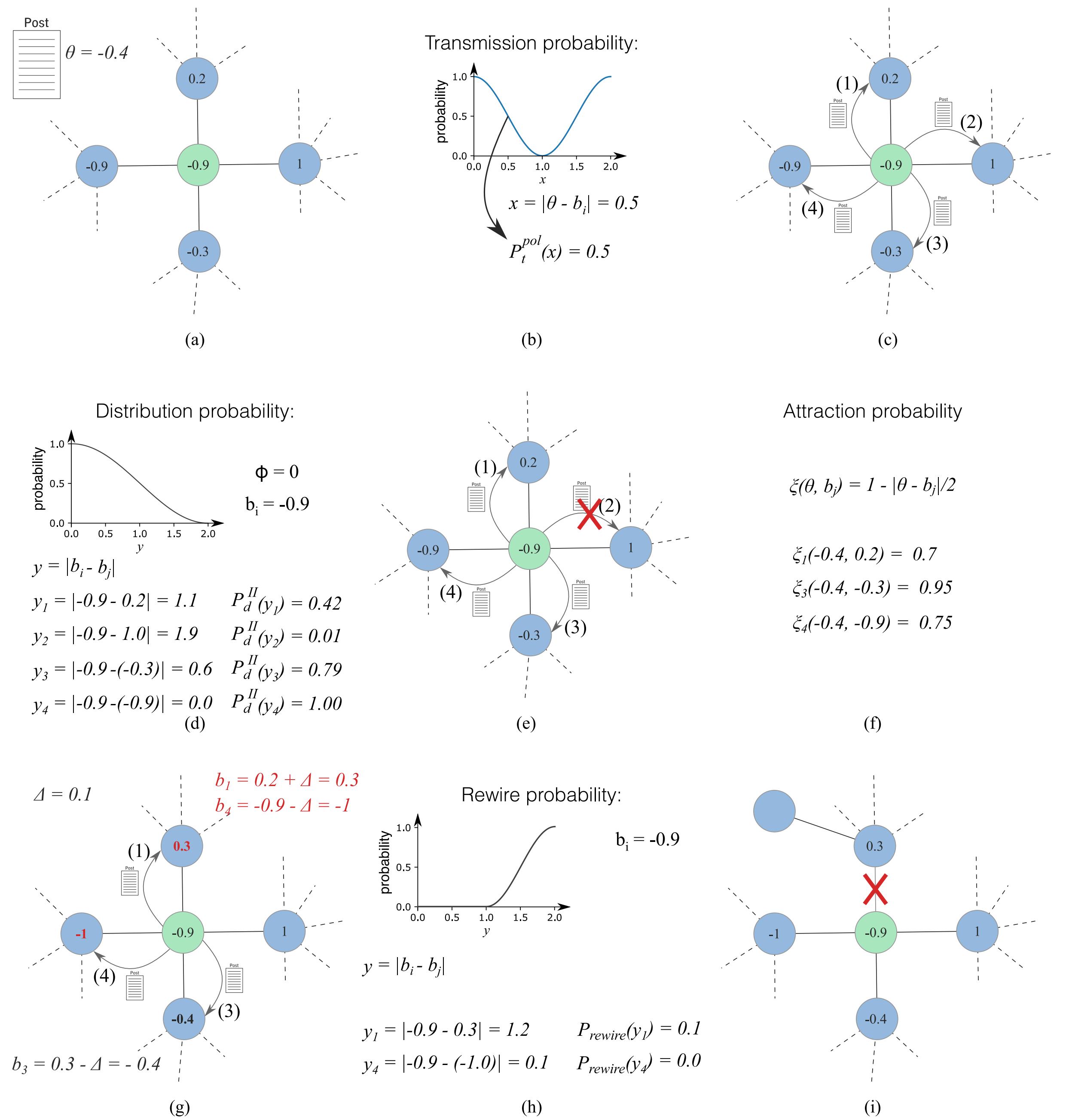


Fig. 1. Example of one step of the proposed dynamics. (a) The post is randomly generated. (b) The transmission probability, $P_t(x)$, is calculated. (c) and (d) *post distribution* calculations. (e) the algorithm chooses if the other users see the post. (f) Attraction probability is calculated. (g) The opinions are changed. (h) The rewiring probability is computed for all repulsed users. (i) the edges can be rewired.

We also considered the users that have a much higher probability of posting information similar to their own opinions and cannot post contrarian information, which can be modelled as

$$P_t^{\mathrm{sim}}(x) = \begin{cases} \cos^2\left(x\frac{\pi}{2}\right), & \text{if } x \leqslant 1 \\ 0, & \text{otherwise.} \end{cases} \tag{3}$$

The third tested strategy is the uniform probability, as follows

$$P_t^{\mathrm{uni}}(x) = 1. \tag{4}$$

This probability simulates the cases in which the users produce posts without taking care of the information. More specifically, all the created posts are spread by the users. This case can also be used as a null model. Fig. 2 illustrates the transmission functions considered.

In addition to the three described possibilities, we considered the combination of them, $P_t^{\mathrm{all}}(x)$. A function is randomly chosen for each individual, with equal probability for all possibilities. For the sake of simplicity, each individual has a fixed behavior. More specifically, the chosen function does not change during the execution of the dynamics.

3.3.2. Post distributions

For the post distribution, we considered several possibilities for the probability functions. The first equation is defined as follows

$$P_d^{\mathrm{I}}(y) = \cos^2\left(y\frac{\pi}{2} + \phi\right), \tag{5}$$

where the parameter ϕ is a real number that controls the starting point of the cosine-squared function and $y = |b_i - b_j|$[1], in which b_i and b_j are the opinions of the individual i and its given neighbor j, respectively. We considered another version, in which the probabilities vary smoother than in Eq. (5), as follows

$$P_d^{\mathrm{II}}(y) = \cos^2\left(\frac{y}{2}\frac{\pi}{2} + \phi\right). \tag{6}$$

Fig. 3 illustrates the relationship between the parameter ϕ and the probability functions. In both cases, the functions can represent a range of algorithms that spread information from a polarized to a depolarized. As the third case, we also employed a null model, that transmits uniformly the information, which is defined as

$$P_d^{\mathrm{III}}(y) = 1. \tag{7}$$

One of the objectives of social networks is to engage users. Among the possibilities of post distribution, given by the parameter ϕ, some choices could be justified to gain the users' attention, including the information bias. Additionally, it is expected that disagreements can promote engagement. Specifically, in our model, the algorithm can promote antagonism if individuals receive posts they strongly disagree with. For this reason, we test both choices, post-distribution giving rise to information bias and disagreement. For instance, information bias and disgreements can be set in our model using P_d^{II} with $\phi = 0$ and $\phi = \frac{\pi}{2}$, respectively, see Fig. 3(b). For the scenarios in which both behaviors occur at the same time, P_d^{I} with $\phi = 0$ or P_d^{II} with $\phi = \frac{\pi}{4}$ can be employed. Furthermore, by changing ϕ, intermediate possibilities can be explored. For instance, for P_d^{I} with $\phi = 3\frac{\pi}{4}$, higher probabilities are obtained for $y < 1$, which can favor the information bias.

3.3.3. Rewiring configurations

In online social networks, individuals can decide to unfollow others with very different opinions. We modeled this behavior by allowing the rewiring of connections only between individuals that strongly disagree. With this purpose, we adopt the following rewiring probability function

$$P_{\mathrm{rewire}}(y) = \begin{cases} \cos^2\left(y\frac{\pi}{2}\right), & \text{if } y > 1 \\ 0, & \text{otherwise}, \end{cases} \tag{8}$$

where y is defined by the difference between the opinions of the individuals i and j ($y = |b_i - b_j|$). Fig. 4 illustrates this function. In order to compare this scenario with a fixed network structure, we also considered the dynamic without the possibility of rewiring ($P_{\mathrm{rewire}}(x) = 0$). Thus, the effect of rewiring can be easily evaluated.

3.4. Opinion polarization analysis

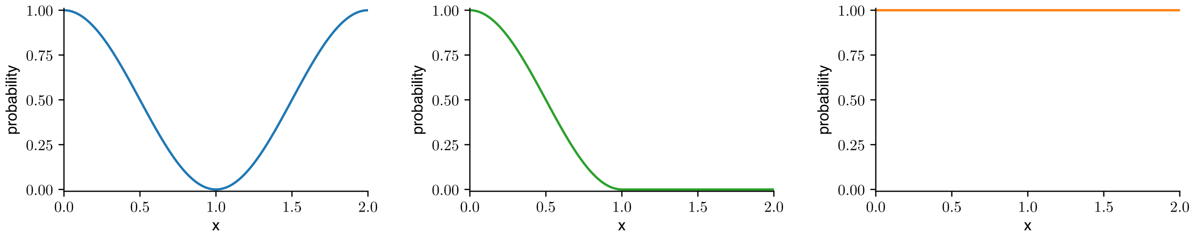
In this section, we describe how the results are analyzed. First, we present the measures used to account for the opinion distributions. Next, we also considered information regarding the relationship between structure and dynamics.

3.4.1. Analysis of opinion distributions

By employing the bimodality coefficient, BC [36], we can quantify the level of polarization of the opinions. This measurement is computed from the opinion distributions and is defined as

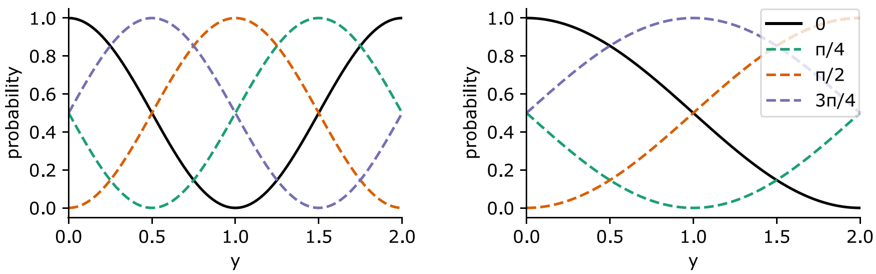
$$BC = \frac{g^2 + 1}{k + \frac{3(n-1)^2}{(n-2)(n-3)}}, \tag{9}$$

[1] This process is similar to a homophily, however, homophily is an outcome of people decisions, while here we are modeling the posts suggestion by the social network algorithm. Homophily could be an extension to our model.



(a) P_t^{pol} (b) P_t^{sim} (c) P_t^{uni}

Fig. 2. Transmission probability functions.



(a) P_d^{I} (b) P_d^{II}

Fig. 3. Examples probability functions for post distribution.

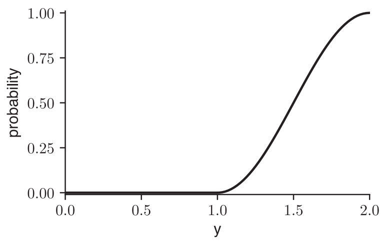


Fig. 4. Function employed to define the rewiring probability.

where n is the number of samples, and g and k are the *skewness* [23] and *kurtosis* [23] of the analyzed distribution, respectively. Furthermore, it was empirically found that for $BC_{\text{critic}} = 5/9$ the distribution tends to be uniform, and for values higher and lower than BC_{critic}, it tends to be bi-modal and uni-modal, respectively [36]. However, in our experiments, BC did not perform well in probability distributions with unbalanced modes. In order to complement this descriptor, we propose a measure that is henceforth called *balance*. First, we divide the resulting opinions into two sets, s_1 and s_2, which contains values lower and greater than zero, respectively. From these sets, we compute the balance, as follows

$$\beta = \frac{\min(c_1, c_2)}{\max(c_1, c_2)}, \tag{10}$$

where c_1 and c_2 represent the number of nodes in s_1 and s_2, respectively.

3.4.2. Relationship between structure and dynamics

Although the bimodality coefficient accounts for the opinion distribution's shape, it cannot quantify individuals' relationships. So, we make another analysis. An interesting characteristic that can be found in opinion dynamics is the presence of echo chambers. Here, we consider the metric used in [3,10,11] to identify if our dynamics leads to the formation of echo chambers, which consists of a density map of the individuals' opinion, b, against the average opinion of its neighbors, b^{NN}. So, when distinct groups are located in the first and third quadrants of the map, for the dynamics converge to echo chambers.

However, other interpretations are possible. In order to illustrate some possible resulting density maps, we created three case examples (see Fig. 5). In Fig. 5(a), a single peak expresses *consensus*, in which all individuals are connected to others with similar opinions. Fig. 5(b) illustrates the formation of echo chambers. Note that the density map in Fig. 5(b) is a necessary condition for the echo chamber formation. However, it is not a sufficient condition as we also need homophilic interactions [3]. Another possible scenario is depicted in Fig. 5(c). In this case, the individuals are essentially connected to others that have the same average opinions. Henceforth, we refer to this scenario as *diverse* since the individuals communicate with others that have diversified opinions. Note that this map describes the initial configuration of opinions of the dynamics. A combination of the already presented density maps can form other possible outcomes. For all of the case examples, the border effect is found in the density map.

4. Results and discussion

Next, we present our results, starting from analysis of the opinion distributions, and varying the post transmission parameters, reception, and rewiring. We also analyze a real scenario in terms of the proposed methodology.

4.1. Model parameters

For each combination of parameters, we executed the dynamics 50 times. In the case of phase ϕ, we considered 33 values between 0 and 2π. The number of iterations was manually defined, guaranteeing that the dynamics reached a steady-state (see Supplementary material S1). More specifically, for the majority of the cases, we considered that the dynamic reached the steady state when there were no significant variations of BC along time. For more details, see Supplementary material S1. Since Δ defines the step for the opinion change, we chose the smaller possible value considering the computational cost of the simulations. In the following results, we adopted $\Delta = 0.1$.

Regarding BC and β, we considered the average and standard deviation. In the case of the density maps of b against b_{NN}, we show the results of single executions because an average map could hide interesting results. For instance, in scenarios where the distribution of b converges to all values close to 1 or -1, the resultant map displays a single lobe. However, the same result could seem to be polarized by considering an average map displaying two lobes.

For the first part of our analysis, described in Section 4.2, we considered the possibility of rewirings of friendships, using Eq. 8. Since the network structure varies with time, we considered only Erdős-Rényi (ER) [14] networks with approximately 1000 nodes, and $\langle k \rangle = 8$. In this network, the connections between nodes are defined according to a probability of connection, p. This parameter was set to give rise to networks approximately with the desired average degree. Furthermore, tests were also executed for $\langle k \rangle = 4$, and the results were similar (results are not shown). Additionally, we varied all possible combinations of parameters, as presented in Sections 3.3.1 and 3.3.2.

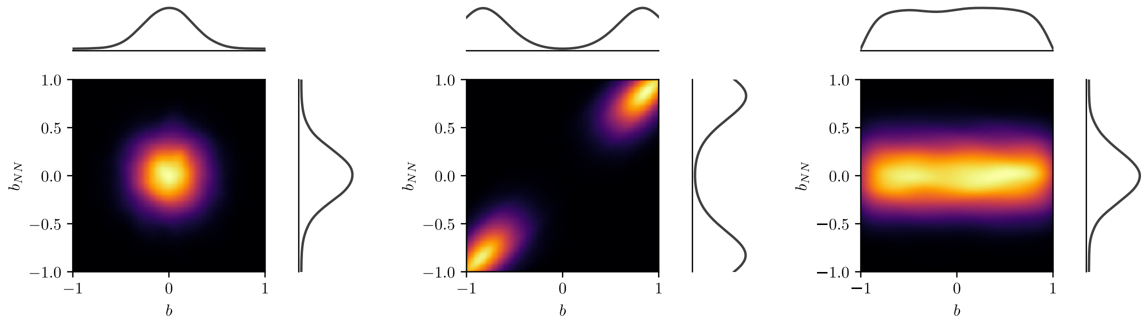
For the dynamics without the possibility of rewiring (described in Section 4.3), the sets of parameters previously described were also employed, except for the rewiring probability function. Additionally, we compared our results with an SBM (Stochastic Block Model) [20]. To this end, the probability of connection between the communities was set to 8×10^{-5}. The comparison with other network models is Section S3 of the Supplementary material.

4.2. Analysis of the dynamics

Many different behaviors can be observed, as illustrated in Fig. 6. For both types of functions for post distribution (P_d^I and P_d^{II}), the type of transmission function that lead to higher values of BC is P_t^{uni}, for values of ϕ close to 1.5. This result means that, according to our model, if the users of social network tend not to express a strong opinion on the posts, the algorithm (via its post distribution) can lead the opinion distributions to be polarized. For P_t^{pol}, P_t^{sim}, and P_t^{uni}, BC presents a relatively high standard deviation, as can be seen in the shaded regions of Fig. 6. These values are obtained because of the fluctuations of the BC for each execution of the dynamic even after the steady-state is reached. Furthermore, the result tends to the average BC. By considering P_d^{III}, for all transmission functions, the b distributions were not found to be bi-modal (more information is shown in Supplementary material S2.1). Again, the function from of the post distribution is found to play an essential role in the polarization.

Also, considering the opinion distributions, for the majority of the results, balance (β) is found to be high (more details are shown in S6 of the Supplementary material). However, for almost all combined functions, β tends to become lower when ϕ is close to $\pi/2$, except for P_t^{uni}. For these values, the distributions can be considered unbalanced.

Next, we describe the relationship between the proposed opinion dynamics and network structure. First, we analyze the density maps of b against b_{NN}. We did not show average density maps because, in some cases, this average could make interesting outcomes less discernible. In general, the obtained density maps indicate results varying from consensus to echo chambers. Items VI, VII, and VIII of Fig. 6 illustrate three different levels of consensus, where the more well-defined scenario is found for $P_d^I, P_t^{\mathrm{sim}}$, and $\phi = 1.47$ (item VI). The echo chamber formation was found only when we employed P_t^{uni}, for both types of post distribution (P_d^I and P_d^{II}) and ϕ values close to 1.5. See an example in item IV of Fig. 6. Furthermore, to lead to echo chambers, the network structure changes and gives rise to separated communities. However, there is no significant



(a) Consensus **(b)** Echo chamber, **(c)** Diverse

Fig. 5. Synthetic examples of possible resulting density maps of opinions b against the average neighbor's opinions b_{NN}, in which the lighter color represents the larger number of users. Furthermore, we depict the probability functions of b and b_{NN} at the top and right sides of the maps, respectively.

change in the degree distributions [2]. Examples of these degree distributions are shown in Fig. S7 of the Supplementary material.

Other interesting results are shown in items I and II of Fig. 6. In the first, BC is slightly higher than BC_{critic}. So the distribution of b tends to be similar to a uniform but with a bi-modal inclination. However, most of the samples are located in the first and third quadrants, which indicates the echo chamber's tendency. In contrast to this result, for item II of Fig. 6, the opinion distribution is bi-modal, but there is no tendency for echo chambers. This density map is more similar to a diverse scenario but with a bi-modal in the distribution of b.

4.3. Analysis of the dynamics without rewiring

In order to better understand the proposed model, we test the dynamics without the possibility of echo chambers. Similar to Section 4.2, here we consider many different combinations of parameters and analyze the steady state of the dynamic. More details are provided in the Supplementary information S2.2.

First, we use the uniform version of the post distribution. In this case, the resulting opinion distributions reflect the post transmission functions. More specifically, for $P_t^{\text{pol}}, P_t^{\text{sim}}$, and P_t^{uni}, the dynamics converged to bi-modal, uni-modal, and uniform distributions of b, respectively. Interestingly, for P_t^{all} the emerging b distribution has an intermediate value of BC, which is slightly lower than BC_{critic}. We also fixed the function of post transmission to the uniform (P_t^{uni}). In the case of P_d^{I}, for $0 \leqslant \phi < \pi/2$, the resulting distribution of b is typically uni-modal, and for $\pi/2 \leqslant \phi < \pi$ bi-modal. A similar result is found for P_d^{II}, but here the uni-modal distributions are found for $\pi/4 \leqslant \phi < 3\pi/4$.

Considering combinations of post transmission and post distribution, other interesting results are found. In contrast to previous results, the BC levels are much higher for some configurations with P_t^{pol}. Furthermore, when we slightly vary ϕ, more abrupt changes of BC are obtained, which can be explained by variations of the opinion distributions from uni-modal directly to bi-modal, or vice versa. These changes happen only for low values of β. More details regarding these analyses are shown in the Supplementary information S2.2.

Because of the large number of reception possibilities, in the remaining of this subsection, we restrict the analysis to two fixed values of ϕ. These values were chosen in line with the variations of BC and β (for more information see Fig. S11 of the Supplementary material). We consider ϕ equal to π and 1.473. By comparing both reception functions, in the case of $\phi = \pi$, the bimodality coefficient does not differ considerably, and high values of β are found. On the contrary, high differences for BC are found for $\phi = 1.473$. In this case, for all tested parameters, except when we employed P_t^{uni}, the resulting distribution of b is found to be unbalanced. We also compared BC for several other networks. All in all, the results were similar (the complete analysis is available in the Supplementary material S3).

In what follows, we focus on the ER networks as they are the simplest and the results for other structures are similar. Fig. 7 shows some examples of density maps. Fig. 7(a) illustrates a scenario in which the opinions are polarized into two groups, but there are no well-defined echo chambers. More specifically, the lack of echo chamber formation is characterized by similar average opinions of the agent neighbors. In other cases, where the opinions of all of the agents are similar between themselves, a single group is found in the density map (see Fig. 7(b) and (c)). Interestingly, for the example of Fig. 7(b), the opinions converge to an extreme. This result is obtained because during the transient the distribution of b becomes bi-modal and converges to uni-modal, in which one of the two peaks increases while the other decreases, giving rise to a uni-modal distribution close to -1 or 1 (this effect is shown in Fig. S2 of the supplementary material). Furthermore, in Fig. 7(d), we pre-

[2] Degree is defined as the number of edges connected to a given node.

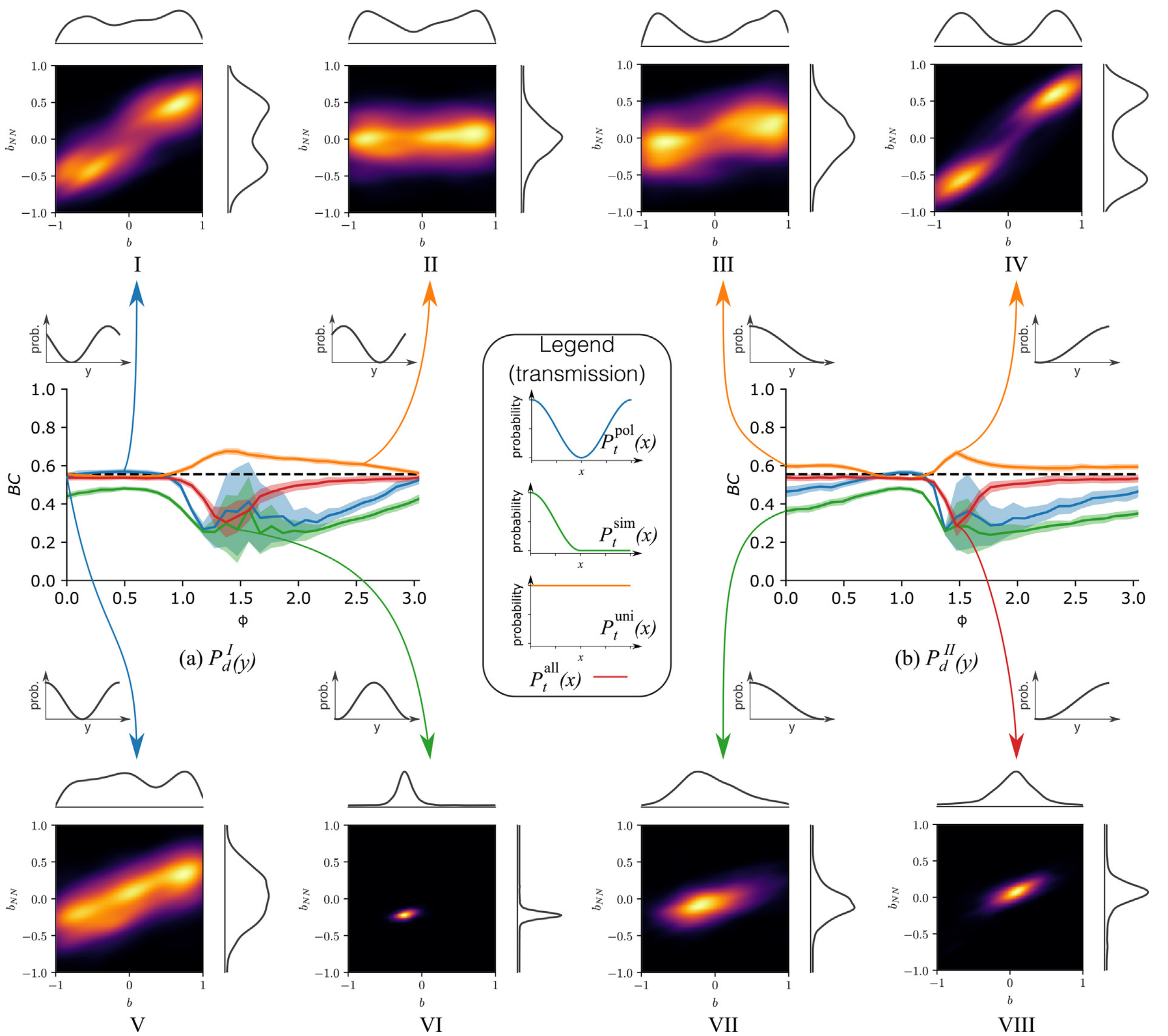
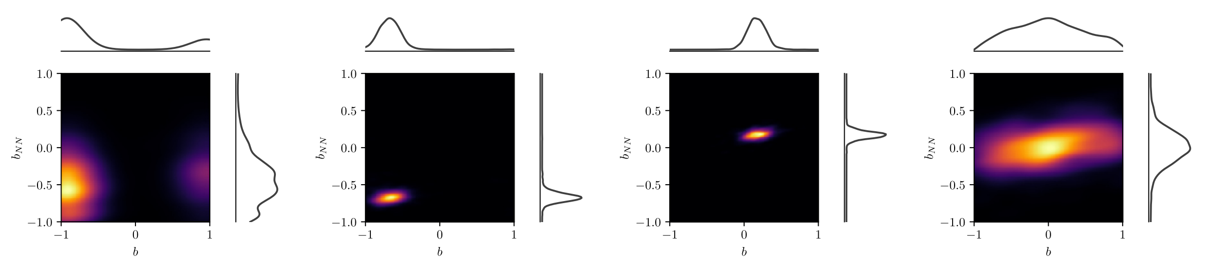
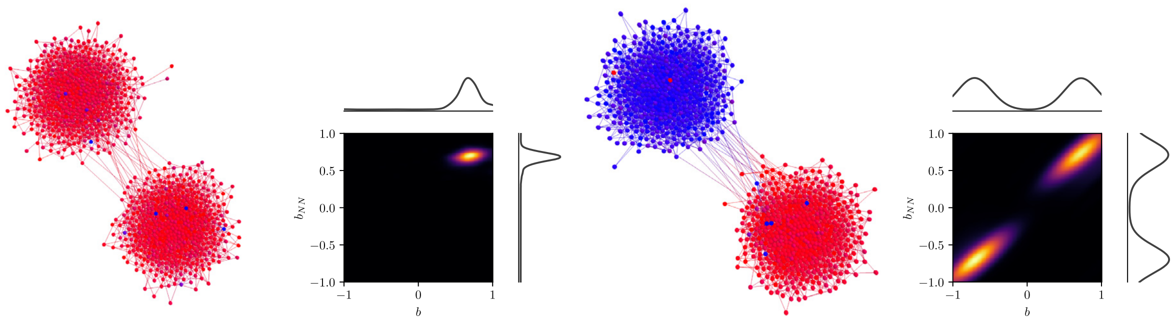


Fig. 6. The average and respective standard deviations of BC (bimodality coefficient) are shown in items (a) and (b). The horizontal dashed lines indicate BC_{critc}. For each item, we illustrate the measurements with four examples of density maps of b against b_{NN}. The plots close to the arrows represent the probability function used for the distribution. The density maps represent the following configurations: I- $P_d^I, P_t^{\mathrm{pol}}$, and $\phi = 0.49$, II- $P_d^I, P_t^{\mathrm{uni}}$, and $\phi = 2.65$, III- $P_d^{II}, P_t^{\mathrm{uni}}$, and $\phi = 0$, IV- $P_d^{II}, P_t^{\mathrm{uni}}$, and $\phi = 1.47$, V- $P_d^I, P_t^{\mathrm{pol}}$, and $\phi = 0$, VI- $P_d^I, P_t^{\mathrm{sim}}$, and $\phi = 1.47$, VII- $P_d^{II}, P_t^{\mathrm{sim}}$, and $\phi = 0$, and VIII- $P_d^{II}, P_t^{\mathrm{all}}$, and $\phi = 1.47$. These density maps illustrate typical results of a single execution of the dynamics.



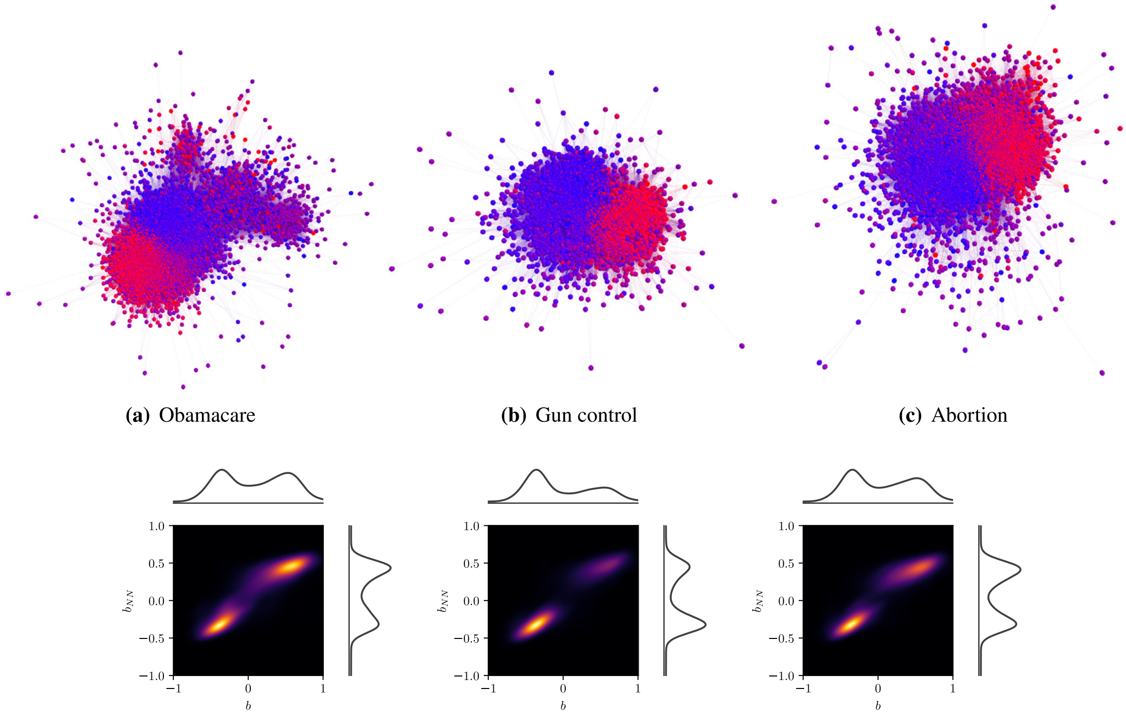
(a) P_t^{pol} and P_d^I with $\phi = 1.47$ **(b)** P_t^{pol} and P_d^{II} with $\phi = 1.47$ **(c)** P_t^{sim} and P_d^{II} with $\phi = 0.0$ **(d)** P_t^{pol} and P_d^{II} with $\phi = 0.0$

Fig. 7. Samples of density maps from the comparison between the b against b_{NN}. Lighter colors represent the denser regions. These results were measured from ER networks with $\langle k \rangle \approx 8$, without considering the possibility of rewiring.



(a) i **(b)** i **(c)** ii **(d)** ii

Fig. 8. Example of two samples (i and ii) of the resulting dynamic executed on a SBM, with P_t^{uni}, P_d^{II}, and $\phi = 1.47$. Items (a) and (c) display the SBM networks. The colors vary from blue to red, which represent left and right wings, respectively. In (b) and (d), we display the density maps of opinions (b) against the average opinion of the neighbors (b_{NN}), where lighter colors represent large numbers.



(d) Obamacare **(e)** Gun control **(f)** Abortion

Fig. 9. Real data visualizations. The three first panels display the Twitter network visualizations, in which the colors vary from blue to red. More specifically, blue and red represent left and right wings, respectively. The second line shows the respective density maps of opinions (b) against the average opinion of the neighbors (b_{NN}), in which lighter colors represent large numbers.

Table 1
Measures of bimodality coefficient, BC, and balance, β, obtained from real Twitter networks.

Subject	BC	β
Obamacare	0.60	0.80
Gun control	0.67	0.70
Abortion	0.60	0.91

sent an example of a density map with a tendency for the diverse scenario. However, individuals are more likely to be connected to neighbors with similar opinions.

Although echo chambers were not found in all the results presented in this subsection, there is the possibility to converge to echo chambers even without rewiring, depending on the network structure. In order to illustrate this possibility, we employ a SBM with two well-separated communities, as shown in Fig. 8. Interestingly, this network structure leads to bistable results. In particular, the dynamic can lead to both consensuses or echo chamber formations, with 46% of the 50 generated samples converging to echo chambers.

4.4. Characterization of real data

We used networks obtained from Twitter that are political examples of polarization in the United States, obtained in [18], and studied in [10]. The topics considered are: Obamacare (8703 nodes and 3,797,871 edges), gun control (3963 nodes and 1,053,275 edges), and abortion (7401 nodes and 2,330,276 edges). In all networks, nodes represent users, and directed connections were created according to followers (from following to follower). See the visualizations of the data in Fig. 9. Furthermore, online news organizations with political inclinations were used to define the individuals' opinions [10]. For the sake of simplicity, in our analysis, we considered the network as being undirected.

The opinions of the network users, shown in Fig. 9(a)–(c), seem to be separated. To compare these opinion distributions with the previous results, we apply the same methodology. Table 1 presents the measurements of BC and β. Comparing the real data with our previous experiments, we observe that in the cases in which we considered rewiring, BC was similar to the real networks' measures (see for instance Fig. 6, panel IV). However, we remark that β is lower for real cases than for our simulation. The abortion network is the case in which our model better approximates real data. This suggests that our model can be helpful both for a quantitative and qualitative analyses of real systems, providing mechanistic interpretations of real phenomena.

5. Conclusions

Due to the rise of social networks, researchers have been studying their opinion dynamics. Here, we proposed a model to study the conditions that can give rise to polarization. Our dynamic is based on some compartmentalized modules, and it is limited to simulating how new information is generated, as well as the individual's friends' reactions. First, in the post transmission phase, the user has contact with an external piece of information and chooses if he/she will post it according to a given probability. For the post distribution, the piece of information is analyzed by the social network algorithm. The opinions of the individuals that receive the post can move closer or farther depending on the content of the piece of news. For the repulsed individuals, a probability function controls if the individual will rewire friendship. Importantly, this study does not only contribute to a novel model but also proposes a new type of analysis. Here, we considered that the network structure is not the single aspect that limits the communication between individuals.

Several interesting outcomes have been observed. For instance, when we considered a rewiring probability, only with the uniform transmission, the dynamic gave rise to echo chambers. According to our model, if the users do not care about the information they post, the post distribution function can lead to polarization and the formation of echo chambers. Furthermore, three opinion outcomes have been observed: consensus, echo chamber, and diverse.

In some cases, high values of the balance and low values of the bimodality coefficient have been found. This result means that the dynamic converged to consensus, but with average values close to -1 or $+1$, which mainly happened when we considered the dynamic without the possibility of rewiring. Also, without including rewirings, polarization can be observed for a wider range of configurations. However, for the majority of these polarized cases, there is no echo chamber formation. One exception is a network with well-separated communities, which can converge to both consensuses or echo chambers for the same set of parameters. Additionally, we calculated and compared bimodality coefficient in synthetic and real data, with similar balance values being obtained, especially in the case of the Abortion data.

One of our model's current limitations is that individuals see received posts. In future works, a probability could be associated with this action, allowing posts to be discarded. Post distribution could also be adaptive and change over time. Furthermore, we considered only undirected networks, but our model can also be implemented using directed structures.

CRediT authorship contribution statement

Henrique Ferraz de Arruda: Conceptualization, Methodology, Software, Validation, Visualization, Investigation, Writing - original draft. **Felipe Maciel Cardoso:** Conceptualization, Methodology, Writing - review & editing. **Guilherme Ferraz de Arruda:** Conceptualization, Methodology, Writing - review & editing. **Alexis R. Hernández:** Conceptualization, Methodology, Validation, Writing - review & editing, Supervision. **Luciano da Fontoura Costa:** Conceptualization, Methodology, Writing - review & editing, Supervision. **Yamir Moreno:** Conceptualization, Methodology, Resources, Writing - review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank Gianmarco De Francisci Morales for sharing the Twitter networks and Michele Starnini for helping us with the data. Henrique F. de Arruda acknowledges FAPESP for sponsorship (grants 2018/10489-0 and 2019/16223-5). Guilherme F. de Arruda and Yamir Moreno acknowledge support from Intesa Sanpaolo Innovation Center. Luciano da F. Costa thanks CNPq (Grant No. 307085/2018-0) and NAP-PRP-USP for sponsorship. Yamir Moreno acknowledges partial support from the Government of Aragón and FEDER funds, Spain through grant ER36-20R to FENOL, and by MINECO and FEDER funds (grant FIS2017-87519-P). Research carried out using the computational resources of the Center for Mathematical Sciences Applied to Industry (CeMEAI) funded by FAPESP (grant 2013/07375-0). This work has been supported also by FAPESP grants 2015/22308-2. The founders had no role in study design, data collection, and analysis, decision to publish, or preparation of the manuscript.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, athttps://doi.org/10.1016/j.ins.2021.12.069.

References

[1] H.F. de Arruda, A. Benatti, F.N. Silva, C.H. Comin, L. da Fontoura Costa, Contrarian effects and echo chamber formation in opinion dynamics, Journal of Physics: Complexity 2 (2021) 025010.
[2] C.A. Bail, L.P. Argyle, T.W. Brown, J.P. Bumpus, H. Chen, M.B. Fallin Hunzaker, J. Lee, M. Mann, F. Merhout, A. Volfovsky, Exposure to opposing views on social media can increase political polarization, Proceedings of the National Academy of Sciences of the United States of America 115 (2018) 9216–9221.
[3] F. Baumann, P. Lorenz-Spreen, I.M. Sokolov, M. Starnini, Modeling echo chambers and polarization dynamics in social networks, Physical Review Letters 124 (2020) 048301.
[4] F. Baumann, P. Lorenz-Spreen, I.M. Sokolov, M. Starnini, Emergence of polarized ideological opinions in multidimensional topic spaces, Physical Review X 11 (2021) 011012.
[5] A. Benatti, H.F. de Arruda, F.N. Silva, C.H. Comin, L. da Fontoura Costa, Opinion diversity and social bubbles in adaptive sznajd networks, Journal of Statistical Mechanics: Theory and Experiment 2020 (2020) 023407.
[6] C. Blex, T. Yasseri, Positive algorithmic bias cannot stop fragmentation in homophilic networks, The Journal of Mathematical Sociology (2020) 1–18.
[7] C. Castellano, S. Fortunato, V. Loreto, Statistical physics of social dynamics, Reviews of Modern Physics 81 (2009) 591.
[9] M. Cinelli, E. Brugnoli, A.L. Schmidt, F. Zollo, W. Quattrociocchi, A. Scala, Selective exposure shapes the facebook news diet, PloS One 15 (2020) e0229129.
[10] M. Cinelli, G.D.F. Morales, A. Galeazzi, W. Quattrociocchi, M. Starnini, The echo chamber effect on social media, Proceedings of the National Academy of Sciences 118 (2021).
[11] W. Cota, S.C. Ferreira, R. Pastor-Satorras, M. Starnini, Quantifying echo chamber effects in information spreading over political communication networks, EPJ Data Science 8 (2019) 35.
[12] G. Deffuant, D. Neau, F. Amblard, G. Weisbuch, Mixing beliefs among interacting agents, Advances in Complex Systems 3 (2000) 87–98.
[13] M. Del Vicario, A. Scala, G. Caldarelli, H.E. Stanley, W. Quattrociocchi, Modeling confirmation bias and polarization, Scientific Reports 7 (2017) 1–9.
[14] P. Erdős, A. Rényi, On the evolution of random graphs, Publications of the Mathematical Institute of the Hungarian Academy of Sciences 5 (1960) 17–61.
[15] J. Fernández-Gracia, K. Suchecki, J.J. Ramasco, M. San Miguel, V.M. Eguíluz, Is the voter model a model for voters?, Phys Rev. Lett. 112 (2014) 158701.
[16] N.E. Friedkin, E.C. Johnsen, Social influence and opinions, Journal of Mathematical Sociology 15 (1990) 193–206.
[17] S. Galam, Minority opinion spreading in random geometry, The European Physical Journal B-Condensed Matter and Complex Systems 25 (2002) 403–406.
[18] K. Garimella, G. De Francisci Morales, A. Gionis, M. Mathioudakis, Political discourse on social media: Echo chambers, gatekeepers, and the price of bipartisanship, in: Proceedings of the 2018 World Wide Web Conference, 2018, pp. 913–922.
[19] C. Gracia-Lázaro, F. Quijandría, L. Hernández, L.M. Floría, Y. Moreno, Coevolutionary network approach to cultural dynamics controlled by intolerance, Physical Review E 84 (2011) 067101.
[20] P.W. Holland, K.B. Laskey, S. Leinhardt, Stochastic blockmodels: First steps, Social Networks 5 (1983) 109–137.
[21] D.J. Isenberg, Group polarization. A critical review and meta-analysis, Journal of Personality and Social Psychology 50 (1986) 1141–1151.
[22] L. Jasny, J. Waggle, D.R. Fisher, An empirical examination of echo chambers in us climate policy networks, Nature Climate Change 5 (2015) 782.
[23] S. Kokoska, D. Zwillinger, CRC Standard Probability and Statistics Tables and Formulae, Crc Press, 2000.
[24] H. Kwak, S. Moon, W. Lee, More of a receiver than a giver: why do people unfollow in twitter?, in: Sixth International AAAI Conference on Weblogs and Social Media, Citeseer, 2012..
[26] H. Liao, X. Li, M. Tang, How to process local and global consensus? a large-scale group decision making model based on social network analysis with probabilistic linguistic information, Information Sciences 579 (2021) 368–387.
[27] R. Ling, Confirmation bias in the era of mobile news consumption: the social and psychological dimensions, Digital Journalism 8 (2020) 596–604.
[28] J. Lorenz, Continuous opinion dynamics under bounded confidence: A survey, International Journal of Modern Physics C 18 (2007) 1819–1838.
[29] Y. Mao, E. Akyol, N. Hovakimyan, Impact of confirmation bias on competitive information spread in social networks, IEEE Transactions on Control of Network Systems (2021).
[30] S. Messing, S.J. Westwood, Selective exposure in the age of social media: Endorsements trump partisan source affiliation when selecting news online, Communication Research 41 (2014) 1042–1063.
[31] G.D.F. Morales, C. Monti, M. Starnini, No echo in the chambers of political interactions on reddit, Scientific Reports 11 (2021) 1–12.
[32] S. Moscovici, M. Zavalloni, The group as a polarizer of attitudes, Journal of Personality and Social Psychology 12 (1969) 125–135.
[33] R.S. Nickerson, Confirmation bias: A ubiquitous phenomenon in many guises, Review of General Psychology 2 (1998) 175–220.

[34] H. Noorazar, Recent advances in opinion propagation dynamics: a 2020 survey, The European Physical Journal Plus 135 (2020) 1–20.
[36] R. Pfister, K.A. Schwarz, M. Janczyk, R. Dale, J. Freeman, Good things peak in pairs: a note on the bimodality coefficient, Frontiers in Psychology 4 (2013) 700.
[37] W. Quattrociocchi, A. Scala, C.R. Sunstein, Echo chambers on facebook, 2016, Available at SSRN 2795110..
[38] S.M. Reia, U.P. Neves, Persistent agents in axelrod's social dynamics model, EPL (Europhysics Letters) 113 (2016) 18003.
[39] K. Sasahara, W. Chen, H. Peng, G.L. Ciampaglia, A. Flammini, F. Menczer, Social influence and unfollowing accelerate the emergence of echo chambers, Journal of Computational Social Science 4 (2021) 381–402.
[41] W. Su, X. Wang, G. Chen, Y. Yu, T. Hadzibeganovic, Noise-based synchronization of bounded confidence opinion dynamics in heterogeneous time-varying communication networks, Information Sciences (2020).
[42] C.R. Sunstein, Republic. com, Princeton University Press, Princeton, NJ, 2001.
[43] K. Sznajd-Weron, J. Sznajd, T. Weron, A review on the sznajd model'20 years after, Physica A: Statistical Mechanics and its Applications 565 (2021) 125537.
[44] O. Tange et al, Gnu parallel-the command-line power tool, The USENIX Magazine 36 (2011) 42–47.
[45] R. Urena, G. Kou, Y. Dong, F. Chiclana, E. Herrera-Viedma, A review on trust propagation and opinion dynamics in social networks and group decision making frameworks, Information Sciences 478 (2019) 461–475.
[46] J.A. Vandello, N.P. Goldschmied, D.A. Richards, The appeal of the underdog, Personality and Social Psychology Bulletin 33 (2007) 1603–1616.
[47] B.E. Weeks, D.S. Lane, D.H. Kim, S.S. Lee, N. Kwak, Incidental exposure, selective exposure, and political information sharing: Integrating online exposure patterns and expression on social media, Journal of Computer-Mediated Communication 22 (2017) 363–379.
[48] M. Wolfowicz, D. Weisburd, B. Hasisi, Examining the interactive effects of the filter bubble and the echo chamber on radicalization, Journal of Experimental Criminology (2021) 1–23.
[49] L. Xiong, X. Su, H. Qian, Conflicting evidence combination from the perspective of networks, Information Sciences 580 (2021) 408–418.
[50] Z. Yu, S. Lu, D. Wang, Z. Li, Modeling and analysis of rumor propagation in social networks, Information Sciences 580 (2021) 857–873.