Modeling algorithmic bias: simplicial complexes and evolving network topologies

Valentina Pansanella[1*], Giulio Rossetti[2] and Letizia Milli[2,3]

*Correspondence:
valentina.pansanella@sns.it

[1] Faculty of Science, Scuola Normale Superiore, Pisa, Italy
[2] Institute of Information Science and Technologies "Alessandro Faedo" (ISTI), National Research Council (CNR), Pisa, Italy
[3] Department of Computer Science, University of Pisa, Pisa, Italy

Abstract

Every day, people inform themselves and create their opinions on social networks. Although these platforms have promoted the access and dissemination of information, they may expose readers to manipulative, biased, and disinformative content—co-causes of polarization/radicalization. Moreover, recommendation algorithms, intended initially to enhance platform usage, are likely to augment such phenomena, generating the so-called *Algorithmic Bias*. In this work, we propose two extensions of the Algorithmic Bias model and analyze them on scale-free and Erdős–Rényi random network topologies. Our first extension introduces a mechanism of link rewiring so that the underlying structure co-evolves with the opinion dynamics, generating the *Adaptive Algorithmic Bias* model. The second one explicitly models a peer-pressure mechanism where a majority—if there is one—can attract a disagreeing individual, pushing them to conform. As a result, we observe that the co-evolution of opinions and network structure does not significantly impact the final state when the latter is much slower than the former. On the other hand, peer pressure enhances consensus mitigating the effects of both "close-mindedness" and algorithmic filtering.

Keywords: Opinion dynamics, Complex networks, Algorithmic bias

Introduction

The tendency to observe political polarization (McCarty 2019; Fiorina and Abrams 2008) on online social networks (Conover et al. 2011; Vicario et al. 2016; Bessi et al. 2016)- e.g., online discourse's tendency to divide users into opposite political factions not aiming at reaching any form of synthesis—has attracted a great deal of attention from different fields in recent years. Polarization not only emerges in political debates but also characterizes various controversial topics (Drummond and Fischhoff 2017; https://thepolarizationindex.com/), and, in some cases, it may affect policymaking (McCarty 2019) and society. It was argued (Hague and Loader 1999) that the advent of the Internet and social media, guaranteeing free access to a massive amount of information, would be a boon for democracy. Nonetheless, recent studies (e.g., Hills 2019) underlined that such a vast and unfiltered access to information could be—at least to some extent—harmful. To enforce such interpretation, in 2020, WHO coined the term "infodemic", a neologism describing a situation where a considerable quantity of false and misleading information (e.g., during a disease outbreak) (Cinelli et al. 2020) may generate social issues (e.g.,

causing severe repercussions on public health systems). Besides infodemics, information proliferation (Hills 2019)—i.e., the capacity to access and contribute to a growing quantity of information—is reducing the quantity and quality of content many people engage with. Indeed, it is evident that individuals have a limited amount of time and attention (Hogan 2001) to dedicate to collecting, processing, and discussing information. Unfortunately, the choice of which users/content to engage with is rarely made to pursue a balanced information diet; instead, it is a process highly affected by preexisting cognitive biases (Benson 2016). People tend to concentrate on the pieces of information that confirm their beliefs and ignore details that may contradict these (Festinger 1957; Knobloch-Westerwick et al. 2020). Moreover, there is evidence that when people are presented with different news articles to read, they tend to make their selection based on anticipated agreement, i.e., focusing on news sources they know to be closer to their political/cultural leaning (Iyengar and Hahn 2009).

These cognitive tendencies are exacerbated by some peculiar mechanisms of online platforms, such as recommender systems and algorithmic filtering: technological systems that implicitly create a polarizing feedback loop by further reducing the amount of diversity in the user experience (Anderson et al. 2020)—and possibly contributing to the creation and maintenance of echo chambers (Sunstein 2007) and filter bubbles (Pariser 2011). Although personalization is essential in information-rich environments to allow people to find what they are looking for and increase user engagement, there is great concern about the negative consequences of algorithmic filtering both from institutions and scientists (Sîrbu et al. 2019).

Algorithms recommending new people on social networking sites may consider the similarity of user-created content (Chen et al. 2009) again possibly reinforcing the underlying system polarization. Despite belief-consistent selection and confirmation bias playing a critical role in opinion formation, the diversity of content/sources encountered during daily activities is not the only driver of such complex realities. Peer pressure-like phenomena play a role in shaping people's opinions (Asch 1956; Haun and Tomasello 2011) and therefore should be considered when addressing the study of how public opinion evolves in social networks. People are likely to experience social pressure in both face-to-face and digital interactions. For example, suppose three individuals are mutual friends, and there is a disagreement on a particular topic. In that case, the majority opinion within the group will likely prevail, and the minority will adopt the majority opinion. Within social networks like Twitter, users participate in binary opinion exchanges with other users, i.e., through direct messaging, which can be modeled as binary interactions. However, the possibility of sharing tweets and engaging in public discussions opens the question of how participants can be influenced by others' opinions expressed in the thread and, in turn, influence their peers' opinions.

Understanding how different social mechanisms may influence the direction of public opinion and the levels of polarization in society has always been a crucial task, and a great challenge for computational social scientists (Conte et al. 2012). Unfortunately, empirical studies on how opinions form and evolve—influenced by environmental and sociological factors—are still lacking (Peralta et al. 2022): Indeed, if on one side moving toward data-driven analyses is necessary, on the other, models are essential to comprehend causes and consequences within controlled scenarios. Unfortunately, classic

opinion dynamics models are often very simplistic and cannot capture the complexity of the observed phenomenon.

In the last twenty years, there was a proliferation of opinion dynamics models that aimed at including more and more characteristics of real social systems and their properties, primarily using the classical models (Holley and Liggett 1975; Castellano et al. 2009; Sznajd-Weron and Sznajd 2000; Degroot 1974; Friedkin and Johnsen 1990, 1999; Deffuant et al. 2001), as baselines. Indeed, "digital era" specific characteristics are being included in recent opinion dynamics models, i.e., algorithmic personalization (Maes and Bischofberger 2015; Sîrbu et al. 2019; Peralta et al. 2021a, b; Perra and Rocha 2019) to understand what changes this new world brought into the way public opinion is shaped; however, several others are still missing making thus leaving room for more accurate modeling. Among such often neglected peculiarities, we can list the temporal dynamic of social interactions. Not only do network topologies evolve, but often this evolution is co-dependent on the dynamic process taking place over the networks, such as opinion exchange (McPherson et al. 2001). For this reason, recent efforts focus on studying opinion dynamics on dynamic/adaptive networks (Sasahara et al. 2019; Kan et al. 2021; Holme and Newman 2006; Iñiguez et al. 2009) also in the context of the Deffuant–Weisbuch and other bounded confidence models (Kozma and Barrat 2008; Kan et al. 2021; Sasahara et al. 2019), describing the effects of the evolution of the underlying network structure on the final state of the population and explaining the effects of the opinion exchange on the structure topology.

Despite group phenomena being present in classical opinion dynamics models (see (Galam 2002)), it has been recently recognized the importance of using higher-order structures to explain and predict collective behaviors that could not be described otherwise (Battiston et al. 2020; Hickok et al. 2022)—e.g., peer-pressure.

Moving from the results discussed in (Pansanella et al. 2022) where static network models and binary interactions were assumed, we aim to study the effects of adaptive networks—where the dynamic of the network depends on the opinion dynamics—and higher-order interactions have on opinion formation and evolution when in the presence of a filtering algorithm. To such an extent, we focus our analysis on the same network models employed in Pansanella et al. (2022), namely Erdős–Rényi (Erdás and Rényi 1959) and Barabási–Albert (Barabási and Albert 1999). Adopting such controlled environments, used to simulate the social structure among a population of interacting individuals, we analyze the behaviors of the two extensions of the Algorithmic Bias model (Sîrbu et al. 2019) and discuss the role of arc rewiring towards like-minded individuals and peer-pressure within 2-cliques.

The paper is organized as follows. In "Methods" section we introduce the Algorithmic Bias model and the two extensions and describe our experimental workflow. "Results and discussion" section discusses the main finding of our simulations, finally "Conclusions" section concludes the paper while opening to future investigations.

Methods

Algorithmic Bias: from Mean-field to Complex Topologies.

Online social networks have become the primary source of information and an excellent platform for discussions and opinion exchanges. However, each user's flow of content is

organized by algorithms built to maximize platform usage. From this comes the hypothesis that there is an algorithmic bias (also called algorithmic segregation) since these contents are selected based on users' precedent actions on the platform or the web, reinforcing the human tendency to interact with content confirming their beliefs (confirmation bias).

To introduce in the study of opinion dynamics the idea of a recommender system generating an algorithmic bias, we started from a recent extension of the well-known Deffuant–Weisbuch model (DW-model henceforth), proposed in Sîrbu et al. (2019) (Algorithmic Bias model or AB model, henceforth).

Definition 1   (Deffuant–Weisbuch model (DW-model)) Let us assume a population of N agents, where each agents i has a continuous opinions $x_i \in [0, 1]$. At every discrete time step, a pair (i, j) of agents is randomly selected, and, if their opinion distance is lower than a threshold ϵ, $|x_i - x_j| \leq \epsilon$, then both of them change their opinion according to the following rule:

$$
\begin{aligned}
x_i(t+1) &= x_i + \mu(x_j - x_i) \\
x_j(t+1) &= x_j + \mu(x_i - x_j).
\end{aligned}
\tag{1}
$$

In the DW-model, the parameter $\epsilon \in [0, 1]$ models the population's confidence bound, assumed constant among all the agents. A low ϵ creates a close-minded population, where the individuals can only be influenced by those whose opinions are similar to theirs; a high ϵ, instead, creates an open-minded population since two agents can influence each other even if their initial opinions are very distant. The parameter $\mu \in (0, 0.5]$ is a convergence parameter, modeling the strength of the influence the two individuals have on each other or, in other words, how much they change their opinion after the interaction. Even if there is no reason to assume that ϵ should be constant across the population or at least symmetrical in the binary encounters, this parameter is often considered equal for every agent (apart from a few exceptions as in Lorenz (2010)). The numerical simulations of this model show that the qualitative dynamic is mainly dependent on ϵ: as ϵ grows, the number of final opinion clusters decreases. As for μ and N, these parameters influence only the time to convergence and the final opinion distribution width.

The AB model introduces another parameter to model the algorithmic bias: $\gamma \geq 0$. This parameter represents the filtering power of a generic recommendation algorithm: if it is close to 0, the agent has the same probability of interacting with all its peers. As γ grows, so does the probability of interacting with agents holding similar opinions while interacting with those who hold distant opinions decreases. Therefore, this extended model modifies the rule to choose the interacting pair (i, j) to simulate a filtering algorithm's presence. An agent i is randomly picked from the population, while j is chosen from i's peers according to the following rule:

$$
p_i(j) = \frac{d_{ij}^{-\gamma}}{\sum_{k \neq i} d_{ik}^{-\gamma}}
\tag{2}
$$

where $d_{ij} = |x_i - x_j|$ is the opinion distance between agents i and j, so that for $\gamma = 0.0$ the model goes back to the DW-model, i.e. the interacting peer j is chosen at random from $i's$ neighbors or—in other words—every neighbor is assigned the same probability to be chosen. When two agents interact, their opinions change if and only if the distance between their opinions is less than the parameter ϵ, i.e., $|x_i - x_j| \leq \epsilon$, according to Eq. 1.

In Sîrbu et al. (2019) the discussed AB model results focus only on the mean-field scenario, i.e., the authors assume a complete graph as the underlying social structure. While considering real-world social interactions, however, we can assume that individuals will likely interact only with whom they are acquainted. Therefore, building from the analysis proposed in Pansanella et al. (2022), we evaluate the effects of two different topological models on the unfolding of the identified opinion formation process: Erdős–Rényi (ER) (Erdás and Rényi 1959) and scale-free Barabási–Albert (Barabási and Albert 1999) networks.

Algorithmic bias: from fixed topologies to adaptive networks

In Sîrbu et al. (2019) and Pansanella et al. (2022), it emerged that the dynamics and final state are mainly determined by ϵ and γ, with the confidence threshold enhancing consensus and the bias enhancing fragmentation. Comparing simulations performed on complete, ER, and scale-free networks, it emerged that the role of the underlying topology is negligible concerning the effects of the model parameters (thus, confirming what was previously observed in Weisbuch (2004); Fortunato (2004); Stauffer and Meyer-Ortmanns (2004) for the scale-free scenario). However, a higher sparsity implies that fragmentation emerges for lower values of the algorithmic bias. Despite this being a crucial step towards reality, assuming that social networks are static during the whole period is unrealistic. Interactions and relationships evolve, and this evolution influences and is influenced by the dynamical process of opinion exchanges and the presence of recommender systems and filtering algorithms for the construction of the social network, reinforcing the tendency toward homophilic choices.

In the present work, we extended the baseline model (Sîrbu et al. 2019) introducing the possibility of arc rewiring, creating the Adaptive Algorithmic Bias model (AABM henceforth), where peer-to-peer interactions are affected by algorithmic biases, and the networks evolve influenced by such interactions, bringing people to connect to peers with opinions within their confidence bound. To incorporate such behavior, we added a new parameter to the Algorithmic Bias model, namely $p_r \in [0, 1]$, indicating the probability that the agent in a situation of cognitive dissonance decides to rewire their link instead of just ignoring their peer opinion. To maintain the model as simple as possible, this parameter is assumed to be constant across the population and does not depend on the opinion distance. Thus, in the model, every time an agent i interacts with a neighbor j and their opinion distance is beyond the confidence threshold, i.e., $|x_i - x_j| \geq \epsilon$:

- with probability p_r the agent rewires the arc looking for a like-minded individual
- with probability $1 - p_r$ the DW-model is applied, i.e. both opinions and network structure remain unchanged.
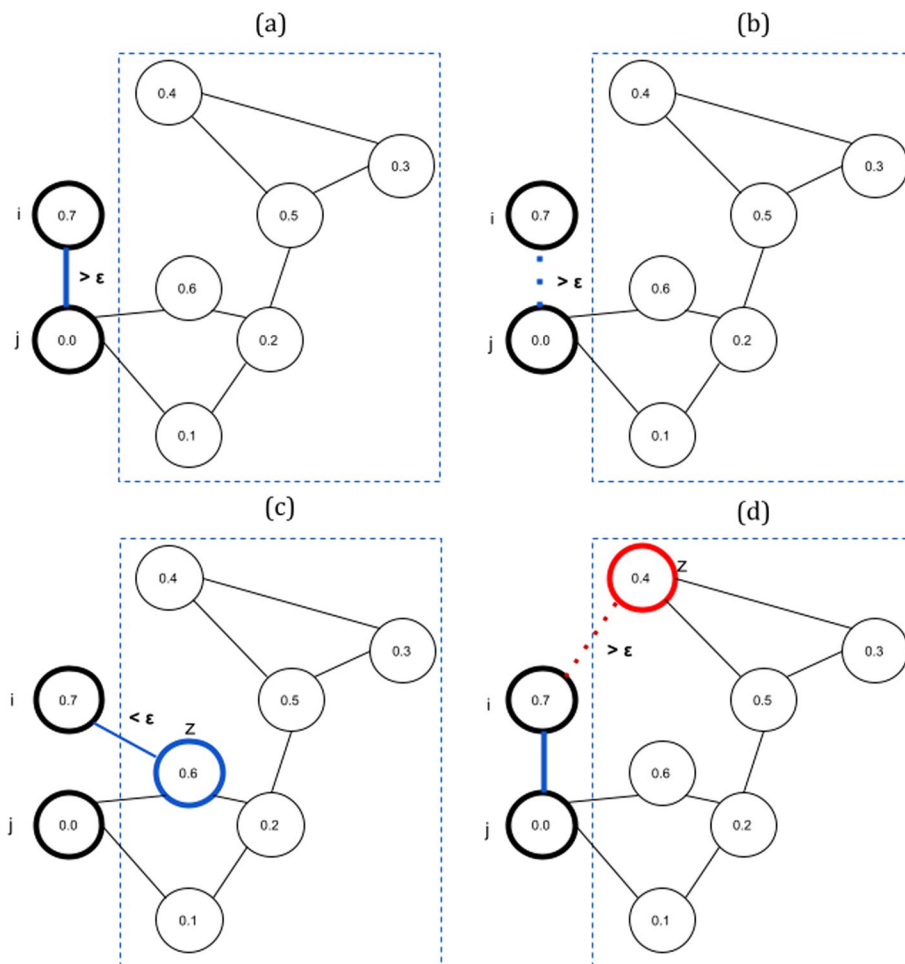
Fig. 1 A schematic illustration of the rewiring step under bounded confidence. In this example the confidence bound is $\epsilon = 0.2$. In **a**, we can see that the interacting pair (i, j) has an opinion distance further than the confidence bound. For this reason **b** node i tries to break the arc (i, j) and form a new arc (i, z) (with probability p_r, with probability $1 - p_r$ nothing happens). Node z is chosen randomly between the remaining nodes in the network. In the case that $|x_i - x_z| < \epsilon$ (**c**) the arc (i, j) is broken and the arc (i, z) is formed. Otherwise, if $|x_i - x_z| \geq \epsilon$ (**d**), the rewiring fails, and the network structure remains the same

To rewire the arc, a node z is randomly selected from the set of non-neighboring nodes, and if $|x_i - x_z| < \epsilon$, the agent z links to the agent i, otherwise the structure of the graph remains unchanged (see Fig. 1 for an example of this process and algorithm 1 for the process pseudocode).

Without considering the algorithmic bias in the choice of the interacting peer, our work is similar to Kozma and Barrat (2008); Kan et al. (2021). In Kozma and Barrat (2008) the process of rewiring works in the same fashion as in the present work, however, every time the rewiring option is chosen over the standard DW-model update rule, the old link (i, j) is broken and a new link (i, z) is formed towards a random non-neighboring agent, even if this agent's opinion is beyond $i's$ confidence threshold. Even in Kan et al. (2021) the rewiring process has a different formulation diverging from the proposed one due to the following specificities: (a) at iteration, a set M of discordant edges is rewired and then a set K of edges undergoes the process of opinion update (i.e.,

if $M < K$ opinion change faster than node rewire, like in the present work); (b) during the rewiring stage the node selection does not happen entirely at random, rather it is "biased" towards similar individuals (still allowing the connection with peers with opinions beyond the confidence threshold); finally (c) the confidence bound and the tolerance threshold for the rewiring are modeled as two independent parameters.

Conversely, from such contributions, our implementation assumes a "zero-knowledge" scenario where agents are not aware of the statuses of their peers beforehand: once rejected the algorithmically biased interaction suggestion, if an agent decides to break the tie and search for a new peer to connect with it will not rely (for that task) on other algorithm suggestions. We adopt such modeling since in social contexts (e.g., in online social networks), the status of unknown peers is hardly known by a user—at least before a first attempt at interacting with them. Moreover, not delegating the identification of potential peers to connect with to the "algorithm," we allow users to react to a first non-successful system recommendation independently (i.e., during the same iteration, the user prefers not to trust the algorithm. Instead, he/she makes a blind connection choice). Therefore, rewiring a link toward a like-minded individual is not always feasible given the limits of users' local views.

Algorithm 1 Rewiring

Given the pair (i, j) where $d_{ij} \geq \epsilon$
Randomly select a vertex z from the remaining vertices of the graph
if $|x_i - x_z| < \epsilon$ **then**
    The arc (i, j) is removed from the graph
5:    The arc (i, z) is added to the graph
else
    The structure of the graph remains unchanged
end if

Beyond pairwise interactions: modeling peer pressure

Classical networks, with vertex and arcs, only capture dyadic relationships, and every collective dynamic is analyzed as a decomposition of pairwise dynamics. However, there are many systems where it is crucial to capture group dynamics and where considering only binary interactions can limit the explanatory power of models. As introduced in "Introduction" section, different structures can be employed to model higher-order interactions. However, in the specific context of this paper, we chose to employ simplicial complexes since the idea is that a triangle of connected agents may experience peer pressure because it constitutes a group of friends, a strong friendship relationships, where in addition to the binary friendships there is a higher-order relationship among these agents. Simplexes are in fact the simplest mathematical object allowing to consider higher-order relationships. A *k-simplex* σ is the convex hull of a set of $k + 1$ nodes $\sigma = [p_0, p_1, ..., p_k]$. The 0-simplex is a single node; the 1-simplex is an arc, the 2-simplex is a triangle, etc. A simplicial complex requires that each face of a simplex is again in the simplicial complex and that the nonempty intersection of two simplices is a face for each. Since a triadic friendship, denoted by a triangle on the social network, does include the binary friendships between each of the three individuals, we propose and analyze an extension of the Algorithmic Bias model to

include second-order interactions: the Algorithmic Bias model on Simplicial Complexes (inspired by Horstmeyer and Kuehn (2020) and adapted to the context of bounded confidence models with continuous opinions). This allows us to incorporate peer pressure in an environment where confirmation and algorithmic biases are still present.

In order to implement peer pressure, i.e., a mechanism for which the majority opinion pressures the individual "minority" one to conform, we first need to define what a majority is in the context of a continuous opinion dynamics model. We choose to consider two nodes "agreeing" if their opinion distance is below the confidence threshold, i.e. $|x_i - x_j| < \epsilon$, similarly to Kozma and Barrat (2008).

The model selects a pair (i, j) according to Eq. 2 at every discrete time step and computes the set of triangles T including (i, j). If the set is nonempty, the model selects a third node z from T according to Eq. 2. Otherwise, the baseline rule is applied, i.e., there is a pairwise interaction between i and j according to the AAB-model rules.

Once the interacting triplet is chosen, if two agents form a majority, two scenarios may arise:

- the third agent already "agrees" with the majority, i.e., its opinion distance from the average opinion of the majority is below the confidence threshold
- the third agent is in a situation of cognitive dissonance with the majority, i.e., its opinion distance from the average opinion of the majority is beyond the confidence threshold

In the former scenario, the attractive behavior of the pairwise model is adapted to the triadic case: the agents take the average opinion of the triplet; in the latter, we implemented peer pressure by making the three agents adopt the average opinion of the majority. These rules are detailed in the algorithm 2.

Algorithm 2 Algorithmic bias on simplicial complexes

     Given two vertex $n1, n2$ linked by an edge $(n1, n2)$
     Compute the set T of the triangles including $(n1, n2)$
3: **if** $T = \emptyset$ **then**
       Baseline rule
     else
6:      Choose $t \in T$ with $t = \{(n1, n2), (n1, n3), (n2, n3)\}$ choosing $n3$ according to 2
       for any possible permutation (i, j, z) of the nodes in T **do**

9:         **if** $|x_i(t) - x_j(t)| < \epsilon$ and $|x_z(t) - avg(x_i(t), x_j(t))| < \epsilon$ **then**
          $newOpinion = avg(x_i(t), x_j(t), x_z(t))$
          $x_i(t+1), x_j(t+1), x_z(t+1) = newOpinion$
12:        return
         else if $|x_i(t) - x_j(t)| < \epsilon$ and $|x_z(t) - avg(x_i(t), x_j(t))| \geq \epsilon$ **then**
          $newOpinion = avg(x_i(t), x_j(t))$
15:        $x_z(t+1), x_i(t+1), x_j(t+1) = newOpinion$
         return
         else
18:        continue
        end if
       end for
21:     If none of the three pairs forms a majority nothing changes
     end if

Our goal here is to understand the effects of higher-order interactions in a biased environment on the degree of fragmentation reached by the population in the final

state. To such an extent, we tested this extended model on the same two graph models: ER (Erdás and Rényi 1959) and a scale-free (Barabási and Albert 1999) network.

We also added the possibility of arc rewiring in this model: to do so, a rewiring takes place with probability p_r between the disagreeing pair (i, j) with $|x_i - x_j| \geq \epsilon$ when T is an empty set or a "majority" cannot be found in T.

Experimental settings

Like in Sîrbu et al. (2019), to avoid undefined operations in Eq. 2, when $d_{ik} = 0$ we use a lower bound $d_\epsilon = 10^{-4}$. The simulations are designed to stop when the population reaches an equilibrium, i.e., the cluster configuration will not change anymore, even if the agents keep exchanging opinions. We also set an overall maximum number of iterations at 10^5 as in Pansanella et al. (2022). We compute the average results over 30 independent executions for each configuration to account for the model's stochastic nature. The initial opinion distribution is always drawn from a random uniform probability distribution in [0;1]. To better understand the differences in the final state concerning the different topologies considered, we study the model on all networks for different combinations of the parameters. We are interested in understanding the effects of a co-evolving topology affected by homophily on the dynamics of public opinion in a population and the consequences of peer pressure when moving from pairwise to higher-order interactions.

Moreover, we are also interested in whether, parameters being equal, different initial network topologies influence the final cluster configuration in such extended models. We tested our model, seeding the co-evolution with two different network topologies: an Erdős–Rényi (random) and a Barabási–Albert (scale-free). We set the number of nodes $N = 250$ in both networks. For the ER network, we fix the p parameter (probability to form a link) to 0.1 (thus imposing a *supercritical* regime, as expected from a real-world network); we obtain a network composed of a single giant component with an average degree of 24.94. In the BA network, we set the $m = 5$ (i.e., the parameter regulating the number of edges to attach from a new node to existing nodes), thus obtaining a network instance with an average degree equal to 9.8.[1]

In our simulations, we evaluated the different models on the different possible combinations of the parameters over the following values:

- ϵ takes a value from 0.2 to 0.4 with a step of 0.1. We chose these values because these are the values for which, in the AB model, we can observe a shift from polarization to fragmentation and from consensus to polarization. Higher values of ϵ lead to consensus regardless of the strength of the algorithmic bias until the bias is high enough and fragmentation explodes.
- γ takes value from 0 to 1.6 with a step of 0.4; for $\gamma = 0$ the model becomes the DW-model. We would see only fragmented final states for higher values of γ.
- $\mu = 0.5$, so whenever two agents interact, they update their opinions to the pair's average opinion if their opinions are close enough

---

[1] Note that the empirical average degree slightly deviates from the expected asymptotic value ($\langle k \rangle = 2m = 10$) due to statistical fluctuations introduced by the random seed used by the generative process.

- p_r (for the Adaptive version of the models) takes a value from 0.0 to 0.5 with a step of 0.1; for $p_r = 0.0$ the model becomes the AB model in the case of the Adaptive Algorithmic Bias model.

The models implementation used to carry out our experiments is the one provided by the NDlib (Rossetti et al. 2018) Python library.[2]

To analyze the simulation results, we start by considering the number of final opinion clusters in the population to understand the degree of fragmentation produced by the different combinations of the parameters. This value indicates how many peaks there are in the final distribution of opinions and provides a first approximation of whether a consensus can be obtained or not. To compute the effective number of clusters, accounting for the presence of major and minor ones, we use the cluster participation ratio, as in Sîrbu et al. (2019):

$$C = \frac{(\sum_i c_i)^2}{\sum_i c_i^2} \tag{3}$$

where c_i is the dimension of the ith cluster, i.e., the fraction of the population we can find in that cluster. In general, for m clusters, the maximum value of the participation ratio is m and is achieved when all clusters have the same size, while the minimum can be close to 1, if one cluster contains most of the population and a small fraction is distributed among the other $m - 1$. To study the degree of polarization/fragmentation, we computed the average pairwise distance between the agents' opinions. Given an agent i with opinion x_i and an agent j with opinion x_j at the end of the diffusion process, the pairwise distance between the two agents is $d_{ij} = |x_i - x_j|$. The average pairwise distance in the final state can be computed as $\frac{\sum_{i=0}^{N} \frac{\sum_{j=0}^{N} d_{ij}}{N}}{N}$. While the asymptotic number of opinion clusters and the degree of polarization are essential metrics to describe the results of the dynamics qualitatively, the time to obtain such a final state is equally so. In a realistic setting, available time is finite, so if consensus forms only after a very long time, it may never actually emerge in the population. Thus, we measure the time needed for convergence (to either one or more opinion clusters) in our extended model, recalling that every iteration is made of N interactions, whether pairwise or higher-order (triadic) (Fig. 2).

Results and discussion

Adaptive algorithmic bias model: close-mindedness leads to segregation in co-evolving networks

Our simulations suggest that allowing users to break friendships that cause disagreement in a biased online environment has little effect on the levels of polarization/fragmentation when the evolution of the network is remarkably slower than the process of convergence towards a steady-state into one or more opinion clusters. However, when two or more opinion clusters form, allowing the rewiring process to continue eventually breaks the network into multiple connected components. To understand the effects

---
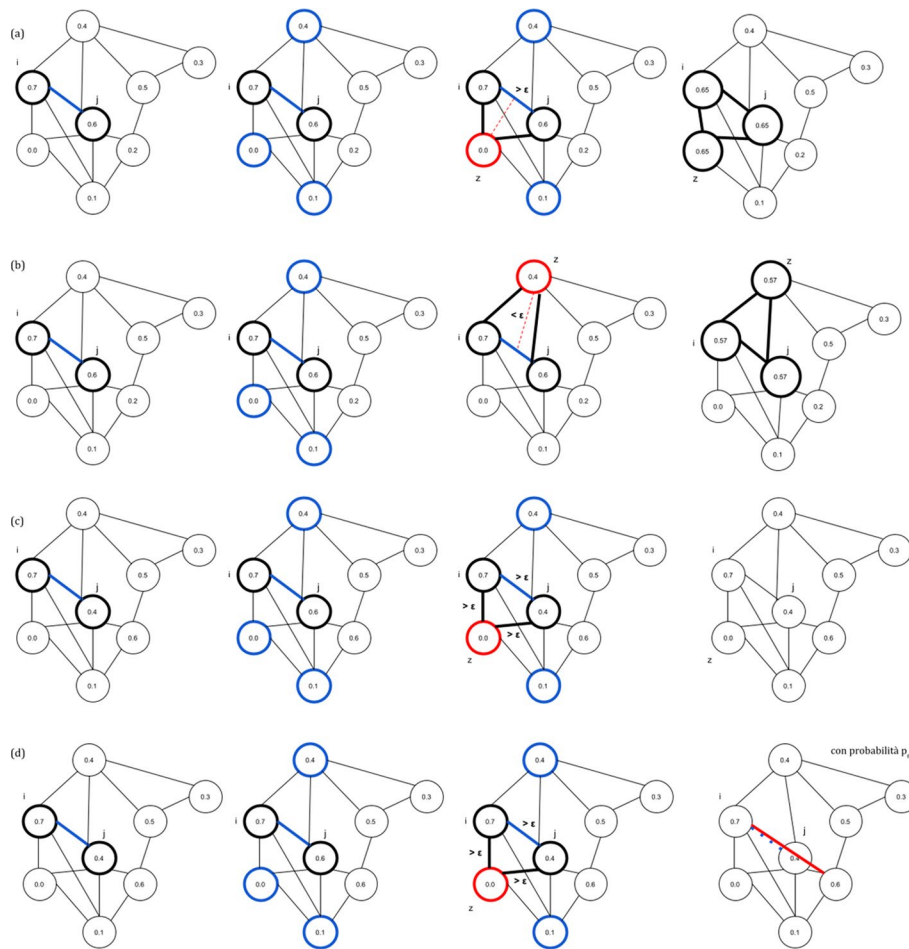
[2] NDlib: http://ndlib.rtfd.io

Fig. 2 Example of the AABSC model. Examples of different cases in the Adaptive Algorithmic Bias Model on Simplicial Complexes. In **a**, a triangle (*i*, *j*, *z*) is chosen, and the minority node adopts the mean opinion of the majority. In **b**, there is no minority, so the three agents adopt their average opinion. In **c**, there is no majority: nothing happens. In **d**, there is no majority, and agent *i* rewires the discording arc with *j* towards a more like-minded agent. The process in **d** is the same described in Fig. 1

of the interplay of cognitive and algorithmic biases and the probability of link rewiring, we start by looking at the average number of final clusters. Figure 3 shows the average number of clusters as a function of p_r and γ for $\epsilon \in \{0.2, 0.3\}$. Results for $\epsilon = 0.4$ and standard deviation analysis can be found in Additional file 1. We can observe from the first row of each heatmap that, without rewiring ($p_r = 0.0$), the behavior of the Algorithmic Bias model discussed in Pansanella et al. (2022) is recovered: fragmentation is enhanced by the bias, while higher values of ϵ counter the effects of the bias and drive the population towards a consensus around the mean opinion of the spectrum. We already saw in Pansanella et al. (2022) that concerning the mean-field case, when the topology is sparser even for $\epsilon \geq 0.4$ for a sufficiently large bias, the final states result in a high number of clusters or even not to be clustered at all, i.e., in the final state opinions are still uniformly distributed across the population since the bias is so strong that even like-minded people (whose opinion distance is below the confidence bound) can never converge to each other because they will unlikely interact.
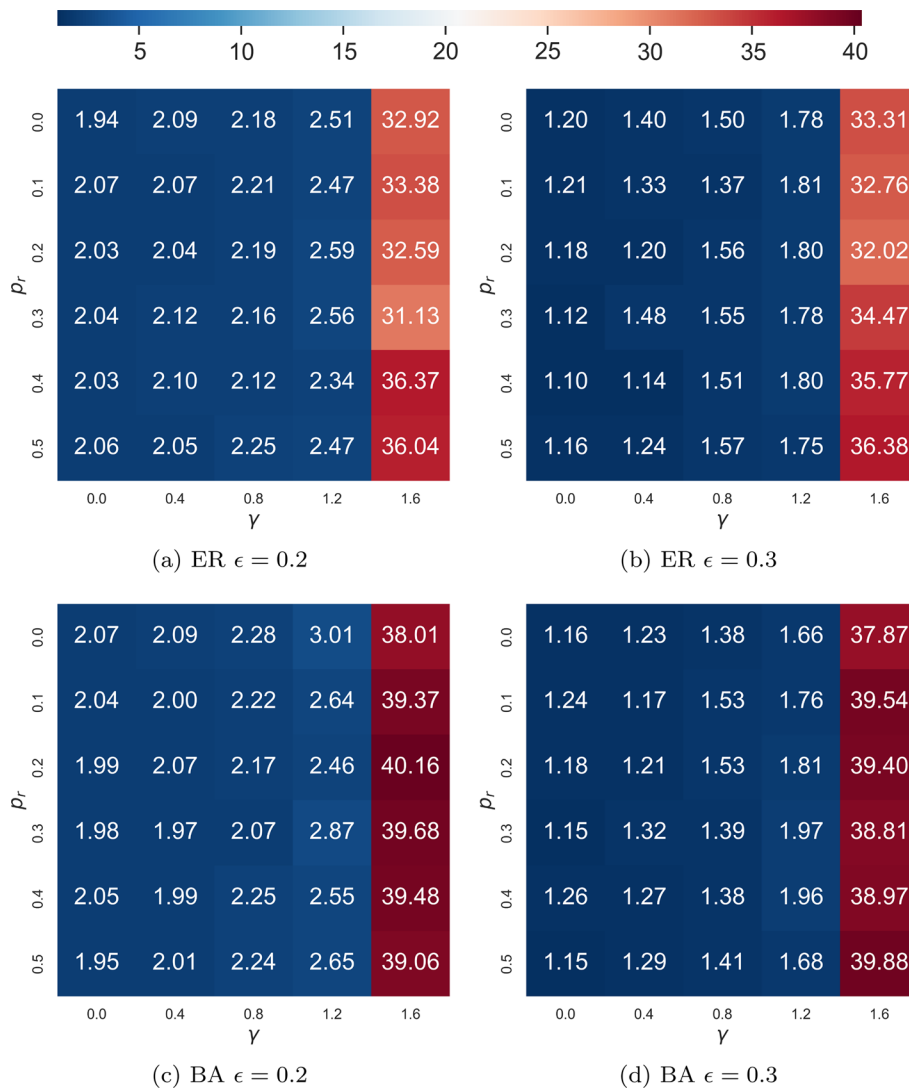
	5	10	15	20	25	30	35	40

(a) ER $\epsilon = 0.2$

p_r	0.0	0.4	0.8	1.2	1.6
0.0	1.94	2.09	2.18	2.51	32.92
0.1	2.07	2.07	2.21	2.47	33.38
0.2	2.03	2.04	2.19	2.59	32.59
0.3	2.04	2.12	2.16	2.56	31.13
0.4	2.03	2.10	2.12	2.34	36.37
0.5	2.06	2.05	2.25	2.47	36.04

γ

(b) ER $\epsilon = 0.3$

p_r	0.0	0.4	0.8	1.2	1.6
0.0	1.20	1.40	1.50	1.78	33.31
0.1	1.21	1.33	1.37	1.81	32.76
0.2	1.18	1.20	1.56	1.80	32.02
0.3	1.12	1.48	1.55	1.78	34.47
0.4	1.10	1.14	1.51	1.80	35.77
0.5	1.16	1.24	1.57	1.75	36.38

γ

(c) BA $\epsilon = 0.2$

p_r	0.0	0.4	0.8	1.2	1.6
0.0	2.07	2.09	2.28	3.01	38.01
0.1	2.04	2.00	2.22	2.64	39.37
0.2	1.99	2.07	2.17	2.46	40.16
0.3	1.98	1.97	2.07	2.87	39.68
0.4	2.05	1.99	2.25	2.55	39.48
0.5	1.95	2.01	2.24	2.65	39.06

γ

(d) BA $\epsilon = 0.3$

p_r	0.0	0.4	0.8	1.2	1.6
0.0	1.16	1.23	1.38	1.66	37.87
0.1	1.24	1.17	1.53	1.76	39.54
0.2	1.18	1.21	1.53	1.81	39.40
0.3	1.15	1.32	1.39	1.97	38.81
0.4	1.26	1.27	1.38	1.96	38.97
0.5	1.15	1.29	1.41	1.68	39.88

γ

Fig. 3 Average number of clusters in the steady-state of the Adaptive Algorithmic Bias model. The average number of clusters in the final state of the Adaptive Algorithmic Bias model as a function of γ and p_r for **a** $\epsilon = 0.2$ and **b** $\epsilon = 0.3$ startingr from the Erdős–Rényi graph and **c**, **d** starting from the scale-free Barabási–Albert graph. These values are averaged over 30 runs

Erdős–Rényi network. In Fig. 3a, b, we can see that in the DW model, i.e., $\gamma = 0.0$ adding the possibility to rewire arcs during conflicting interactions does not change the final number of clusters on average. For $\epsilon = 0.2$ we obtain a polarised population for every value of p_r. A consensus is always reached for $\epsilon \geq 0.3$, specifically for $\epsilon = 0.3$, the main cluster coexists with smaller clusters. In comparison, for $\epsilon = 0.4$, a perfect consensus around the mean opinion is always reached (see Additional file 1).

The co-evolving topology does not impact the dynamics of the Algorithmic Bias model either: the adaptive topology does not change the fact that the system reaches consensus or polarization. Plots in Fig. 3a, b show the population moving from two to three cluster in $\epsilon = 0.2$ and from one to two clusters in $\epsilon = 0.3$ and always reaching a consensus for $\epsilon = 0.4$ (see Additional file 1), even if we consider $p_r = 0.5$.
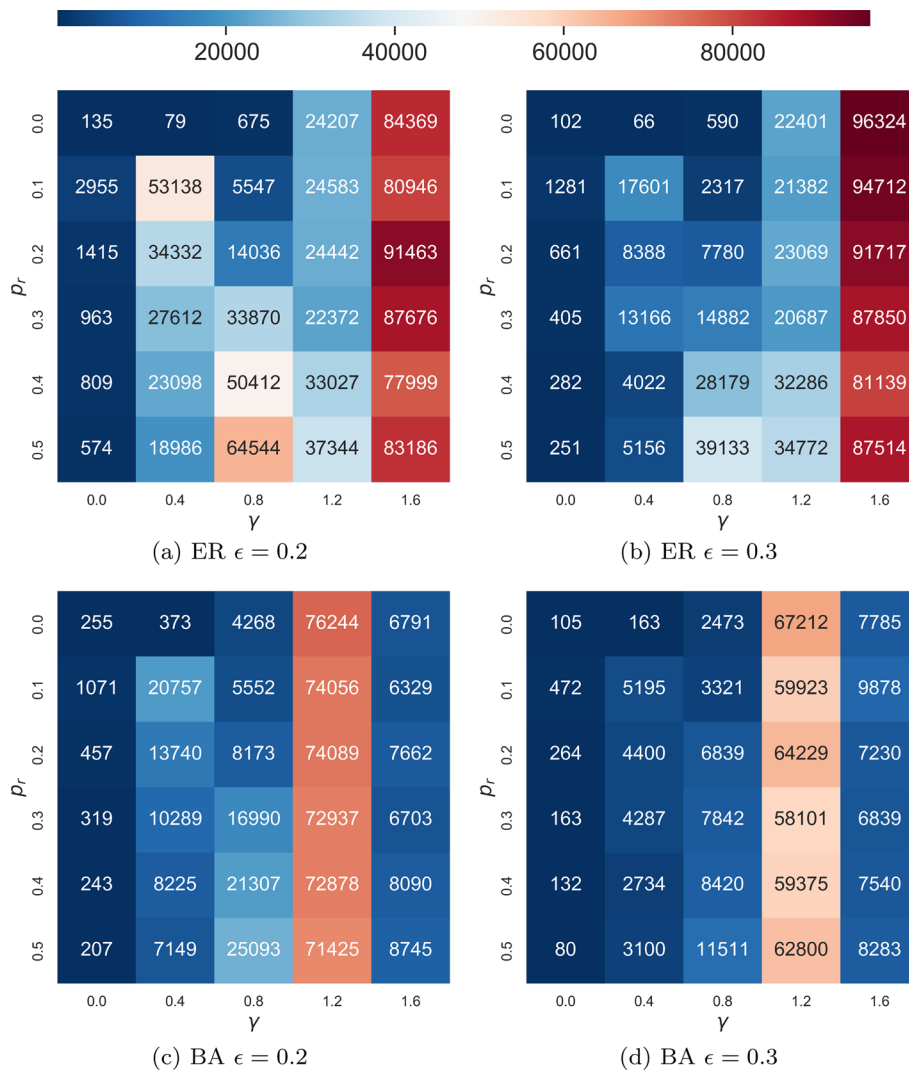
Fig. 4 Average number of iterations to convergence in the Adaptive Algorithmic Bias model. Average number of iterations to convergence in the Adaptive Algorithmic Bias model as a function of γ and p_r for **a** $\epsilon = 0.2$, **b** $\epsilon = 0.3$ and **c, d** in a scale-free Barabási–Albert graph. These values are averaged over 30 runs

For what concerns the time (i.e., the number of total interactions) the population needs to reach an equilibrium, we can see how the general behavior of the baseline model is kept, even when the network co-evolves with a biased opinion dynamic. Convergence is relatively fast when interactions are not biased, while it slows down as the bias grows until it reaches a peak, after which it speeds up again. Until the peak, a higher number of iterations positively correlates with a higher number of clusters. In contrast, even if convergence is faster, the population is spread across the opinion space after the peak. Since the bias is strong, two agents cannot get closer in the opinion distribution after the first few interactions, and the condition to reach the steady-state is met very quickly. Moreover, every node has a limited network of agents to interact with; with a strong bias, they always exchange opinions with the same agents. Not much can change once they adopt their average opinion, even long-term. We can see from Fig. 4a, b that this measure is

less dependent on the population's open-mindedness as the bias level mainly influences it. However, we can see from the average values that an increase in ϵ often means a small convergence speed-up, all other parameters being equal. Results for $\epsilon = 0.4$ and standard deviation analysis can be found in Additional file 1.

When we introduce the possibility of rewiring, convergence is generally slower. Deleting edges beyond one's confidence bound denies agents the possibility of participating in a possible path towards convergence. This does not mean that the population cannot converge; instead, a higher number of total interactions is needed. Without bias, it is worth noticing that a small probability of arc rewiring $p_r = 0.1$ in a close-minded population ($\epsilon = 0.2$) has the slowest time to convergence. Arc rewiring towards like-minded individuals and selection bias combined slow down convergence, especially in close-minded populations (i.e., $\epsilon = 0.2$): we can see that even for a relatively small bias and a relatively small probability of arc rewiring, the steady-state needs tens of thousands of iterations, while without arc rewiring less than one hundred would be enough.

Barabási–Albert network. In Fig. 3c, d, we can see that the same results that were drawn for the ER network also hold for scale-free networks, though, on the latter, fragmentation is higher on average and the same level of fragmentation arises for lower values of the bias.

Also, for the time at convergence, similar conclusions can be drawn. However, in the scale-free network, the peak is always reached for $\gamma = 1.2$, regardless of the value of p_r. As p_r grows, the convergence slows down so much that the system can no longer reach a steady state, and for higher values of the bias, convergence is much faster, while in the ER network, it is overall slower.

To sum up, we can say that the process of co-evolution of the network along with the diffusion of opinions in the population does not affect the final opinion distribution in terms of the number of opinion clusters. This is because when there is no bias—or the bias is low—despite a lot of conflicting interactions happen, the process of convergence is too fast with respect to the process of link rewiring to separate the network into many different opinion clusters, enhancing fragmentation; when the bias is high-instead—despite this already has a fragmenting effect, it also reduces the number of conflicting interactions and therefore slows down the process of network evolution, even more, leaving the network structure practically unchanged when the population reaches its steady state.

Despite not changing the final number of opinion clusters, it impacts the network's topology, as we can observe from the examples in Figs. 5 and 6. In this case, we performed experiments with the same initial configuration of the Erdős–Rényi network (i.e., 250 nodes, p = 0.1, and uniformly distributed initial opinions). We stopped the simulations when no opinion change (nor arc rewiring) happened for 1000 consecutive iterations. In these case we set $\epsilon = 0.2$ and we compared results for $\gamma \in \{0.0, 0.5\}$ and $p_r \in \{0.0, 0.5\}$. As we can observe, polarization occurs when the network is static, meaning that there are two opinion clusters in the population while the network structure remains unchanged. As we can observe from Fig. 5e, f, the two opinion clusters tend to separate into two different components or—at least—into two different communities on the network, with fewer and fewer inter-communities links, when rewiring is allowed. After 100 iterations (Fig. 5e) there is still one connected
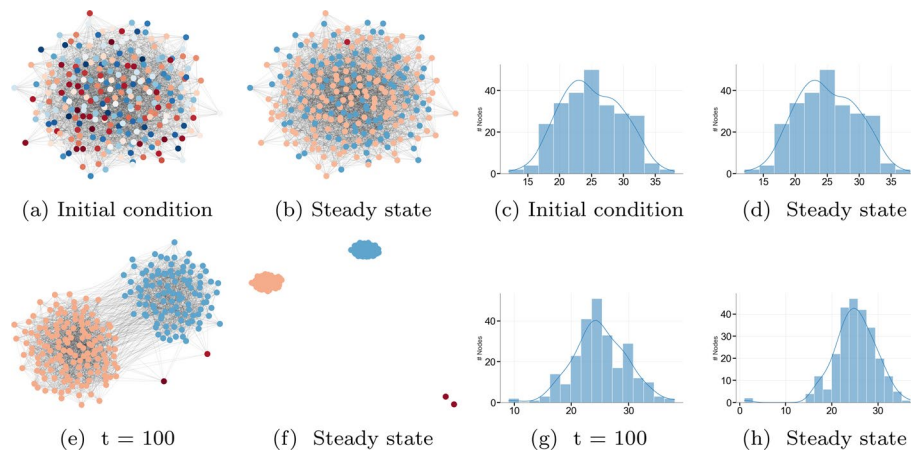
Fig. 5 Example of the effects of the adaptive topology on the Adaptive Algorithmic Bias Model on the Erdős–Rényi graph with $\gamma = 0.0$. An example of the effects of the co-evolution of network structure and opinions in the Adaptive Algorithmic Bias model on the Erdős–Rényi graph for $\epsilon = 0.2$, $p_r \in \{0.0, 0.5\}$ and $\gamma = 0.0$
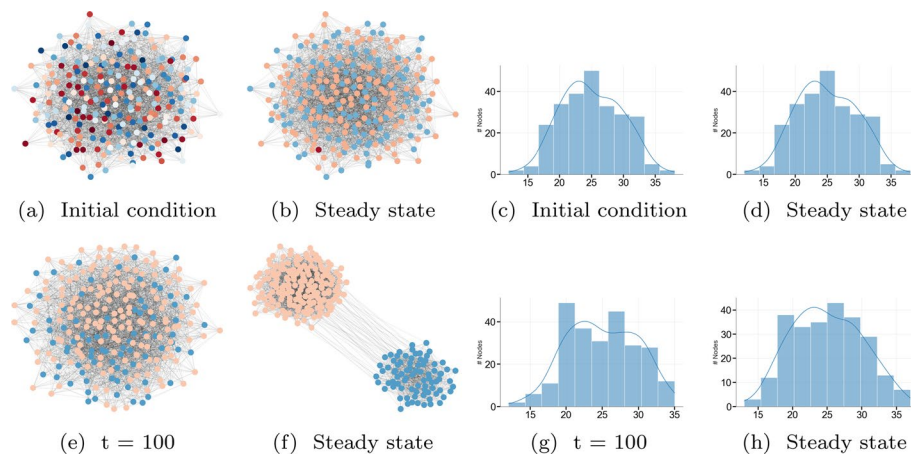

Fig. 6 Example of the effects of the adaptive topology on the Adaptive Algorithmic Bias Model on the Erdős–Rényi graph with $\gamma = 0.5$. An example of the effects of the co-evolution of network structure and opinions in the Adaptive Algorithmic Bias model on the Erdős–Rényi graph for $\epsilon = 0.2$, $p_r \in \{0.0, 0.5\}$ and $\gamma = 0.5$

component, but two polarized communities started to form . It is also worth noticing that, in the steady-state (Fig. 5f), every node is connected only to agents holding identical opinions, since there are three separated components, each holding perfect consensus. The long left tail of the degree distribution in Fig. 5h is due to the two-nodes component. If we do not consider that component, the final degree distribution is substantially similar to the distribution in Fig. 5h with a slightly lower variance. Introducing an algorithmic bias in the process slows down the convergence, as we can see from the example in Fig. 6e where—starting from the same initial configuration—the network does not present node clusters holding similar opinions after 100 iterations (while this was the case in the absence of bias). This is because bias skews interactions towards more like-minded nodes, further slowing down the process of arc rewiring by
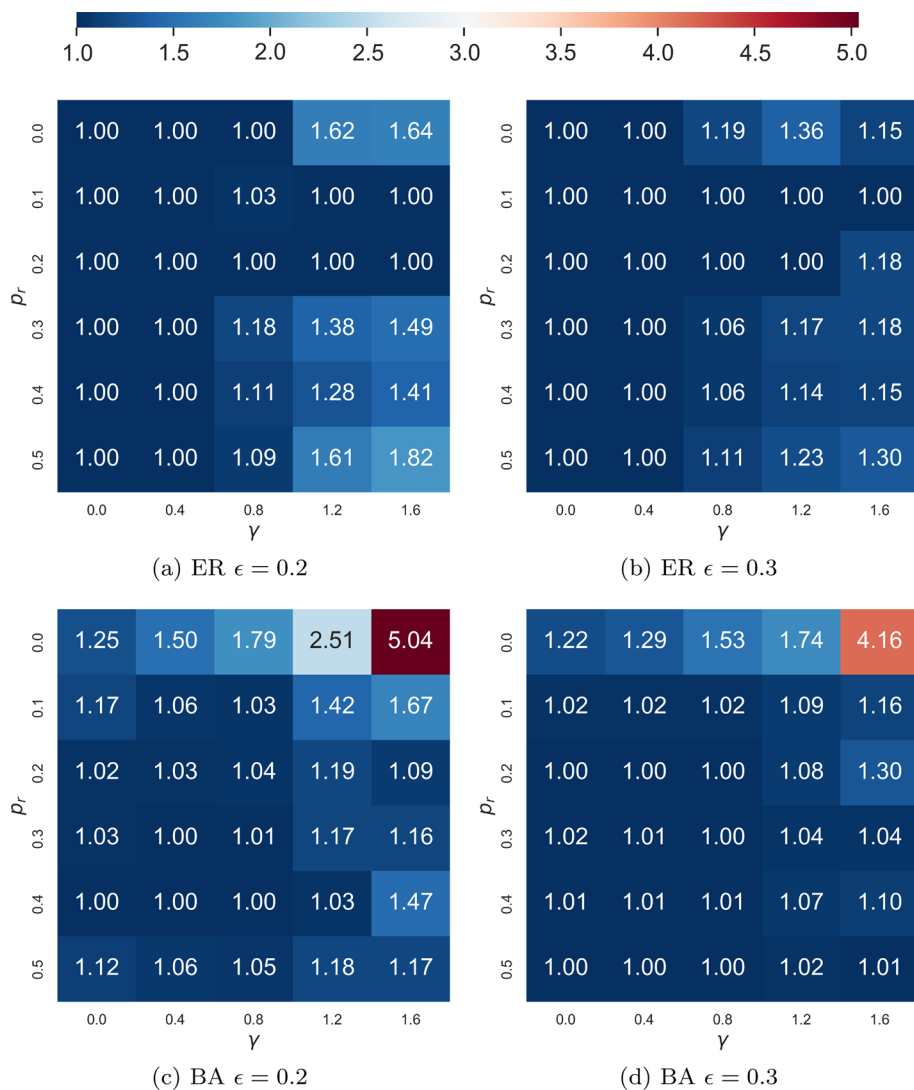
Fig. 7 Average number of clusters in the steady-state for the Adaptive Algorithmic Bias model on Simplicial Complexes. Average number of clusters in the final state for the Adaptive Algorithmic Bias model on Simplicial Complexes as a function of γ and p_r for **a** $\epsilon = 0.2$, **b** $\epsilon = 0.3$ and **c**, **d** in a scale-free Barabási–Albert graph. These values are averaged over 30 runs

reducing the amount of discording encounters. We can see from Fig. 6f that, in this case, equilibrium is reached before two components could form on the network, but there are two well-separated communities, each holding a separate opinion. While a steady state in terms of opinion clusters may be reached within a few iterations, which are not enough to separate the network into different components, if we allow the process to go on until there are no possible rewirings—since every node is connected to agreeing nodes - the network eventually splits into two or three components when opinions are clustered. Even when the maximum number of iterations set in such simulations is not enough, we can still see that links between polarized opinion clusters are fewer and fewer over time.
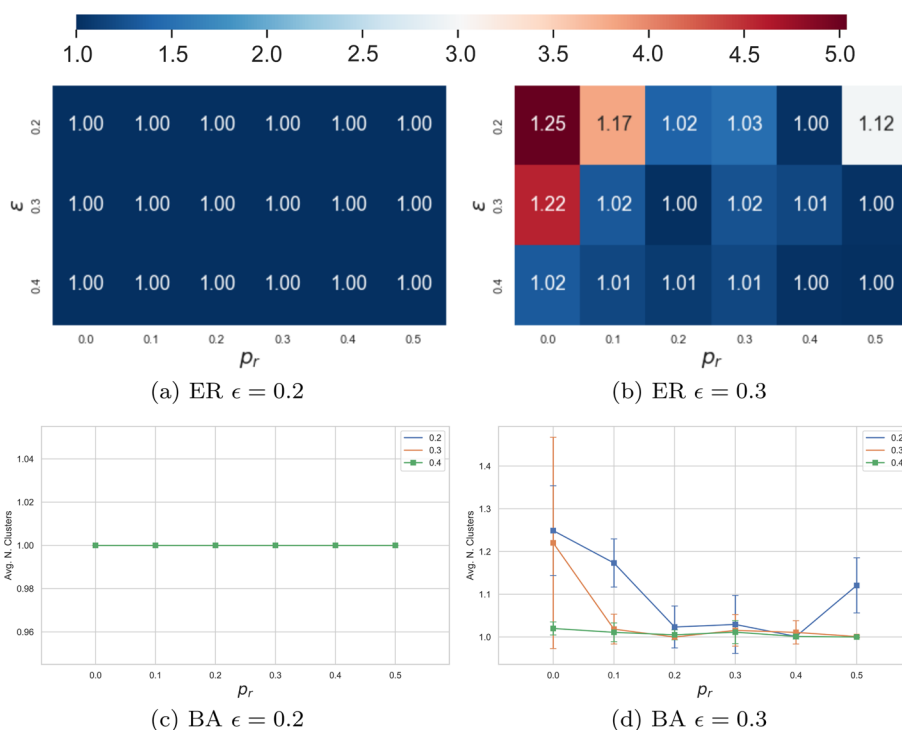
Fig. 8 Average number of clusters in the steady-state for the Adaptive Deffuant–Weisbuch bounded confidence model on Simplicial Complexess. The average number of clusters in the final state for the Adaptive Deffuant–Weisbuch bounded confidence model on Simplicial Complexes fixing $\gamma = 0.0$, as a function of ϵ and p_r for **a**, **b** a Erdős–Rényi graph and **c**, **d** a scale-free Barabási–Albert graph. These values are averaged over 30 runs

Adaptive algorithmic bias model on simplicial complexes: peer pressure enhances consensus

In the Adaptive Algorithmic Bias model on Simplicial Complexes, we introduced a simple form of higher-order interaction where three agents can influence each other—as a group—if they form a complete subgraph. Introducing higher-order interactions lets us model the phenomenon of peer pressure, where the majority edge pushes the minority node to conform to their ideology. If there is no minority opinion, we assume there would be an attractive dynamic similar to the one present in the binary case, i.e., the three nodes attract each other and adopt the mean opinion of the group.

The main result from our simulations is that peer pressure promotes consensus and reduces fragmentation with respect to the binary counterpart. Besides this general conclusion, we can observe in Fig. 7 the model's behavior is different in the two chosen networks and that γ and p_r still play a role in shaping the final state of the population.

Adaptive Deffuant–Weisbuch model on Simplicial Complexes on complex topologies. Before analyzing the effects of peer pressure and algorithmic biases on the Algorithmic Bias model, we briefly analyze the results for the Deffuant–Weisbuch model, i.e., $\gamma = 0.0$. We can observe in Fig. 8a that in the ER network, a perfect consensus is always reached, regardless of the level of bounded confidence and rewiring probability (the number of clusters is 1 in every execution of the model and the standard deviation of the final distribution is always 0). Results for $\epsilon = 0.4$ and standard deviation analysis can

be found in Additional file 1. Also, in the scale-free network (Fig. 8b), the consensus is always reached, but it is not always perfect and depends on both the confidence bound ϵ and the probability of rewiring p_r. In a static network ($p_r = 0.0$), introducing peer-pressure reduces fragmentation: in the case of $\epsilon = 0.2$, for example, the baseline model would lead to polarization, on average. Changing the update rule to account for group interactions and social pressure reduces the level of fragmentation in the final state. It leads almost the whole population to converge on a common opinion. Not surprisingly, increasing the confidence bound enhances consensus in the same way as in the baseline model. However, in this case, increasing the probability of rewiring reduces fragmentation leading the population to a perfect consensus. In particular we can see from Fig. 8b that for $\epsilon = 0.2$ perfect consensus is reached for $p_r > 0.2$, for $p_r > 0.0$ in the case of $\epsilon = 0.3$ and always in the case of $\epsilon = 0.4$.

Erdős–Rényi network. Simulating the Algorithmic Bias model on Simplicial Complexes on the chosen ER graph, for $\epsilon = 0.2$ we can see how the population always reaches a consensus for low values of the bias, while for $\gamma \geq 1.2$, peer pressure is not enough to stop the population from polarizing into two opposing clusters. However, if compared with the same results with only pairwise interactions (Fig. 3), it is clear that fragmentation is strongly reduced. For $\epsilon = 0.3$, the qualitative dynamic remains the same. However, the average number of clusters is overall reduced due to a higher open-mindedness of the population, and the population splits into two clusters only in a few cases. In contrast, in most simulations, a majority cluster forms along with a few smaller ones. When the population is open-minded, i.e., $\epsilon \geq 0.4$ consensus is always reached around the mean opinion (i.e., 0.5). The only effect of a higher algorithmic bias is that a few agents cannot converge into the main cluster. Introducing the possibility of rewiring towards a more like-minded individual after a conflicting interaction enhances polarization when combined with a mild or high selection bias (i.e., $\gamma > 0.8$). The population converges into two or three clusters when the confidence threshold is low (either two polarized clusters or two polarized clusters and a moderate one). When the population is mildly open-minded ($\epsilon = 0.3$), the system converges into one or two clusters (either two polarized clusters or a moderate cluster). At the same time, it always reaches a consensus for higher values of the confidence bound (around the mean opinion).

As we did for the previous model, we also analyzed the average time to convergence. From Fig. 9 we can see that a higher bias slows down convergence like in the binary model. It is slowed down so much that the population cannot reach a steady state within the imposed time interval. While in the model by Sîrbu et al. (2019) the level of open-mindedness did not play a crucial role in the time at convergence, in this case, we can see that increasing the open-mindedness of the population also means a faster convergence towards an equilibrium.

Barabási–Albert network. In the scale-free network, the model's behavior is slightly different: a higher probability of arc rewiring seems to reinforce consensus: we can see that when $p_r = 0.0$, the number of clusters in the final opinion distribution is higher as γ grows. This small fragmentation is reduced as p_r grows. For example, in the case of a close-minded population, i.e., $\epsilon = 0.2$, we can see that, without rewiring, a consensus is possible until the algorithmic bias is not very strong. However, it is not a perfect consensus (like in the ER network), but there is a major cluster coexisting
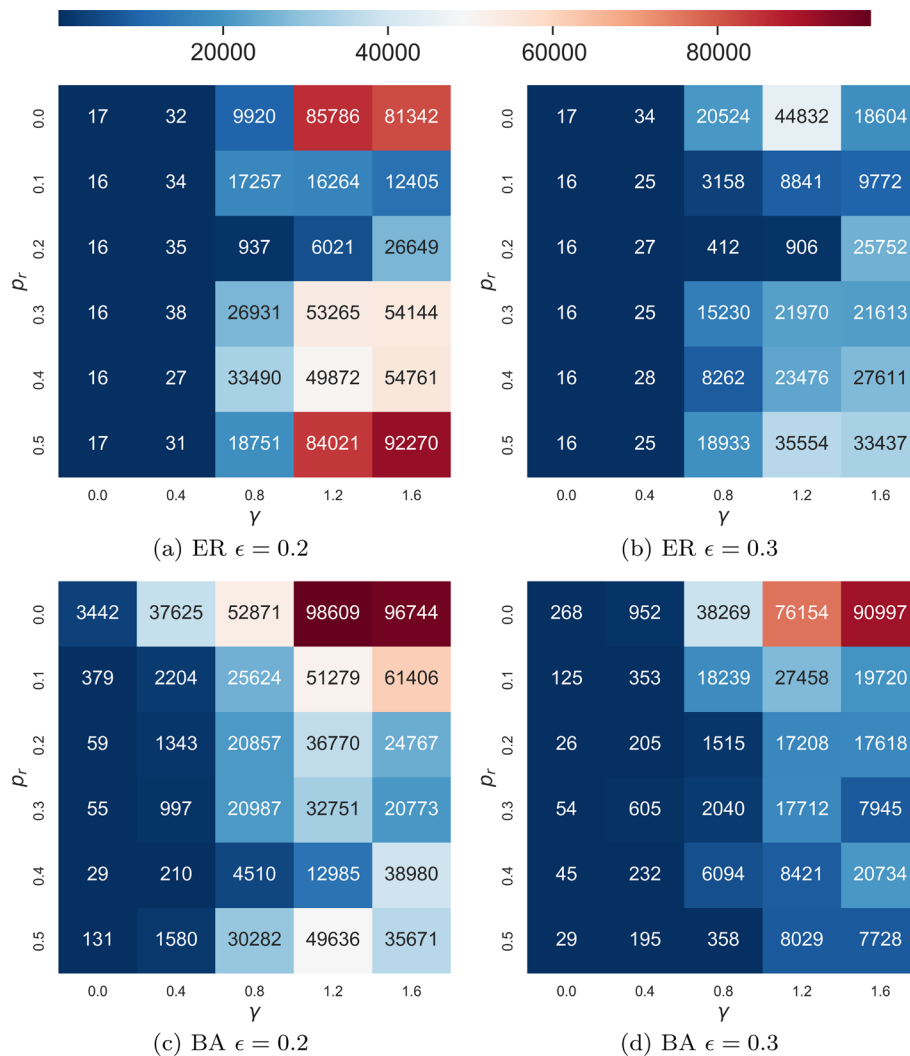
Fig. 9 Average number of iterations at convergence for the Adaptive Algorithmic Bias model on Simplicial Complexes. The average number of iterations at convergence for the Adaptive Algorithmic Bias model on Simplicial Complexes as a function of γ and p_r for **a** $\epsilon = 0.2$ and **b** $\epsilon = 0.3$ in a Erdős–Rényi graph and **c, d** in a scale-free Barabási–Albert graph. These values are averaged over 30 runs

with many agents scattered across the opinion space. Moreover, such a cluster does not necessarily form around the mean opinion but can be pretty extreme (with the final consensus below 0.2 or above 0.7). For $\gamma = 1.2$, the population becomes polarized: two homogeneous and opposed clusters form, and, in some cases, there are few "outlier" agents around the mean opinion or further at the extremes. Finally, for $\gamma = 1.6$, the population splits into multiple clusters: still, in most cases, two major polarized clusters form, alongside a variety of minor clusters below, between, and above the two. Two cohesive groups coexist with a population of individuals scattered across the opinion space so that the final distribution is not so different from the initial one: multiple opinions are still present in the population and cover the whole range [0,1]. Raising p_r to 0.1 prevents fragmentation, but for strong biases the population polarizes. For $p_r \geq 0.2$ consensus is always reached. However, as in the baseline
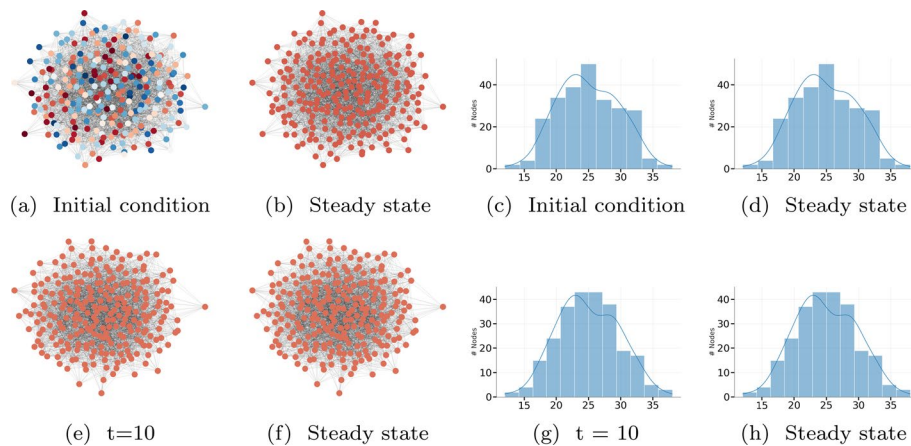
Fig. 10 Example of the effects of the adaptive topology on the Algorithmic Bias Model on Simplicial Complexes on the Erdős–Rényi graph. An example of the effects of the adaptive topology on the Algorithmic Bias Model on Simplicial Complexes on the Erdős–Rényi graph for $\epsilon = 0.2$, $p_r \in \{0.0, 0.5\}$ and $\gamma = 0.0$. The convergence towards consensus is faster and is always reached before the network can cluster around different opinions
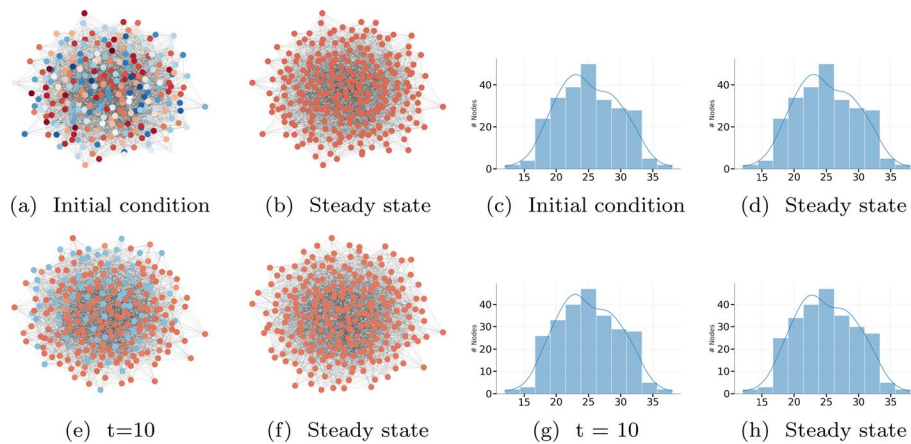

Fig. 11 Example of the effects of the adaptive topology on the Algorithmic Bias Model on Simplicial Complexes on the Erdős–Rényi graph An example of the effects of the adaptive topology on the Algorithmic Bias Model on Simplicial Complexes on the Erdős–Rényi graph for $\epsilon = 0.2$, $p_r \in \{0.0, 0.5\}$ and $\gamma = 0.5$. Bias slightly slows down the convergence process

case (without rewiring), consensus does not necessarily form around the center of the opinion space but can vary and form on strongly extreme opinions. An increase in open-mindedness also counters the fragmenting effects of the algorithmic bias. The average number of clusters reduces as ϵ grows, all other parameters being equal. In the case of a highly mildly open-minded population, i.e., $\epsilon = 0.4$ consensus can be prevented only with an extreme algorithmic bias ($\gamma = 1.6$) and without the possibility of arc rewiring. Moreover, in the scale-free topology, convergence is faster with respect to the ER network.

As we can see from Figs. 10 and 11 since, in this case, the consensus is enhanced by peer-pressure and triadic interactions that also fasten convergence, opinions

reach a steady state before the topology of the network can impact the process. We can observe that neither opinions nor nodes segregate during the process. Figure 10 shows that in the absence of an algorithmic bias, consensus can be reached within a few iterations, even with low confidence bound. Figure 11 shows how introducing an algorithmic bias does not prevent the population from reaching consensus but slows down the process even with the help of peer pressure mechanisms. Comparing $\gamma = 0.0$ and $\gamma = 0.5$ after 10 iterations, we can see how in the first case population has already reached a consensus, while in the second, two opinion clusters are still present in the network. Due to the fast convergence process towards consensus, even if rewiring is allowed, it does not significantly impact the network structure, as shown in Figs. 10e–h and 11e–h.

Conclusions

Algorithmic bias is an existing factor affecting several (online) social environments. Since interactions occurring among agents embedded in such realities are far from being easily approximated by a mean-field scenario, in our study, we aimed to understand the role played by alternative network topologies on the outcome of biased opinion dynamic simulations. From our study in Pansanella et al. (2022) emerged that the qualitative dynamic of opinions remains substantially in line with what was observed assuming a mean-field context: an increase in the confidence bound ϵ favors consensus. In contrast, introducing the algorithmic bias, γ hinders it and favors fragmentation. Conversely, both simulations' time to convergence and opinions fragmentation appears to increase as the topology becomes sparser and the hub emerges. Therefore, our analysis underlines that, alongside the algorithmic bias, the network's density heavily affects the degree of consensus reachable, assuming a population of agents with the same initial opinion distribution. The present work extends the work in Pansanella et al. (2022), proposing two extensions of the model and analyzing such extension on the same two complex networks as in Pansanella et al. (2022), leaving out the complete network. The first extension considers a straightforward mechanism of arc rewiring so that the underlying structure co-evolves with the opinion dynamics, generating the Adaptive Algorithmic Bias model. The second adds a peer-pressure mechanism, considering triangles as simplicial complexes, where a majority—if there is one—can attract a disagreeing node pushing them to conform. We found that—in general—the role of bounded confidence and algorithmic bias remains the same as in the baseline models, with the former enhancing consensus while the latter enhancing fragmentation. Going from a static to an underlying adaptive topology does not strongly affect the dynamics, leading to the same number of opinion clusters in the steady-state. However, suppose we allow the agents to continue interacting with each other. In that case, opinion clusters eventually lead to the formation of mesoscale network ones, then finally separating the network into different connected components. On the other hand, peer pressure enhances consensus, reducing the effects of low bounded confidence and high algorithmic biases. Such a model suggests how different sociological and topological factors interact with each other, thus leading populations towards polarization and echo chamber phenomena, contributing to the creation and maintenance of inequalities on social networks. These models can also be employed to study different phenomena besides opinion diffusion, such as the effects

of peer pressure on the adoption of different behaviors where social network structure and psychological factors play a role. The present work presents some of the limitations already considered in Sîrbu et al. (2019) while overcoming others. The existence of bounded confidence, for example, and the fact that it is constant across the population is an assumption that should be empirically validated along with the role of algorithmic filtering in influencing the path towards polarization/fragmentation. We went beyond the concept of static networks considering an adaptive topology; however, to further investigate the role of arc rewiring, a more thorough analysis of the model's parameters' effects on the network's topology should be made. The present implementation of a rewiring mechanism is just one way to incorporate the fact that users hardly know their neighbors' state before interacting with them; however, a mechanism considering only the set of agents with opinions within the confidence threshold would be a useful comparison to the present model. Moreover, to better understand the role of homophily in the sense of friendship formation and its relation to the online social network environment, the role of the recommender system—and therefore algorithmic bias—a biased mechanism simulating "link recommendations" could be implemented—as in Kan et al. (2021). Finally, the importance of social interactions in opinion formation is undeniable. However, external media can be essential in polarizing opinions or driving the population towards consensus. For this reason, we feel their role needs to be further investigated while embedded in an algorithmically biased environment.

Abbreviations

DW model	Deffuant–Weisbuch model
AB model	Algorithmic bias model
ER network	Erdős–Rényi network
BA network	Barabási–Albert network
AABM	Adaptive Algorithmic bias model
AABSC model	Adaptive algorithmic bias model on simplicial complexes

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1007/s41109-022-00495-7.

Additional file 1. Additional figures for the average number of clusters and the average number of iterations at convergence with standard deviation values for the two models introduced in the present work.

Declarations

Ethics approval and consent to participate
Not applicable.

References

Anderson A, Maystre L, Anderson I, Mehrotra R, Lalmas M (2020) Algorithmic effects on the diversity of consumption on spotify. In: Proceedings of the web conference 2020

Asch SE (1956) Studies of independence and conformity: I. A minority of one against a unanimous majority, vol 70. American Psychological Association, Washington, p 1

Barabási AL, Albert R (1999) Emergence of scaling in random networks. Science 286(5439):509–512

Battiston F, Cencetti G, Iacopini I, Latora V, Lucas M, Patania A et al (2020) Networks beyond pairwise interactions: structure and dynamics. Elsevier, Amsterdam

Benson B (2016) Cognitive bias cheat sheet. https://betterhumans.pub/cognitive-bias-cheat-sheet-55a472476b18

Bessi A, Zollo F, Vicario MD, Puliga M, Scala A, Caldarelli G et al (2016) Users polarization on Facebook and Youtube. PLoS ONE 11:e0159641

Castellano C, Muñoz MA, Pastor-Satorras R (2009) Nonlinear q-voter model. Phys Rev E 80(4):041129

Chen J, Geyer W, Dugan C, Muller M, Guy I (2009) Make new friends, but keep the old: recommending people on social networking sites. In: Proceedings of the SIGCHI conference on human factors in computing systems, pp 201–210

Cinelli M, Quattrociocchi W, Galeazzi A, Valensise CM, Brugnoli E, Schmidt AL et al (2020) The COVID-19 social media infodemic. Sci Rep 10:1–10

Conover M, Ratkiewicz J, Francisco M, Gonçalves B, Menczer F, Flammini A (2011) Political polarization on twitter. In: Proceedings of the international AAAI conference on web and social media, vol 5, pp 89–96

Conte R, Gilbert N, Bonelli G, Cioffi-Revilla C, Deffuant G, Kertész J et al (2012) Manifesto of computational social science. Eur Phys J Spec Top 214:325–346

Deffuant G, Neau D, Amblard F, Weisbuch G (2001) Mixing beliefs among interacting agents. Adv Complex Syst 3:11

Degroot M (1974) Reaching a consensus. J Am Stat Assoc 69:118–121

Drummond C, Fischhoff B (2017) Individuals with greater science literacy and education have more polarized beliefs on controversial science topics. Proc Natl Acad Sci 114:9587–9592

Erdás P, Rényi A (1959) On random graphs. I. Publ Math 6:290–297

Festinger L (1957) A theory of cognitive dissonance, vol 2. Stanford University, Redwood City

Fiorina MP, Abrams SJ (2008) Political polarization in the American public. Annu Rev Polit Sci 11:563–588

Fortunato S (2004) Universality of the threshold for complete consensus for the opinion dynamics of Deffuant et al. Int J Mod Phys C 15(09):1301–1307

Friedkin NE, Johnsen E (1990) Social influence and opinions. J Math Sociol 15:193–206

Friedkin N, Johnsen E (1999) Social influence networks and opinion change. Adv Group Process 01:16

Galam S (2002) Minority opinion spreading in random geometry. Eur Phys J B Condens Matter Complex Syst 25:403–406

Hague BN, Loader BD (1999) Digital democracy: an introduction. Discourse and decision making in the information age. Digital Democracy, Routledge, pp 3–22

Haun DBM, Tomasello M (2011) Conformity to peer pressure in preschool children. Child Dev 82(6):1759–67

Hickok A, Kureh YH, Brooks HZ, Feng M, Porter MA (2022) A bounded-confidence model of opinion dynamics on hypergraphs. ArXiv. 2022;abs/2102.06825

Hills TT (2019) The dark side of information proliferation. Perspect Psychol Sci 14:323–330

Hogan EA (2001) The attention economy: understanding the new currency of business. Acad Manag Perspect 15:145–147

Holley RA, Liggett TM (1975) Ergodic theorems for weakly interacting infinite systems and the voter model. Ann Probab 3(4):643–663

Holme P, Newman MEJ (2006) Nonequilibrium phase transition in the coevolution of networks and opinions. Phys Rev E 74(5):056108

Horstmeyer L, Kuehn C (2020) Adaptive voter model on simplicial complexes. Phys Rev E 101(2):022305

Iñiguez G, Kertész J, Kaski KK, Barrio RA (2009) Opinion and community formation in coevolving networks. Phys Rev E 80(6):066119

Iyengar S, Hahn KS (2009) Red media, blue media: evidence of ideological selectivity in media use. J Commun 59:19–39

Kan U, Feng M, Porter MA (2021) An adaptive bounded-confidence model of opinion dynamics on networks. ArXiv. 2021;abs/2112.05856

Knobloch-Westerwick S, Mothes C, Polavin N (2020) Confirmation bias, ingroup bias, and negativity bias in selective exposure to political information. Commun Res 47:104–124

Kozma B, Barrat A (2008) Consensus formation on coevolving networks: groups' formation and structure. J Phys A 41:224020

Lorenz J (2010) Heterogeneous bounds of confidence: meet, discuss and find consensus! Complexity 15(4):43–52

Maes M, Bischofberger L (2015) Will the personalization of online social networks foster opinion polarization? Available at SSRN 2553436

McCarty N (2019) Polarization: what everyone needs to know®. Oxford University Press, Oxford

McPherson M, Smith-Lovin L, Cook JM (2001) Birds of a feather: homophily in social networks. Ann Rev Sociol 27(1):415–444

Pansanella V, Rossetti G, Milli L (2022) From mean-field to complex topologies: network effects on the algorithmic bias model. In: Gaito S, Quattrociocchi W, Sala A (eds) Complex networks and their applications X. Springer, Berlin, pp 329–340

Pariser E (2011) The filter bubble: what the Internet is hiding from you. Penguin UK, London

Peralta AF, Kertész J, Iñiguez G (2021b) Opinion formation on social networks with algorithmic bias: dynamics and bias imbalance. IOP Publishing, Bristol

Peralta AF, Kertész J, Iñiguez G (2022) Opinion dynamics in social networks: from models to data. arXiv:2201.01322

Peralta AF, Neri M, Kertész J, Iñiguez G (2021a) Effect of algorithmic bias and network structure on coexistence, consensus, and polarization of opinions. APS, New York

Perra N, Rocha LEC (2019) Modelling opinion dynamics in the age of algorithmic personalisation. Sci Rep 9:1–11

Rossetti G, Milli L, Rinzivillo S, Sîrbu A, Pedreschi D, Giannotti F (2018) NDlib: a python library to model and analyze diffusion processes over complex networks. Int J Data Sci Anal 5(1):61–79

Sasahara K, Chen W, Peng H, Ciampaglia GL, Flammini A, Menczer F (2019) Social influence and unfollowing accelerate the emergence of echo chambers. J Comput Soc Sci 4:381–402

Sîrbu A, Pedreschi D, Giannotti F, Kertész J (2019) Algorithmic bias amplifies opinion fragmentation and polarization: a bounded confidence model. PLoS ONE 14(3):e0213246

Stauffer D, Meyer-Ortmanns H (2004) Simulation of consensus model of Deffuant et al. on a Barabasi–Albert network. Int J Mod Phys C 15(02):241–246

Sunstein CR (2007) Republic.Com 2.0. Princeton University Press, Princeton

Sznajd-Weron K, Sznajd J (2000) Opinion evolution in closed community. HSC Research Reports

The Polarization Index. https://thepolarizationindex.com/

Vicario MD, Vivaldo G, Bessi A, Zollo F, Scala A, Caldarelli G et al (2016) Echo chambers: emotional contagion and group polarization on Facebook. Sci Rep 6:1–12

Weisbuch G (2004) Bounded confidence and social networks. Eur Phys J B 38:339–343

Publisher's Note