

# Unfriending facilitates cooperation: Co-evolution of opinion and network dynamics

Mikkel Werling, (201706722)

## Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
1.1	Polarization, echo chambers and the threat to democracy . . . . .	4
<b>2</b>	<b>Why choose agent-based modelling</b>	<b>6</b>
<b>3</b>	<b>Central Mechanisms</b>	<b>6</b>
3.1	Homophily . . . . .	6
3.1.1	Explaining Homophily . . . . .	7
3.2	Social Influence . . . . .	8
<b>4</b>	<b>Opinion Dynamics</b>	<b>8</b>
<b>5</b>	<b>Network Formation</b>	<b>9</b>
5.1	Triadic Closure . . . . .	9
<b>6</b>	<b>Co-evolution via tie-deletion</b>	<b>10</b>
6.1	Tie-deletions effect on cooperation . . . . .	11
<b>7</b>	<b>Model Description</b>	<b>12</b>
7.1	Conceptual overview . . . . .	12

7.2 Assumptions . . . . .	12
7.3 Components and their properties . . . . .	12
7.4 Initialization . . . . .	13
7.5 Dynamics . . . . .	13
7.5.1 Sampling agents and interactions . . . . .	13
7.5.2 Interaction between neighbors . . . . .	14
7.5.3 Tie-dissolution . . . . .	15
7.6 Outcome metrics . . . . .	15
7.6.1 Time-dependent Metrics . . . . .	15
7.6.2 Final State Metrics . . . . .	15
<b>8 Model Evaluation</b>	<b>16</b>
8.1 Fitting the model . . . . .	16
8.2 Data . . . . .	17
8.3 Hyperparameter Optimization . . . . .	17
8.4 Results . . . . .	18
8.5 Why is co-evolution a better explanation? . . . . .	20
<b>9 Results</b>	<b>21</b>
9.1 Alpha, beta, threshold, tie-deletion and chaotic systems . . . . .	21
9.2 Correlations . . . . .	25
<b>10 Discussion</b>	<b>27</b>
10.1 Reiterating previous literature . . . . .	27
10.2 The effect of initial opinions . . . . .	27
10.3 Tie-deletion facilitates cooperation . . . . .	27

---

10.4 Distance in opinion and network space . . . . .	29
<b>11 Contrast to previous work</b>	<b>33</b>
<b>12 Short comings and future work</b>	<b>34</b>
12.1 Including Empirical Data . . . . .	34
12.2 Accurate perceptions of opinions . . . . .	34
12.3 Shrinking Variance . . . . .	35
12.4 Social influence exclusively . . . . .	35
<b>13 A broader perspective</b>	<b>36</b>
<b>14 Conclusion</b>	<b>36</b>
<b>References</b>	<b>36</b>

# 1 Introduction

## 1.1 Polarization, echo chambers and the threat to democracy

Democracies around the world are experiencing increased amounts of polarization (Boxell, Gentzkow, & Shapiro, 2020; McCoy, Rahman, & Somer, 2018; Somer & McCoy, 2018). As a result, making decisions by cooperating with other people becomes harder which can severely damage the ability of democratic systems to solve problems (Andris et al., 2015; Levin, Milner, & Perrings, 2021; McCoy et al., 2018). A striking recent example is uniquely severe rise of polarization of the political system in the United States (Dimock & Wike, 2020). During the last two decades, the amount of overlap of opinions between the two political parties have decreased substantially, which has led to gridlock, government shutdowns and failure to enact new legislation (Andris et al., 2015; Center, 2014b). A similar increase can be observed in terms of affective polarization. Affective polarization refers to how much citizens carry more negative feelings towards other political parties than their own (Boxell et al., 2020; Iyengar, Lelkes, Levendusky, Malhotra, & Westwood, 2019). Once again, the United States is reported to have uniquely high amounts of affective polarization (Boxell et al., 2020). Although the United States is unique in the extent of affective polarization, they are not the only country experiencing an increase in affective polarization. Countries such as France, Switzerland and Denmark has seen increased levels of affective polarization in the last two decades (Boxell et al., 2020). Polarization is in other words a general trend, affecting countries around the world (McCoy et al., 2018; Somer & McCoy, 2018; Wilson, Parker, & Feinberg, 2020). With that said, polarization of opinions does not seem to be an inevitable state for democracies. During the last two decades, a number of European democracies have experienced higher degrees of consensus and less affective polarization, which includes Sweden, Norway and Germany (Boxell et al., 2020).

There is reason to believe that the problem of polarization might be getting worse due to technology. Informational filters can create situations where almost all opinions one is exposed to are congruent with one's belief, popularized under the term echo chambers on social media (F. Baumann, Lorenz-Spreen, Sokolov, & Starnini, 2020; Sasahara et al., 2021; Tsai, Tao, Chuan, & Hong, 2020). Relatedly, higher involvement in one's echo chamber correlates with more negative emotions in terms of valence, which suggests that echo chambers are largely driven by outrage (Del Vicario et al., 2016). With regards to polarization, the internet can be seen as a "mixed blessing" – it can drive opinions closer together by diversifying information, but it also enables encounters of opposing opinions by chance, which can increase polarization (Lev-On & Manin, 2009). This points to a certain tension between how much exposure to different opinions affects polarization. When presented with diversifying views, people generally decrease negative affects to other political parties (Levy, 2021), but exposure to opposing views can cause opinions to become more polarized (Bail et al., 2018).

Although polarization is usually referred to as a political phenomenon, polarization is important for any social system in order to cooperate successfully. It is therefore important that we understand what mechanisms can lead a system to polarization, to consensus or anywhere in between the two extremes. This is no easy task, as opinion formation is a highly complex phenomena (F. T. W. Baumann, 2021). Nevertheless, an entire field has been dedicated

to study how opinions change. The study of the dynamics shaping opinions is aptly named opinion dynamics. The models from opinion dynamics are often analogous models, drawing inspiration from statistical physics in how they build their models (Galesic, Olsson, Dalege, van der Does, & Stein, 2021). The models considered in opinion dynamics include models of discrete and continuous opinions. The field has already a rich history of research on how opinions change using experimental data in combination with formal models (F. T. W. Baumann, 2021; Chacoma & Zanette, 2015; Flache et al., 2017; Friedkin & Johnsen, 1990; Noorazar, Vixie, Talebanpour, & Hu, 2020; Spears, 2021; Turner & Smaldino, 2018). These models normally focus on mechanisms of social influence, where interactions between agents shape the agents' opinions over time (Flache et al., 2017). Of special interest to this paper, the models typically assume a static network structure in which agents are situated (Galesic et al., 2021). This fails to take into account that networks are dynamic and co-evolves with opinions (Ferraz de Arruda et al., 2022; Galesic et al., 2021). The interdependence between the dynamics of social networks and opinion dynamics is largely understudied, despite the fact that the two processes seem to influence each other in important ways (Asikainen, Iñiguez, Ureña-Carrión, Kaski, & Kivelä, 2020; Bruch & Atwell, 2015; Galesic et al., 2021; Kossinets & Watts, 2009; Noorazar et al., 2020). The dynamic relation between the network structure and social influence has often been reported as an important avenue for future research (Flache et al., 2017; Galesic et al., 2021).

The field as a whole seems to have an issue of including empirical data in their models. This is most likely because opinions are hard to measure, but also because the field of opinion dynamics predominantly uses simulations and formal models to study their systems of interest. This is important because it often leads to models which are not evaluated in terms of how well they correspond to the real world (Flache et al., 2017; Galesic et al., 2021). If we want to get closer to the answers of how opinions are formed in social networks, we need to make better use of the empirical data available as well as evaluate how well our models can explain the observed patterns of opinionated social networks.

This paper seeks to investigating the co-evolution of the dynamics of opinions and networks by combining simple models based on insights from cognitive science, social psychology, computational biology and network science (Asikainen et al., 2020; Flache et al., 2017; Ilany & Akçay, 2016; Jackson & Rogers, 2007; Santos, Pacheco, & Lenaerts, 2006). This paper seeks to start filling the gap in the literature concerning the interconnected dynamics of networks and opinions. This is done by investigating a co-evolutionary agent-based model of opinion and network dynamics. The model is specifically designed to allow for the considering of the impact that opinions have on social networks and vice versa. In order to investigate how opinion dynamics influences social networks, this paper attempts to estimate quantitatively how well the proposed network formation models fit empirical social networks. The investigation primarily focuses on the difference in performance between co-evolutionary models and their network formation counterparts. To investigate how network dynamics influences opinion dynamics, the dynamics of the networks are changed and the resulting change in opinions are examined. This leads to the central question that this paper will try to provide answers to. Namely, this paper's central interest is to investigate what the effect of tie-deletion to dissimilar others has on facilitating cooperation.

## 2 Why choose agent-based modelling

In order to get sufficiently good answers to our questions, we need to build sufficiently good models. When it comes to models of opinion dynamics, this is no easy task. The social systems underlying opinion dynamics are complex and complicated. The change of opinions is both a cognitive matter of updating one's belief, but also a social matter of being exposed to different perspectives (Flache et al., 2017; Friedkin & Johnsen, 1990; Spears, 2021). Models that do incorporate realistic social and cognitive mechanisms are typically verbal models. Although these verbal models often provide rich and detailed accounts with realistic mechanisms, they lack the precision of formal models (Fogarty, Ammar, Holding, Powell, & Kandler, 2022; Galesic et al., 2021; P. E. Smaldino, 2020). Agent-based models are well-equipped to strike a balance between the two types of models (Flache, 2018; Galesic et al., 2021; Epstein, 1999; Mäs, Flache, & Kitts, 2014). This is done by investigating the macro-behaviors of a system, where the micro-behaviors are specified (Bruch & Atwell, 2015; Epstein, 1999; Flache, 2018). As is the case with any model, the results critically hinge on the assumptions of the model. It is therefore the role of the modeler to provide as empirically plausible mechanisms as possible for the system of study (Crooks & Heppenstall, 2012; Epstein, 1999; Page, 2010). At the same time, the value of a model is in its simplification of reality. The model should be as simple as it can be and as complicated as it needs to be in order to answer the questions of interest (P. E. Smaldino, 2020). The hope is that by simplifying the system to a sufficient extent, we can observe and understand some important features of even extremely complex systems (Fogarty et al., 2022; P. E. Smaldino, 2020). The modeler's job is to strike this balance between realism and simplicity by including only the most central mechanisms of the system in their model (P. Smaldino, 2016).

The rest of the introduction will review the empirical literature and try to find the central mechanisms that needs to be included in the agent-based model. To incorporate these mechanisms in the model of this paper, we turn to identifying key aspects of how social interactions shape opinions and how networks change over time. We start by considering key features of social interactions. Here, we will focus on homophily as well as positive and negative social influence. Thereafter, we turn to network dynamics and discuss the central role of triadic closure in how networks are shaped over time.

## 3 Central Mechanisms

### 3.1 Homophily

One of the most robust findings of the social world is summed up in the now famous phrase "birds of a feather flock together" (McPherson, Smith-Lovin, & Cook, 2001). This phrase refers to the fact that assortment in humans is non-random, and is often based on similarity (Asikainen et al., 2020; Crandall, Cosley, Huttenlocher, Kleinberg, & Suri, 2008; McPherson et al., 2001). Almost everywhere researchers look, what they find is that you are more likely to engage with other individuals that are similar to you (Taylor, Mantzaris, & Garibay, 2018). This is true for almost all

demographic variables which have been investigated, which includes gender, race, religion, and socioeconomic class (Asikainen et al., 2020; McPherson et al., 2001). This is what is referred to as observed homophily in the literature (Kossinets & Watts, 2009; McPherson et al., 2001). Observed homophily seems very intuitive. It seems reasonable to assume that you have more in common with your friends than you have with strangers. Kossinets & Watts (2009) sought to estimate how observed homophily manifested in real dynamical social networks. They recorded social interactions on a large university over the course of a year. Their findings not only suggests that this intuitive notion is true, but that it is a special case of a more general phenomenon. The findings suggests that distance in similarity is a function of distance in the social network (see Fig 1). The further you are removed from someone in the social network, the less you will have in common (Kossinets & Watts, 2009). The connection between distance in similarity and distance in the social network seems to be a robust finding of social networks (Bener, Çağlayan, Henry, & Prałat, 2016; Crandall et al., 2008). But why is this the case? The normal explanation includes two central dynamics. The first is that people tend to choose new friends that are similar to themselves. The second is that friends tends to become more similar as a result of social interactions (Crandall et al., 2008).

### 3.1.1 Explaining Homophily

Homophily is traditionally explained as a psychological preference for similar people (Asikainen et al., 2020; McPherson et al., 2001; Winter & Kataria, 2020). Interacting primarily with similar people ensures easier communication and enhances the individual's ability to predict the other person's behavior (Kossinets & Watts, 2009; Winter & Kataria, 2020). Assorting based on similarity might facilitate cooperation, because coordination between similar individuals is simpler and less costly for the individual (Winter & Kataria, 2020). This psychological explanation of the homophilic preference of individuals seems sufficient to explain the phenomenon of homophily, but structural constraints could also limit how dissimilar choices of interactions are available to a person (Peixoto, 2022). If you work as a female nurse, chances are that most of your interactions at work will be with other females. In this case, even if you do not have a strong psychological preference for interacting with people like yourself, the social interactions available to you might primarily be people like yourself. In this scenario, there will be high observed homophily by sheer chance. The existing assortment in the network makes the probability of interacting with similar individuals much higher than with non-similar individuals (Peixoto, 2022). These structural constraints are of course not exclusive to gender or workplaces. Other examples include socio-economic status and the neighborhood you live in and preference in music and what concerts you attend (McPherson et al., 2001). In other words, the psychological and the structural factor contributing to observed homophilic interactions are not mutually exclusive but might instead facilitate each other (Asikainen et al., 2020). Even a little psychological homophilic preference could turn into structural factors over the course of time (Asikainen et al., 2020; Kossinets & Watts, 2009; Taylor et al., 2018). Empirical work points to both psychological and structural homophily being important mechanisms, but suggests that structural homophily might be a more powerful mechanism for finding new connections in a social network (Bener et al., 2016; Kossinets & Watts, 2009). As mentioned, the observed general finding of assortment in networks are not only based on selecting friends based on similarity. It is also due to the fact that humans interact with their social ties, which typically increases simi-

larity (Friedkin & Johnsen, 1990; Spears, 2021). The combined mechanism of social interactions in shaping opinions is called social influence and is what we turn our attention to next.

### 3.2 Social Influence

Humans are inherently a social species. We think and form our opinions by discussing them with other people. Either via social media, debates or discussions with friends, opinions are constantly being shared between individuals. As a result of these social interactions, the opinions of people involved in the interaction might change. The effect that social encounters and social relations has on the opinions of the involved agents is referred to as social influence. Social influence has been proposed as the most important effect in shaping an individual's opinions (Chacoma & Zanette, 2015). In this paper, we will distinguish between positive and negative social influence. Positive social influence refers to interactions that result in more similarity between agents after interacting (Flache et al., 2017; Levin et al., 2021). This can be thought of as reaching a compromise, finding common ground or seeing the other person's perspective on a topic. Negative social influence refers to interactions that result in less similarity between agents after interacting (Flache et al., 2017). Examples include moral outrage, outgroup aversion or similar distancing effects. It is important to note that although positive influence seems to be a robust empirical finding, negative influence is more elusive (Flache et al., 2017; Takács, Flache, & Maes, 2014; Turner & Smaldino, 2018). With that said, there is evidence that suggests that exposure to opposing views can lead to increased polarization (Bail et al., 2018).

## 4 Opinion Dynamics

The effects of homophily and social influence has been studied extensively with different types of formal models, primarily agent-based models (Flache et al., 2017; Flache, 2018; Noorazar et al., 2020). These simplified systems are normally studied to identify which conditions can give rise to polarization or consensus in terms of the opinions of simulated agents (Flache et al., 2017). Previous models have largely only focused on how social influence affects a system over time, without considering how the social interactions change as a result of changes in opinions (Galesic et al., 2021; Holme & Newman, 2006; Jalili, 2015). Central to most classical models of opinion dynamics is that agents are situated in static, theoretical networks (Flache et al., 2017). The assumption of static networks is not inconsequential. Social networks are inherently dynamic, with new ties being formed and deleted over time. But as has already been established, empirical work suggests that these two processes are largely interdependent (Bener et al., 2016; Kossinets & Watts, 2009). New social interactions are more likely between similar agents, and tie deletion is also more likely between dissimilar agents (Kossinets & Watts, 2009). Although all models do some violence on the world by making assumptions, models should try to include the vital mechanisms of the system they are representing, while simplifying the system enough to understand it (Epstein, 1999; P. Smaldino, 2016). Considering the empirical evidence (Bener et al., 2016; Crandall et al., 2008; Kossinets & Watts, 2009), it is one of this paper's central claims that coevolution should be considered a vital mechanism. In order to support this claim, we need to expand our

understanding of the basic underlying mechanisms of social influence to also incorporate the underlying mechanisms of network formation. We therefore turn our attention to identifying empirically plausible micro-behaviors of network formation.

## 5 Network Formation

In order to identify realistic mechanisms of network behavior, it is fruitful to first survey previously considered networks in opinion dynamics. Classic examples of the static, theoretical networks used in most models include ring lattices, small-world networks and scale-free networks (Barabási & Bonabeau, 2003; Watts & Strogatz, 1998). Both small-world and scale-free networks are famous for being able to capture essential features of real social networks. Namely, small-world networks are able to capture the general tendency of large average clustering coefficient and small average path length (Watts & Strogatz, 1998). Famously, this is achieved by starting with a ring lattice and rewiring a small percentage of the ties randomly, which creates highways of information that decreases the average path length considerably (Watts & Strogatz, 1998). On the other hand, scale-free networks can capture the general tendency of long tails in degree distributions of social networks. That is, the degree distribution of social networks often follows a power law, where a few nodes have an extreme amount of edges, while most nodes only have a few (Barabási & Bonabeau, 2003). Recent work seems to call into question how universal power laws are in empirical social networks, and finds instead that most networks follow a lognormal distribution, and not a power law (Broido & Clauset, 2019). It is noteworthy that neither of the theoretical networks can capture the essential features of social networks simultaneously (Jackson & Rogers, 2004). As these features are likely to be critical to the behavior of how the agents interact in such a network, we need more empirically plausible models of social network formation.

### 5.1 Triadic Closure

An alternative to the preferential attachment of scale-free networks and the rewiring of small-world networks is to consider what mechanisms govern how real social networks change over time. Both scale-free and small world networks attempt to answer the question of how networks characteristics could be generated, not why social networks have the characteristics that they do (Jackson & Rogers, 2004). An attempt at answering the why-question is the work by Jackson & Rogers (2004), which emphasizes the role of triadic closure in network generation. Reliably, triadic closure is found to be the most important and robust mechanism for how new connections are made in social networks (Asikainen et al., 2020; Bianconi, Darst, Iacovacci, & Fortunato, 2014; Kossinets & Watts, 2009; Peixoto, 2022). Triadic closure refers to generating new connections by selecting from “friends of friends”. Empirical studies on dynamical networks find that the probability of creating a new tie is a decreasing function of the distance in the network (Bener et al., 2016; Kossinets & Watts, 2009). The less separated you are from someone, the more likely you are to form a social tie to this person. Specifically, the empirical study suggests that when you are removed one rather than two degrees of separation from someone, you are 30 times more likely to form a connection. This increase in likelihood only increases with

distance. When you are 5 degrees of separation away, you are 2.500 times less like to form a connection than you would have been were you only removed by one degree of separation (Kossinets & Watts, 2009).

Using triadic closure as the generating principle for network formation has shown great promise in explaining some of the key characteristics of social networks. In formal models, this is implemented by making new connections predominantly by triadic closure while letting a small percent of new ties be formed at random (Ilany & Akçay, 2016; Jackson & Rogers, 2004, 2007). These models can generate the important characteristic findings of high average clustering coefficient, low average path length and lognormal degree distributions (Jackson & Rogers, 2004, 2007). It is worthwhile to take time to understand why these measures are important and why these models generate the patterns that they do. Average clustering coefficient is a measure of connected local communities are in the network (Watts & Strogatz, 1998). Models of triadic closure are likely to exhibit high average clustering coefficients as triadic closure will increase the number of local links. This is closely linked to the fact that social networks are often made of tight-knitted, well-connected local communities (Peixoto, 2022). Average path length refers to average length of the shortest path between any node in the network. Here, the implemented formal models of triadic closure benefit from similar mechanisms of randomness as small-world networks to achieve low average path lengths (Jackson & Rogers, 2007; Watts & Strogatz, 1998). Having even a small amount of randomness can create long-range connections, that acts like shortcuts between nodes, decreasing the average path length substantially (Watts, 1999). Path length is important for how fast information or contagious diseases can spread between individuals in a network. Finally, triadic closure will lead to lognormal or scale free degree distributions, as high degree nodes have more possibilities for being selected for tie generation simply because they have larger degrees (Jackson & Rogers, 2007). The principle of triadic closure seems to not only characterize human social networks, but could be a robust organizing principle for many different social species (Ilany & Akçay, 2016).

## 6 Co-evolution via tie-deletion

Now that we have established the fundamental principles of opinion and network dynamics, we now turn to the mechanisms that could enable them to co-evolve. It is easy to see how opinions could be influenced by network dynamics. As the social network changes, the available interactions for each agent will change. As one of the primary forces of opinion dynamics is social influence, changes in the social network will have large effects on the opinions of the social agents of the network. This one side of the coin. As the social network already influences opinion dynamics, we need to identify mechanisms that can change the ties of the social network based on the opinions of its agents. Any mechanism that changes the ties in the network based on the opinions of the agents will cause opinion and network dynamics to co-evolve. There are good reasons to believe that there are many effects that live up to this criteria, as we have seen with the effect of observed homophily. Similarity is a strong predictor of which pairs form new ties (Bener et al., 2016; Kossinets & Watts, 2009). In the same vein, dissimilarity also predicts the breaking of social ties (Bener et al., 2016; Kossinets & Watts, 2009; Levin et al., 2021). Although both mechanisms are at play in social networks, we will focus our efforts on tie-deletion of dissimilar ties. The reason for this choice is that there are already existing

work focusing on the effects of tie-deletion in similar models (Santos et al., 2006; Sasahara et al., 2021). This allows us to examine the robustness of the effect of tie-deletion on opinion dynamics. Understanding the previous findings concerning co-evolution and the effect of tie-deletion is critical for understanding this paper. Because of this, we turn to presenting the main findings from the previous literature.

## 6.1 Tie-deletions effect on cooperation

In addition to the empirical work investigating the interplay between social networks and the opinions of their constituents, theoretical work also points to the primacy of the co-evolution of networks (Holme & Newman, 2006). These studies come specifically from another field than opinion dynamics interest in identifying the mechanisms that facilitate cooperation. The field I am referring to is the field of computational biology. In many ways, this field should be considered an adjacent field to opinion dynamics, as they share much in their interests and in their methods. The literature from computational biology suggests that co-evolution can play a vital role in facilitating cooperation (Dakin & Ryder, 2018; Melamed & Simpson, 2016; Pepper & Smuts, 2002; Santos et al., 2006). Intuitively, one might expect that when connectivity increases in such a network, cooperation increases as well. However, the opposite effect is observed. In their paper, Santos et al. (2006) considers agents playing game-theoretic games, where agents either cooperate or defect. In relation to the mechanism of tie-deletion, Santos et al. (2006) proposes a way for agents to adjust their ties based on interactions made with other agents. When cooperators interact with defectors, the social tie between cooperators and defectors has a chance of being rewired to a different agent. This keeps the number of edges constant, but decreases connectivity (Santos et al., 2006). Their results indicate that as the propensity for deleting ties between cooperators and defectors increases, cooperation flourishes. The more likely tie-deletion is, the more cooperation evolves in the system. The reason given for this result is that removing ties between cooperators and defectors increases positive assortment between cooperators and decreases positive assortment between defectors and cooperators. One of the striking attributes of this finding is that without any propensity for tie-removal, cooperation does not evolve for conditions equivalent to the Prisoner's Dilemma (Santos et al., 2006). More generally, positive assortment has been found to be a robust facilitator of cooperation in computational biology (Boyd, Gintis, & Bowles, 2010; Dakin & Ryder, 2018; Melamed & Simpson, 2016; Pepper & Smuts, 2002). The results also indicate that even relatively simple topological network dynamics that reflect the individual agent's reaction to their social interactions can produce realistic networks (Santos et al., 2006).

Although positive assortments effect on cooperation seems general in the computational biology literature, the opinion dynamics literature points to an opposite result. When combined with social influence, they report that tie deletion can lead to echo chambers, that stifle cooperation (Sasahara et al., 2021). This is more in line with a more intuitive explanation, where a decrease in the communication between agents leads to a decrease in cooperation. Obviously, computational biology and opinion dynamics are not same field, but it is still noticeable that the same underlying mechanism gives rise to effects of opposite directions on a networks ability to cooperate. Both studies are interested in how the dynamics of the network influences how well agents cooperate in the network, and they both find that tie-deletion is critical to developing cooperation (Santos et al., 2006; Sasahara et al., 2021). It is precisely because of

this dispute that this paper is especially interested in how the deletion of negative ties can impact cooperation. Like the work of Santos et al. (2006), the probability of deleting negative ties will therefore be a key aspect of our model.

## 7 Model Description

### 7.1 Conceptual overview

The model consists of agents situated in an undirected network. Nodes on the network represent social agents. Edges represent social ties to other agents. The model simulates social interactions between the agents without agents having any strategy or task.

### 7.2 Assumptions

It is assumed that the opinion of an agent is shaped only by her initial opinion and the influence of her peers. The model assumes that agents will compromise their opinion when interacting with agents with similar opinions. When agents interact with agents dissimilar from them, the model assumes that agents will either not change their opinion or they will become more opposed by the interactions. Furthermore, it is assumed that connections between agents are not static but dynamic in nature. Specifically, the model assumes that agents will find new connections primarily through their already existing connections via triadic closure. Finally, it is assumed that agents will tend to relinquish ties to agents too dissimilar from themselves. All these assumptions can vary in the strength of the proposed effect. For instance, similar agents could reach a perfect compromise or only convince each other slightly. As the results of the model critically hinges on the strength of these assumptions, I introduce model parameters that control the strength of these mechanisms.

### 7.3 Components and their properties

The model is specified by five different parameters. The first three parameters are the opinion dynamics parameters. They describe how interactions change opinions ( $\alpha, \beta, T$ ). The fourth parameter is the network dynamics parameter. It specifies how the probability of generating random ties rather than ties via triadic closure ( $R$ ). The fifth parameter is the co-evolutionary parameter. It specifies the probability of dissolving negative ties ( $P(D)$ ). The specifics of how they influence the model is specified in "Dynamics".

To limit the combinations of different values, we limit the possible values of the different parameters with the following:

$$\begin{aligned}\alpha &\in \{0.05, 0.10, 0.15, 0.20, 0.25\} \\ \beta &\in \{0.00, 0.05, 0.10, 0.15, 0.20, 0.25\} \\ T &\in \{0.8, 0.9, 1.0, 1.1, 1.2\} \\ R &\in \{0.1, 0.3, 0.5\} \\ P(D) &\in \{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}\end{aligned}$$

## 7.4 Initialization

A random Watts-Strogatz small-world graph is created with  $N = 500$ ,  $k = 7$ , and  $p = 0.5$ . For each node in the network, an agent is initialized with a value ( $O_i$ ) which is taken to represent the opinion of the agent. The opinion is modelled as a continuous number between -1 and 1. The idea is to represent that opinions can be anywhere inbetween extremely pro and extremely against a proposition. To initialize the model, we start off with a diverse set of opinions. All values of  $O_i$  are initialized by drawing from a uniform distribution between -1 and 1:

$$O_i \sim U(-1, 1)$$

## 7.5 Dynamics

### 7.5.1 Sampling agents and interactions

The model is an attempt at representing social encounters and their change over time. To represent this idea, agents interact with each other over a certain amount of time-steps. During such a time-step, an agent will always create a new connection. At every timestep ( $t$ ) a random agent ( $A_t$ ) is sampled from the network ( $N_t$ ). This agent creates a new edge to another agent in the network. As the dynamics also allows for tie-deletion, I control for the number of edges in the network. To keep the number of edges in the network approximately constant, edges are either created or rewired to account for deleted edges. To do this, let  $E_1$  be the number of edges of  $N_1$  and let  $E_t$  be the number of edges of  $N_t$ . If  $E_t < E_1$ ,  $A_t$  will not rewire one of its existing edges, but instead create a new edge. If  $E_t \geq E_1$ ,  $A_t$  will rewire one of its existing connections to the new agents. With a probability of  $R$ ,  $A_t$  connects to a random agent, that  $A_t$  is not currently connected to. With a probability of  $1 - R$ ,  $A_t$  connects to one of its edge's edges. Regardless of whether the connection was sampled randomly or through an edge's edges, let  $C_t$  define the newly connected agent.

In order to calculate the average path length of the network, the model is restricted to being only one component. Simply put, having only one component guarantees that there exists a path between any two nodes of the network. To ensure this property, we check the degree of  $A_t$  and  $C_t$ . If the degree of either  $A_t$  or  $C_t$  is 0, a new edge is created randomly to a new node from  $N_t$ . If both  $A_t$  and  $C_t$  have degrees larger than 0, and the network has more than one component, a new edge is created to restore the network. Specifically, we restore the edge that  $A_t$  rewired to  $C_t$ ,

while keeping the edge from  $A_t$  to  $C_t$ . I will refer to the process of ensuring only one connected component simply as component ensurance.

### 7.5.2 Interaction between neighbors

After finding a new connection, the agents interact with each other. This is meant to represent discussion or other social interactions between agents. Social interactions between similar individuals will lead to compromise between their initial opinion and their opinion after the encounter. As a consequence, the more you disagree with a person, the more malleable your opinion is. If you reach a compromise with a person with a very different opinion than yours, you will change your opinion more dramatically than if you reached a compromise with someone that you almost agree with. To model this kind of behavior, the randomly sampled agent,  $A_t$ , interacts with all its connected agents and changes its opinions according to the opinion of the connected agents. This is done by iteratively interacting with every edge of  $A_t$ . Let  $B$  denote an edge of  $A_t$ . Let  $O(\cdot)$  define a function with agents as inputs and their opinions as outputs. The interaction between two different agents is determined by a threshold value,  $T$ .

When  $T \geq |O(A_t) - O(B)|$ , the interaction will pull the opinions of the two agents closer to each other. The force with which they are pulled is defined as a fraction of their distance from each other:

$$V_p = (|O(A_t) - O(B)|) \cdot \frac{\alpha}{2}$$

Let  $O_{max} = \max(O(A_t), O(B))$  and  $O_{min} = \min(O(A_t), O(B))$  and update the values of opinions by:

$$O'_{max} = O_{max} - V_p$$

$$O'_{min} = O_{min} + V_p$$

When  $T < |O(A_t) - O(B)|$ , the interaction will push the opinions of the two agents further apart by a similar principle as illustrated above:

$$V_n = (|O(A_t) - O(B)|) \cdot \frac{\beta}{2}$$

Using the same definition for  $O_{max}$  and  $O_{min}$ , we update the opinions of each agent:

$$O'_{max} = O_{max} + V_n$$

$$O'_{min} = O_{min} - V_n$$

This process concludes when  $A_t$  has updated her values for all her edges.

### 7.5.3 Tie-dissolution

After interacting with her connected agents,  $A_t$  might delete the connection to some of these agents based on their similarity. In this model, social interactions can result in opinions becoming more distant than they were initially. This can be because what the agents are discussing is divisive or controversial. When this happens, connections can be severed. I call this process tie-dissolution. If  $T < |O(A_t) - O(B)|$ , there is a chance that their tie is dissolved after interacting. This probability of dissolution between dissimilar agents in the model is described by the parameter  $D$ . When  $D = 1$ , all dissimilar ties are severed. When  $D = 0.5$ , there is 50% chance that dissimilar ties will be deleted. After ties are dissolved, the process of component insurance is performed where  $C_t$  is replaced with  $B$ . When all edges of  $A_t$  have been evaluated, the time-step concludes. This process is repeated for 10.000 time-steps.

## 7.6 Outcome metrics

### 7.6.1 Time-dependent Metrics

Every 20th time-step, the current state of the network is recorded. To track the polarization of opinions over time, the mean and standard deviation of the absolute value of opinions are recorded. To track the similarity of an agent's opinion to the opinion of their neighbors, the average distance to all edges are recorded. To evaluate the effect of tie-dissolution, the cumulative frequency of tie-dissolutions are recorded. For characterizing the network, I record the average clustering coefficient, average path length and degree assortativity coefficient.

### 7.6.2 Final State Metrics

After 10.000 time-steps, the network reaches its final state and its characteristics are recorded. For every agent in the network, I record:

- The initial opinion of the agent
- The opinion of the agent at the final state of the network
- The mean distance to all neighbors' opinions
- The degree of the agent
- The betweenness centrality of the agent
- The clustering coefficient of the agent

The measures related to the opinion of agents can then be correlated with clustering, centrality and degrees to see how changes in network topology relate to changes in opinions.

## 8 Model Evaluation

### 8.1 Fitting the model

As previously mentioned, previous models have often not integrated empirical data of social networks sufficiently to test the validity of their models (Flache et al., 2017). What is meant by this critique? Arguably, the most pressing issue is that models rarely consider how the results of their computational models compare with the real world. This is largely a product of the types of agent-based models typical in opinion dynamics as well as the phenomena of opinion dynamics themselves. Most models of opinion dynamics are low-dimensional (Bener et al., 2016), in that they don't seek to explain exhaustively how opinions are formed, but rather to point out interesting interactions between key variables of the process. This makes empirical validation especially tricky for these models. Compounding to this problem is that the time-scales considered by these models are massive, and that kind of data is very hard to come by. And even if they were available to us, opinions are very hard to measure. Any self-reported measure is going to be subject to the usual problems of the unreliability of self-reports. Using political voting patterns might be a usable proxy for polarization, but these data are very hard to come by. This does not however mean that incorporation of empirical data is impossible. Because the model considered in this paper is coevolutionary, it allows us to investigate how the generated networks compare to real empirical social networks. By borrowing methodology from computer science, we will treat this as a machine learning problem, and investigating the goodness of fit using hyperparameter optimization (Akiba et al., 2019; Bergstra et al., 2011). What we really want to know is to what degree the model can explain the patterns observed in real world social networks. To do this, we want to estimate the model parameters by generating networks from the model, and then see how much these generated networks resemble real networks. I will refer to the process of finding these parameter values as agent-based model fitting, as it is analogous to estimation of parameter values in statistics. Beyond knowing whether the model can generate the observed structure of real-world networks, this paper is specifically interested in how critical co-evolution is to the explanatory power of these models. Ideally, we want to compare models which include co-evolution to models without co-evolution and investigate whether co-evolutionary models are better explanations than their network formation counterparts. I will refer to this process as agent-based model comparison, which is again reasoned by its analogy to model comparison in statistics. Although these steps are analogous to model fitting and model comparison, there are important differences that are worth mentioning. Especially important is that model comparison normally penalizes the inclusion of additional parameters in the model (Vrieze, 2012). In other words, simple explanations are preferred when the performance is comparable (Emiliano et al., 2014; Vrieze, 2012). The primary reason for this is that including additional parameters allows the model to overfit to the data. In classical regression models, models with additional parameters will always give a better fit, as the scalar for the new parameter can simply be set very close to zero. In this case, the model becomes mathematically identical to the simpler explanation. Although we don't have a formal way of calculating agent-based equivalents to Akaike Information Criteria or Bayesian Information Criteria to determine which model is the better explanation (Vrieze, 2012), we can try to avoid situations of mathematical identity and overfitting. This is done by framing the problem of agent-based model fitting as a hyper parameter optimization problem.

## 8.2 Data

To fit the model to data, we consider 6 empirical social networks; a network of representing social networks of Dolphins (Lusseau, 2003), the Karate Club Network (Zachary, 1977), a citation network (Newman, 2006), co-purchase network of political books on Amazon (Shi et al., 2017), political blogs (Adamic & Glance, 2005) and network of politicians based on shared likes on Facebook (Rozemberczki et al., 2018). For all networks, only the main component of the network was considered. This was done as average path length is not well-defined for unconnected graphs. As the model assumes undirected networks, the network of political blogs was transformed from a directed to an undirected network. This was done by making any directed link between two agents be undirected.

## 8.3 Hyperparameter Optimization

To evaluate how closely the generated networks resemble the real networks, we must first define a distance metric between networks. Although many features are important for characterizing a network, we will follow previous suggestions that identify the average path length, the average clustering coefficient and the degree distribution as the most defining aspects of a network (Jackson & Rogers, 2004, 2007). We therefore seek to minimize the mean difference on these parameters between the generated and the real networks. Before we can simply calculate the mean of difference, there are some specifics that needs to be clarified. These primarily concern the average path length and how to operationalize the difference in degree distributions. To identify the distance of degree distributions of networks, we calculate the Jensen Shannon Divergence between the two distributions. The Jensen Shannon Divergence is a distance metric between probability distributions (Fuglede & Topsøe, 2004). In terms of the average path length, there is a problem of computational costs. Calculating the average path length can be very expensive computationally, especially for large networks (Matsumura et al., 2018). Instead of calculating the actual average path length, we approximate it by taking 1000 samples of random pairs of nodes and calculating the average path length of those 1000 nodes. Although this is still only an estimation of the actual average path length, this method generally approximates the average path length fairly well (Matsumura et al., 2018). Both the correlation coefficient and the Jensen Shannon Divergence are values between 0 and 1. Average path length on the other hand is not restricted in the same way. If it is not normalized, it will therefore be weighted much higher than the correlation coefficient and the Jensen Shannon Divergence in the distance metric. We therefore normalize the average path length by

$$APL^* = \frac{|APL(G) - APL(A)|}{APL(A) + 2}$$

where  $APL^*$  is the normalized average path length,  $APL(G)$  is the approximated average path length of the generated network and  $APL(A)$  is the apprioximated average path length for the actual network. Adding two to the denominator is somewhat arbitrary. The choice was made to be more certain that  $APL^*$  would still be below 1. With this in place, the final distance metric considered between the actual network,  $A$  and the generated network  $G$  is

$$O(A, G) = \frac{1}{3} \cdot (JSD(D(A), D(G)) + |C(A) - C(G)| + APL*)$$

where  $O(A, G)$  is the distance between two networks,  $JSD$  is the Jensen Shannon Divergence,  $D(A)$  is the degree distribution of  $A$ ,  $D(G)$  is the degree distribution of  $G$ ,  $C(A)$  is the average clustering coefficient of  $A$  and  $C(G)$  is the average clustering coefficient of  $G$ . This distance metric is just the average difference on the considered network characteristics between  $A$  and  $G$ . This is also why the sum of the difference is multiplied by  $\frac{1}{3}$ . This distance metric between networks,  $O(A, G)$  will serve as the objective function that we want to minimize. The hyperparameter optimization algorithm search the parameter space for the combination of parameters that results in the lowest values of  $O(A, G)$  (Akiba, Sano, Yanase, Ohta, & Koyama, 2019; Wu et al., 2019).

All models were fit by first calculating the ratio between edges,  $E$ , and vertices,  $V$ , and calculating  $K = \lfloor (2 \cdot \frac{E}{V}) \rfloor$ . The value  $K$  is then used as the  $K$ -parameter in the initially generated small world network to ensure that the generated network has approximately the same number of edges per vertex as the target network. For both models, we let randomness be a free parameter between 0 and 1. For the co-evolutionary model, we restrict the threshold to be between 0.1 and 1.3. The lower limit is chosen as values of threshold close to 0 results in agents that don't agree with anyone else but themselves. The upper limit was chosen to ensure that the behavior of the model differs significantly from the model without opinion dynamics. The values of positive learning rate are restricted to be between 0.05 and 0.5. This is justified by values close to 0 being akin to no positive social influence and values above 0.5 assumes social influence to be an unrealistically powerful force. Similarly, we let the negative learning rate be restricted to be between 0 and 0.5. Here, we allow values close to 0 to reflect that negative social influence is still a less established fact. Finally, we force the model to exhibit co-evolution by restricting the probability of tie-deletion to be between 0.1 and 1.

Fitting the model to the different networks estimates the best combination of parameters for each given network. These parameters are then used to simulate 10 different networks using different random seeds but the same parameter values. This is done as the model is largely stochastic. Having 10 simulations of the same parameters gives a more robust estimate of how closely the model can recreate the patterns from the empirical networks. After finding the best parameter values, the importance of each parameter value and each network is calculated using a fANOVA (Hutter, Hoos, & Leyton-Brown, 2014).

## 8.4 Results

The results indicate that the network formation algorithm of triadic closure can explain many of the considered characteristics of social networks. The model without opinion dynamics performs well on most networks and is on par with the model with opinion dynamics. It even performs better than the model with opinion dynamics in the Karate Club network. On the political networks investigated here, the model with opinion dynamics performs much better. Parameter importance shows that the randomness parameter was the most important parameter for all networks except for the network of politicians and political blogs. In these networks, the most important parameter is instead the threshold.

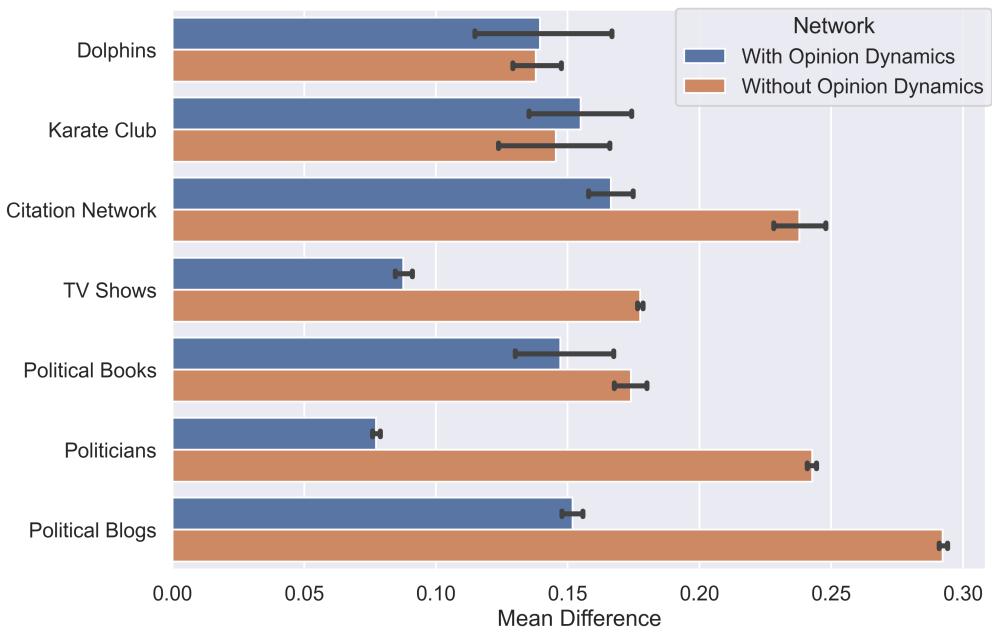


Figure 1: Difference in model performance. X-axis shows the mean difference between the generated network and actual networks, ( $O(A, G)$ ). Y-axis shows the different empirical networks considered. Colors indicate whether the model includes or excludes opinion dynamics. Black lines indicate 95% confidence intervals.

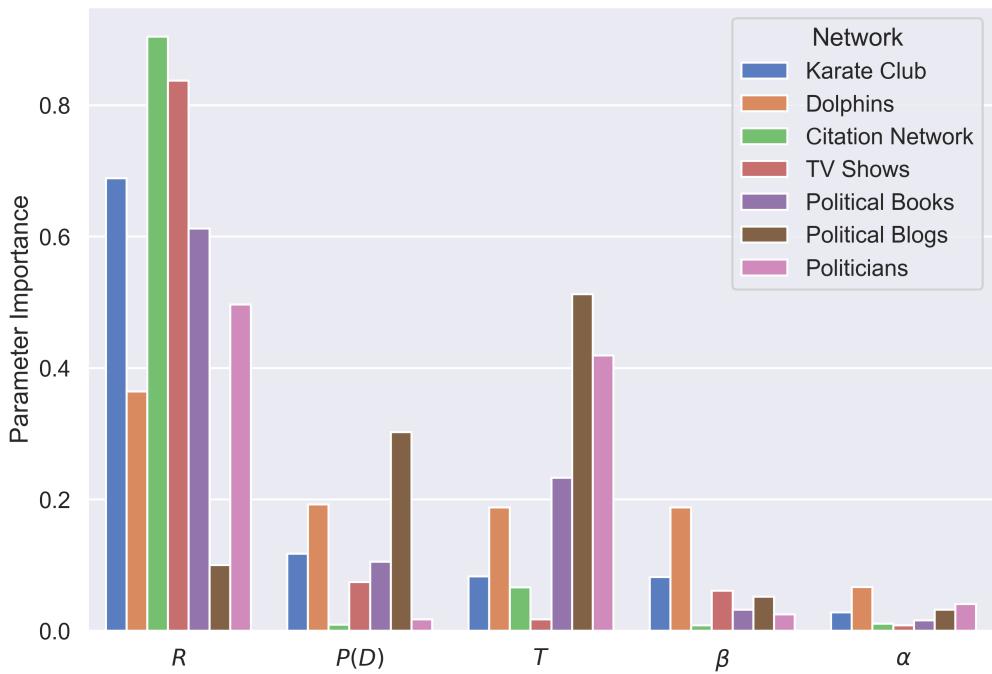


Figure 2: Hyper Parameter Importance. X-axis shows different hyperparameters. Y-axis shows parameter importance. Colors indicate different networks. Parameter importance is calculated using an fANOVA (Hutter et al., 2014).

Table of best parameter values					
Network	R	T	$\alpha$	$\beta$	$P(D)$
Dolphins	0.003	0.531	0.202	0.095	0.111
Karate Club	0.870	0.137	0.282	0.090	0.911
Citation Network	0.060	0.655	0.389	0.493	0.150
TV Show	0.038	0.157	0.349	0.093	0.751
Political Books	0.004	0.222	0.075	0.144	0.508
Politicians	0.014	0.101	0.050	0.014	0.229
Political Blogs	0.109	0.100	0.246	0.007	0.249

Table 1: Table of best parameter values

Table of network characteristics						
Network	N	E	$D_\mu$	$D_\sigma$	C	$L_G$
Dolphins	62	159	5.129	2.932	0.259	3.257
Karate Club	34	78	4.588	3.820	0.571	2.326
Citation Network	379	914	4.823	3.927	0.741	6.052
TV Shows	3892	17262	8.871	12.557	0.374	6.324
Political Books	105	441	8.400	5.449	0.488	3.015
Politicians	5908	41729	14.126	20.096	0.385	4.628
Political Blogs	1222	16717	27.360	38.402	0.320	2.737

Table 2: Table of network characteristics.  $N$  is the number of nodes,  $E$  is the number of edges,  $D_\mu$  is the mean degree,  $D_\sigma$  is the standard deviation of the degree distribution,  $C$  is the average clustering coefficient, and  $L_G$  is the average path length.

## 8.5 Why is co-evolution a better explanation?

Of particular interest is the political networks, as we know that these networks are likely to be highly opinionated. As the model performance shows, these are also the networks where the difference in performance between the network formation and the co-evolutionary model become most stark. Most notably, the difference in performance for the network of politicians and of political blogs is drastic. It is therefore worthwhile to consider why the co-evolutionary model fits these types of networks so much better than the network formation model. One of the primary reasons for why the co-evolutionary model performs so much better on political social networks is that it enables the network to have multiple central hubs. By letting the threshold,  $B$ , be relatively low, the network balkanizes into different communities, exhibiting connected caveman properties (Watts, 1999). This is also in line with the fact that these two networks are the only networks considered, where threshold is the most important parameter. This suggests that

the political networks are organized by clearly delineated communities. The balkanization obviously affects all the considered network characteristics considerably. Having low values of thresholds makes for networks with a much larger average path length than if all nodes were in one cluster. As the communities become smaller in size, this affects the degree distribution as well. When the network is no longer organized around one central cluster, the resulting degree distribution will exhibit a shorter tail, as agents cannot connect to a large portion of the agents in the network. When a network is characterized by larger average path length and shorter tail degree distributions, the co-evolutionary model can outperform the model without opinion dynamics. Because of the co-evolutionary model's ability to create multiple hubs, it can capture these phenomena well. In contrast, the network formation model only has one parameter,  $R$ . It has no way of reliably creating multiple clusters in the network, and therefore it fails when it encounters networks with clearly separated communities. The network formation model has to balance average path length and high clustering coefficient, as it can only vary how random new connections are, similar to small world networks (Watts & Strogatz, 1998). The co-evolutionary model can change the size of different clusters by adjusting the confidence boundary and the tie-deletion parameter. This explanation seems to be in line with the importance of the different parameters for the model fit. A congruent point comes from considered the best parameters for the different political social networks. Most of them are characterized by very low confidence boundaries, very low amount of randomness, very low negative learning rate and moderate values for tie-deletion. As we've established, low confidence boundaries reflect that the political scene consists of different groups, that are locally very well connected, but doesn't have much to do with each other. To achieve the high average clustering coefficients, the randomness must be low. For the network to not just become two clusters, negative learning rate must be low, so that opinions do not cluster around 1 and -1. Finally, to ensure that there is still some global connectivity between the different clusters, tie deletion is not high enough to limit the network to only have one or two informational highways.

In conclusion, the agent-based model fitting and model comparison suggests two important notions. The first notion is that triadic closure alone seems to be able to generate many of the characteristic patterns of social network. This echoes previous findings concluding the same (Ilany & Akçay, 2016; Jackson & Rogers, 2004, 2007). The second notion is that when opinions are central to the social networks, including opinion dynamics makes for a markedly better model of why social networks are like they are. This second notion points to the fact that the effects of network and opinion dynamics are largely interdependent. We have seen that opinion dynamics can drastically alter the dynamics of networks. We now turn to the other side of this interplay by considering how network dynamics can change opinions.

## 9 Results

### 9.1 Alpha, beta, threshold, tie-deletion and chaotic systems

Higher values of the threshold decrease polarization. Higher values of alpha decreases polarization. Higher values of beta increases polarization. Notably, when beta is zero, almost all simulations reach consensus, regardless of all other parameters. Similarly, there is no effect of negative tie-deletion when  $\beta = 0$ . The drivers of polarization is low

values of  $T$  and high values of  $\beta$ . When  $T$  is sufficiently small and  $\beta$  is sufficiently large, all agents will have more extreme opinions at the end of the simulations than they had in the beginning. With intermediate values of  $\beta$  and  $T$ , whether the opinions of a simulation becomes polarized or not is largely dictated by  $P(D)$ . The results indicate that higher values of  $P(D)$  results in lower values of  $|O|$  on average. The strength of the effect is modulated by how random new connections are made in the network. When randomness is higher, the effect of cooperation facilitation by high tie-deletion is diminished. Notably, when  $P(D) = 0$  and  $\beta > 0$ , the network almost always polarizes.

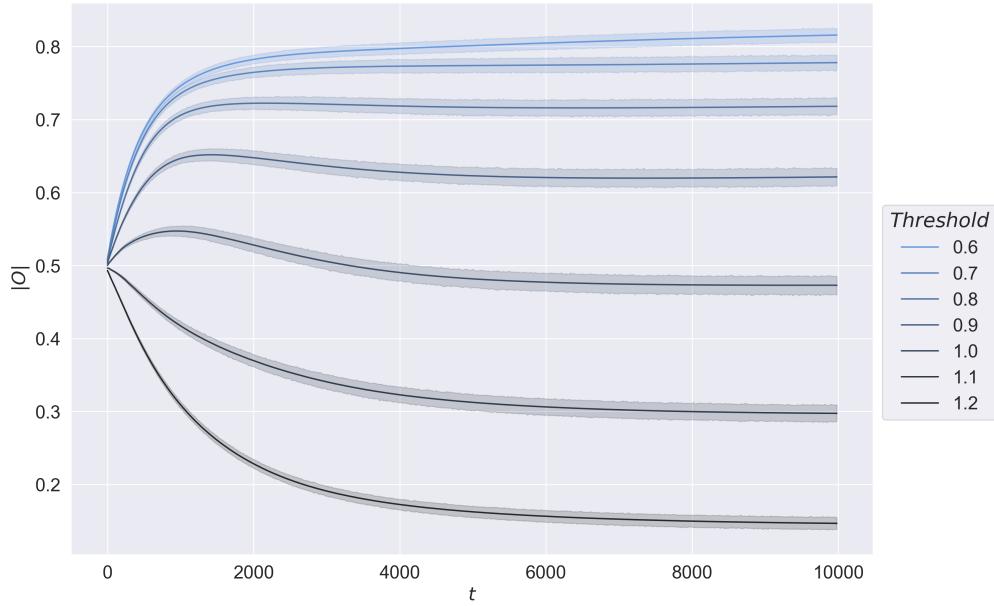


Figure 3:  $T$ 's effect on polarization over time. X-axis shows the time-step and the y-axis shows the absolute value of the opinions of agents. Colors indicate different values of  $T$ . Lines indicate the mean value of each time-step, aggregated over all parameters excluding  $T$ . Shaded regions are the 95% confidence intervals.

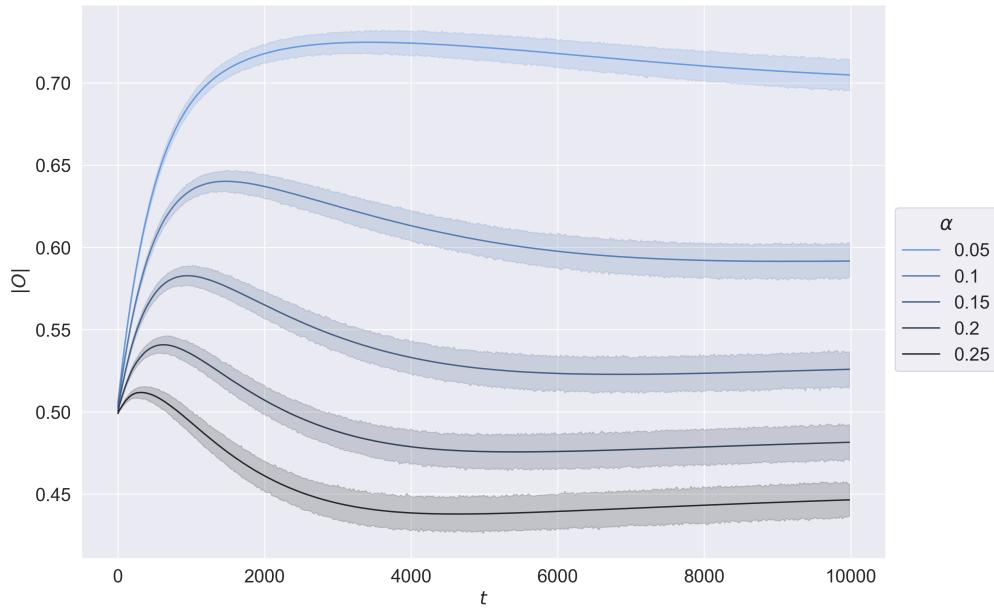


Figure 4:  $\alpha$ 's effect on polarization over time. X-axis shows the time-step and the y-axis shows the absolute value of the opinions of agents. Colors indicate different values of  $\alpha$ . Lines indicate the mean value of each time-step, aggregated over all parameters excluding  $\alpha$ . Shaded regions are the 95% confidence intervals.

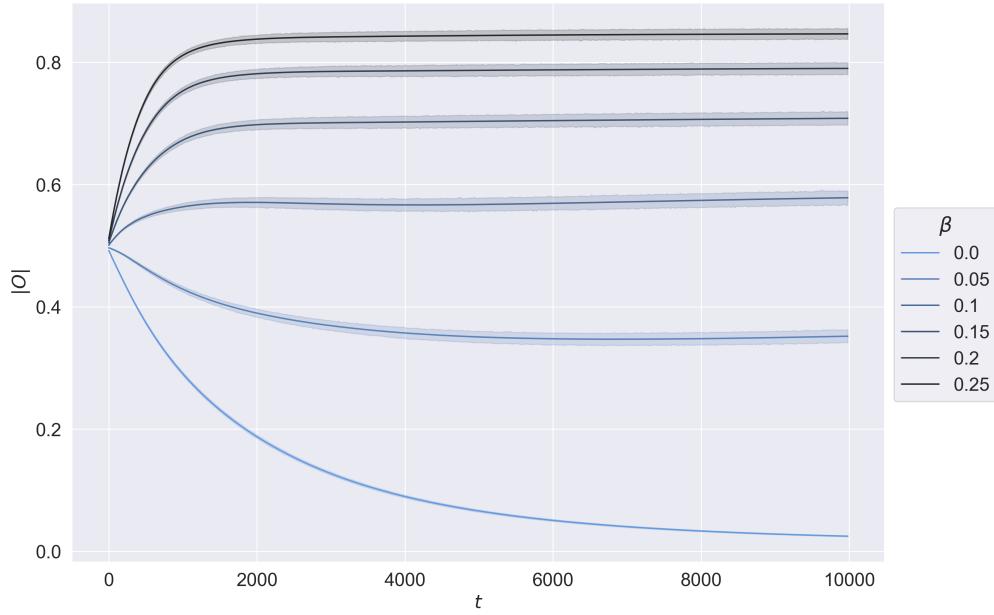
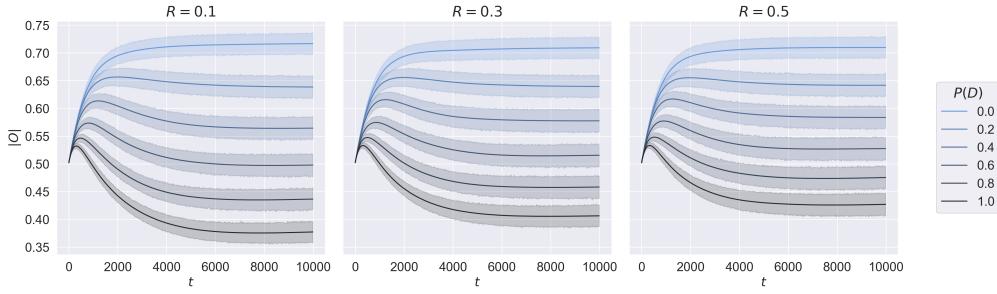
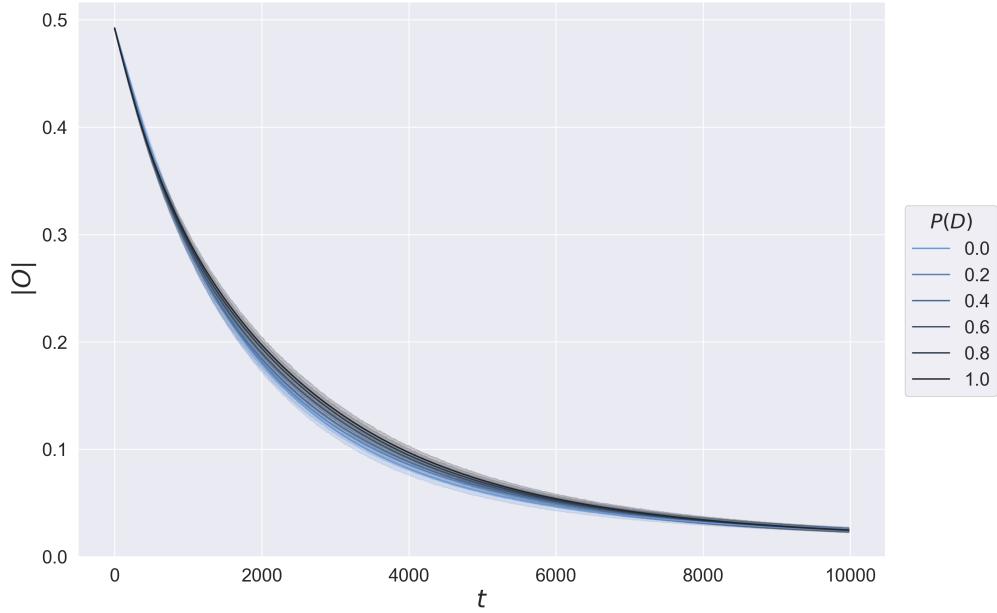


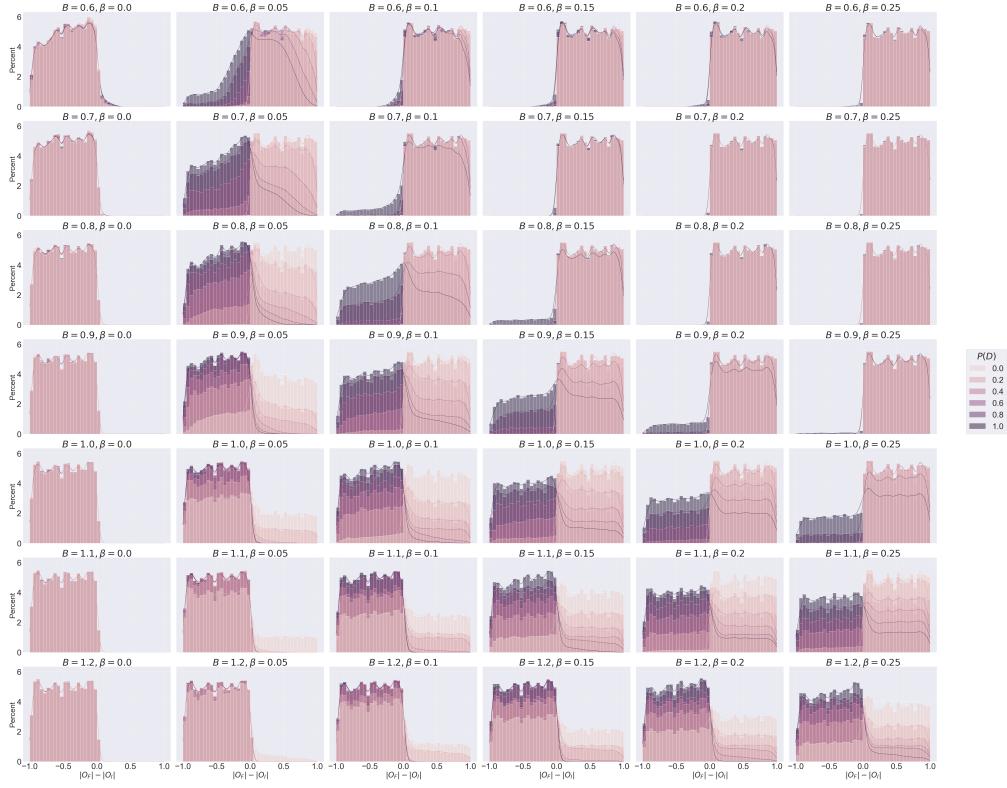
Figure 5:  $\beta$ 's effect on polarization over time. X-axis shows the time-step and the y-axis shows the absolute value of the opinions of agents. Colors indicate different values of  $\beta$ . Lines indicate the mean value of each time-step, aggregated over all parameters excluding  $\beta$ . Shaded regions are the 95% confidence intervals.



*Figure 6:*  $P(D)$ 's effect of polarization over time. X-axis shows the time-step and the y-axis shows the absolute value of the opinions of agents. Colors indicate different probabilities of tie-deletion of dissimilar agents. Lines indicate the mean value of each time-step, aggregated over all parameters excluding the probability of tie-deletion and randomness. Shaded regions are the 95% confidence intervals. The three different plots show the different values of probability of new connections to be generated randomly.



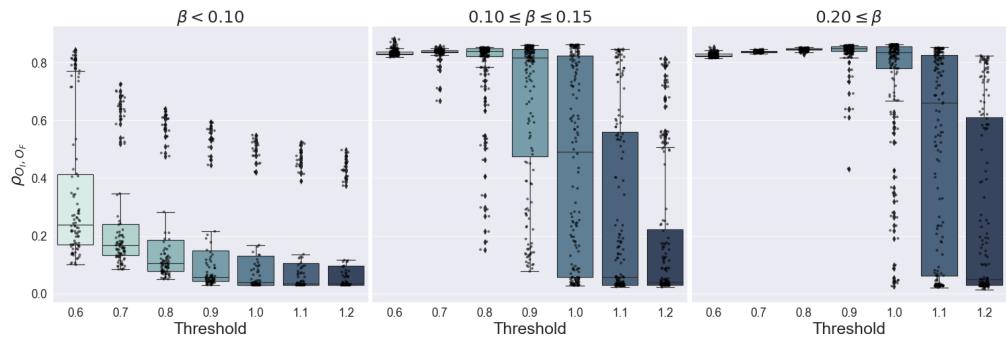
*Figure 7:* The effect of  $P(D)$  without negative learning. X-axis shows the time-step and the y-axis shows the absolute value of the opinions of agents. Colors indicate different probabilities of tie-deletion of dissimilar agents. Lines indicate the mean value of each time-step, aggregated over all parameters with  $\beta = 0$ , excluding the probability of tie-deletion. Shaded regions are the 95% confidence intervals.



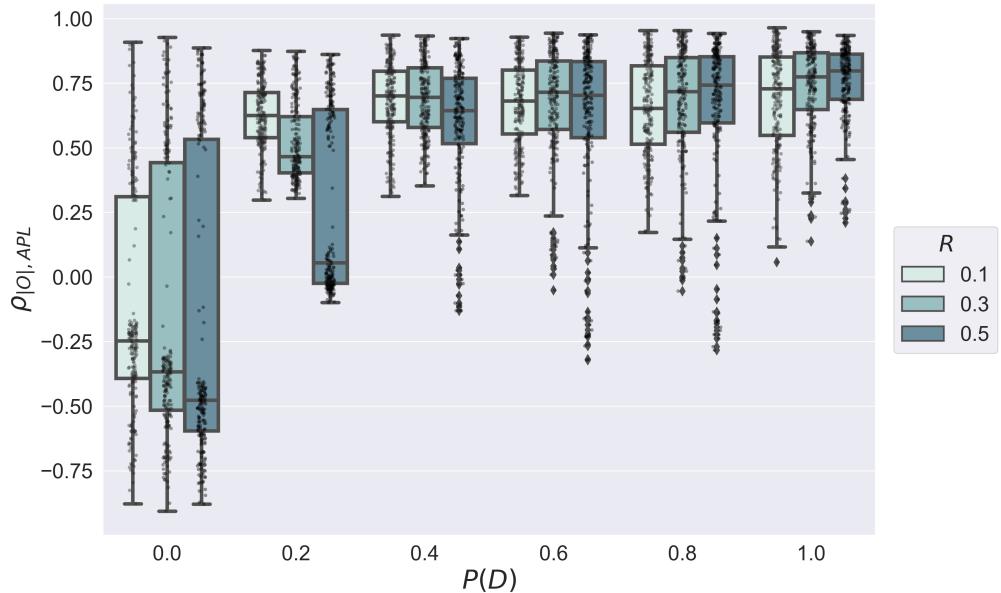
*Figure 8:* Distribution of opinion changes across different conditions. Rows display different values of thresholds, columns indicate different values of negative influence. Different colors show different values of probability of tie-deletion. X-axis shows the difference between the absolute value of the final and initial opinions of agents. This measure can be thought of as a measure of “radicalization”. Higher values indicate that the final value of agents is more extreme than the initial values, and vice versa. Y-axis exhibit the percent of values within each bin.

## 9.2 Correlations

We calculate two Pearson Correlation Coefficients of interest for all simulations. The first of which is the correlation coefficient between the initial opinions and the final opinions of agents,  $\rho_{O_I, O_F}$ . The second is the correlation between the absolute value of opinions and the average path length,  $\rho_{|O|, APL}$ . We start by reporting the findings for  $\rho_{O_I, O_F}$ . When  $\beta$  is low and threshold is high, there is little correlation between initial and final opinions. When the opposite is true, final and initial values are highly correlated. Next, we turn to  $\rho_{|O|, APL}$ . As the probability of tie-deletion increases, so does the correlation between the average path length of the network and the absolute value of opinions. When the probability of tie-deletion is 0, there is no consistent correlation between the two variables.



**Figure 9:** Correlations between initial opinions and final opinions. X-axis shows different values of threshold and the y-axis shows the Pearson Correlation Coefficient between the initial and final opinion of agents. Boxes indicate the quartiles, whiskers indicate the 1.5 of the inter-quartile range outside of the quartiles. Points outside this range is depicted as diamonds. Smaller dark dots show the correlation coefficient for individual simulations. Columns show low, medium and high values of  $\beta$ .



**Figure 10:** Correlations between absolute opinion and average path length. The Pearson Correlation Coefficient between the initial and final opinion of agents was calculated for all simulations. Boxes indicate the quartiles, whiskers indicate the 1.5 of the inter-quartile range outside of the quartiles. Points outside this range is depicted as diamonds. Smaller dark dots show the correlation coefficient for individual simulations. First plot has negative learning rate on the x-axis. Second plot shows the threshold on the x-axis.

## 10 Discussion

### 10.1 Reiterating previous literature

We first discuss the main effects of  $\alpha$ ,  $T$  and  $\beta$ . All of these are intuitive and are in line with previous research on similar models (Flache et al., 2017). When  $\alpha$  is high, there is a stronger force pulling the opinion of agent's closer together. Higher values of  $\alpha$  therefore results in lower values of polarization. But as can be seen from the scale of the y-axis, although this is large effect, it is small in comparison to the effect of  $\beta$  and  $T$ . Higher values of  $\beta$  has more drastic effects on the polarization of the population. This is primarily caused by the considered values of  $T$  and its operationalization. When  $\beta = 0$ , all simulations reach a consensus. If this simulation had considered lower values of  $T$ , this would most likely not be true. Lower values of  $T$  are known to be able to produce clusters of opinions (Flache et al., 2017; Sasahara et al., 2021). When  $\beta$  is large, almost all simulations polarize. The reason for this is partly that in its operationalization, it is a stronger force than  $\alpha$ . The effect of pushing and pulling opinions is defined as being proportional to the distance between agents. By definition, agents that are pushed away from each other are further apart than agents that are pulling closer together. In this sense,  $\beta$  is a stronger force than  $\alpha$ . Only in aggregate can the effect of  $\alpha$  become stronger than  $\beta$ . This is only possible when the number of positive interactions is much higher than the number of negative interactions. The effect of  $T$  is therefore not surprising. With larger values of  $T$ , agents are more "open-minded" - they can cooperate with more diverse agents, and will therefore have much more positive interactions than negative. Notice that if an agent has an opinion of 0 and  $T \geq 1$ , this agent will not be able to have any negative interactions. The opposite effect is witnessed when  $T$  is low. This limits the subset of agents you can cooperate with, which makes negative interactions much more plausible.

### 10.2 The effect of initial opinions

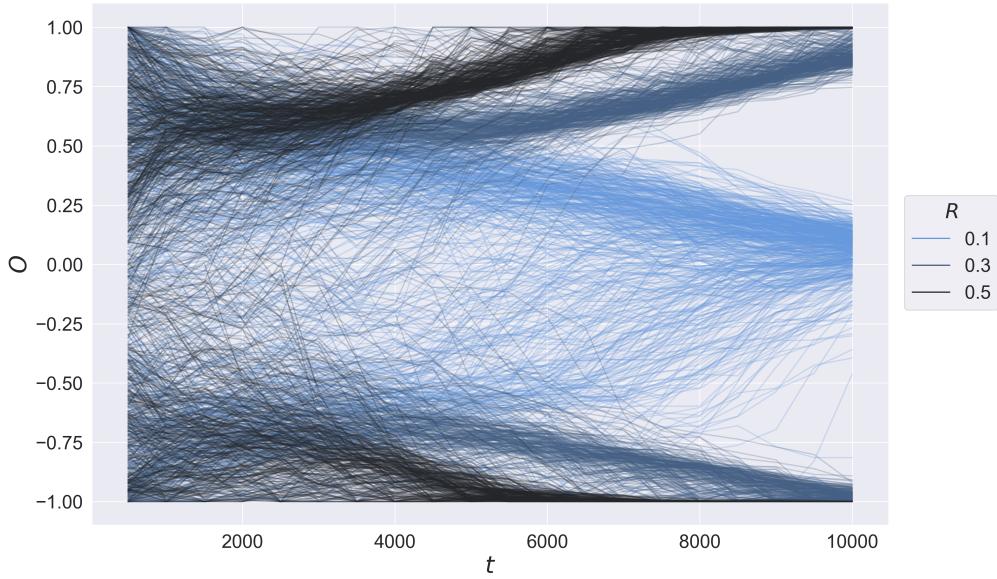
The results indicate that in some conditions, there is a strong positive correlation between starting opinion of an agent and the final opinion of the agent. These are conditions of high negative influence and of low thresholds. These conditions are likely to lead to high degrees of polarization. The results indicate that in polarized conditions, agents will tend to move less than in conditions where they agents reach consensus. In conditions where agents do reach consensus, your starting position does not matter much. All agents will tend to be close to zero, no matter the starting condition. In polarizing conditions, agents will be very unlikely to move far from their starting position and will instead be pushed to the nearest extreme.

### 10.3 Tie-deletion facilitates cooperation

Finally, one of the main findings of this paper is that higher probabilities of tie-deletion can prevent polarization and facilitate cooperation. This is especially the case when new ties are created via triadic closure and not randomly. This

happens because of two effects in the system. The first is that deleting negative ties will cause fewer negative social interactions, which limits the effect of  $\beta$ . When negative ties are deleted, agents will assort based on their opinions. Deleting negative ties is important in its own right because it deletes negative ties, that would otherwise cause further polarization. But because of assortment, the effect of tie-deletion interacts with how random new ties are created. When assortment on opinions is high, the heuristic that "a friend of a friend is also friend" is very likely true. Triadic closure will result in even more positive interactions, because similar to what is found in empirical networks, distance in the social network is highly indicative of distance in similarity (Kossinets & Watts, 2009). When randomness is high, the probability of connecting to a very dissimilar agent is much higher than through triadic closure. In other words, increasing randomness decreases assortment, which leads to more negative encounters. This also echoes the finding from the opinion dynamics literature stating that when new connections are made randomly, opinions will tend to polarize more quickly when they include negative influence (Flache & Macy, 2006, 2011; Turner & Smaldino, 2018). Here we show that reason for this effect is largely due to positive assortment of opinions. Positive assortment of opinions is also closely linked to the reason for the correlation between the absolute value of opinions and the average path length when tie-deletion is high. Because the network co-evolves with opinions when the average distance between opinions increase, the average distance in the network will increase as well. When opinions polarize, negative ties will tend to be deleted. Because of positive assortment, long-range ties will suffer especially when the network polarizes. This will make highways of information which connects distant parts of the network much less likely, which increases the average path length substantially. When the network depolarizes, the opposite happens. Some long range ties will be made by random, and because the network is depolarized, these will stay intact. This will make further long-range connections more probable, as the two clusters can connect via triadic closure.

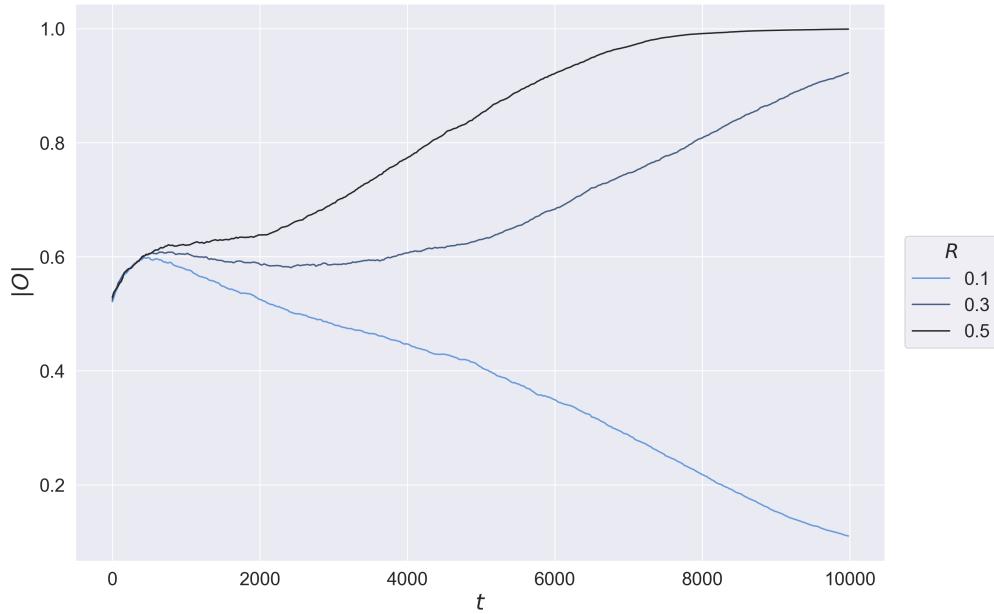
## 10.4 Distance in opinion and network space



*Figure 11:* Lineplot of opinions over time. X-axis shows time-steps of the simulation, y-axis shows the opinion of agents in the simulation. Colors show different values of  $R$ . All simulations were  $T = 0.8$ ,  $\alpha = 0.15$ ,  $\beta = 0.1$ ,  $P(D) = 1$  and with the same random seed.

To further appreciate why tie-deletion facilitates cooperation, it is beneficial to examine representative simulations in more detail. Here we investigate one of the conditions, where threshold is relatively low ( $T = 0.8$ ), there is some negative social influence ( $\beta = 0.1$ ) and positive social influence ( $\alpha = 0.15$ ). This combination of parameters only avoids polarization when tie-deletion is high for the reasons highlighted in the previous section. By focusing on these cases, we can inspect directly how the network evolves over time in relation to the opinions of its agents. The first thing to notice is the stark difference between the trajectory of opinions of different levels of randomness. Only the simulation with  $R = 0.1$  avoids polarization. Additionally, when  $R = 0.5$ , the polarization is much faster than when  $R = 0.3$ . The second thing to notice is how closely the average path length follows the absolute value of opinions. As mentioned, the mechanism that links these two variables is positive assortment. After only 2000 time-steps, there is a clear gradient in the networks, which is the result of positive assortment. Positive assortment is achieved directly by negative tie-deletion in this model. Additionally, positive assortment is increased indirectly by triadic closure. This is the same mechanism referred to as structural homophily in the literature (Asikainen et al., 2020; Peixoto, 2022). In other words, tie-deletion is not necessarily an important mechanism in itself. Rather, it is important because it increases positive assortment. I presume that any mechanism that will increase positive assortment in these kinds of models will decrease polarization. This point is also strengthened by the difference between the simulations considered here. Increasing randomness causes the network to polarize, because it decreases positive assortment. Counterintuitively, tie deletion in combination with triadic closure facilitates cooperation because it isolates groups, not despite it. Here we can also clearly see that polarization in opinions causes the network to become less connected, which increases the path length. When tie-deletion is high, the network behaves analogous to a spring. As long-range ties are deleted, the

tension in the spring rises, and the network starts to be better characterized as a network of two factions, connected mostly by a group of agents with values around 0. As distant parts of the network reconnect, the tension of the spring is lowered. When polarization increases, there will be less of these long-range ties, but there might also be fewer middlemen to facilitate cooperation across the aisle. These middlemen are the nodes of cyan colors, that lie in the middle of the two major clusters of dark and light colors. These agents are critical for cooperation. Without them, the different clusters cannot have positive interactions with each other as they are too far apart. By interacting only through these middlemen, there is a chance of bridging the gap. If these middlemen become part of either cluster before polarization has decreased sufficiently for the different groups to have positive interactions, the only possible outcome is polarization. Middlemen are therefore necessary conditions for cooperation in this model. It is also possible to give an alternative interpretation of what the probability of tie-deletion represents. The probability of tie-deletion can also be seen as how quickly the structure of the network is adjusted by opinions. In this interpretation, consensus can be achieved if the social network adjusts fast enough. Having faster adjustment of the structure of social interactions is critical, because the system is highly path dependent (Turner & Smaldino, 2018). When the network doesn't adjust fast enough, agents will already have had several negative interactions, which will have pushed them further towards an extreme value of opinions. This will make them more likely to polarize other agents in the future. This will also be the case for middlemen. Adjusting too slowly can therefore burn the only bridge that is available for facilitating cooperation.



*Figure 12:* Correlations between absolute opinion and average path length. The Pearson Correlation Coefficient between the initial and final opinion of agents was calculated for all simulations. Boxes indicate the quartiles, whiskers indicate the 1.5 of the interquartile range outside of the quartiles. Points outside this range are depicted as diamonds. Smaller dark dots show the correlation coefficient for individual simulations. First plot has negative learning rate on the x-axis. Second plot shows the threshold on the x-axis.

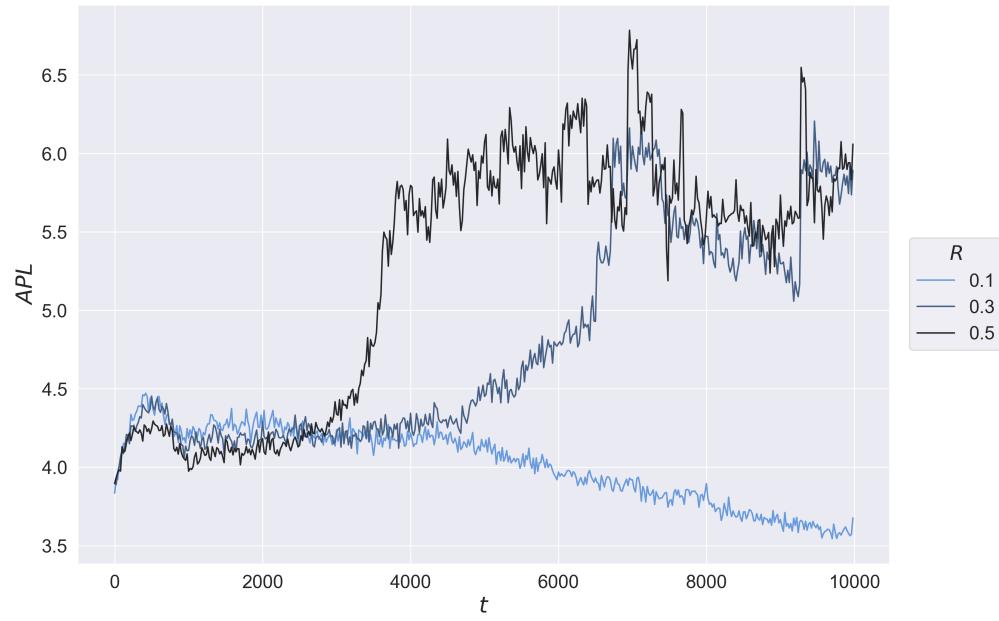


Figure 13: Correlations between absolute opinion and average path length. The Pearson Correlation Coefficient between the initial and final opinion of agents was calculated for all simulations. Boxes indicate the quartiles, whiskers indicate the 1.5 of the interquartile range outside of the quartiles. Points outside this range is depicted as diamonds. Smaller dark dots show the correlation coefficient for individual simulations. First plot has negative learning rate on the x-axis. Second plot shows the threshold on the x-axis.



*Figure 14:* Network evolutions for different levels of randomness. Columns show different values of randomness,  $R$ . Rows are different time-steps of the simulations, increasing in increments of 2000. Agents are colored by their opinion. Dark blue colors represent values close to -1, and lighter colors represent opinions close to 1.

## 11 Contrast to previous work

It is important to understand why the results found here are in so stark contrast to the recent opinion dynamics models, which proposes that unfollowing dissimilar individuals accelerates polarization (Sasahara et al., 2021). First, the network they are considering is a directed network, meaning that relations need not be reciprocal. This could have important effects for the effects observed here, because it allows for social influence to be asymmetric. Secondly, the number of agents considered is only 100 instead of the 500 considered here. This could also be consequential. Tie-deletions can only facilitate cooperation if there are "middlemen", that can reunite the different clusters. Having enough diversity in opinions is critical for bridging the gap between distant opinions. Having more agents in the situations increases the number of expected agents with values close to 0, which can help in facilitating cooperation between agents of too different opinions. Finally, a critical difference between the model proposed here and the model that they considered is that this model includes negative social influence,  $\beta$ . Because of this, social relations that are outside of your boundary of confidence have no influence on you opinion. In other words, there are no "costs" associated with negative social interactions. For the parameter values considered in the model of this paper, tie-deletion is only effectively facilitating cooperation when there is negative social influence in the system. The reason for this is primarily that this paper does not consider very low values of the confidence boundary as it somewhat obscures the analysis, as the system's behavior changes qualitatively below these values. Although their model is also a variation of a classic confidence boundary model (Flache et al., 2017; Sîrbu, Loreto, Servedio, & Tria, 2017), like the one proposed in this paper, they analyze much lower levels of boundaries of confidence. This has benefits and drawbacks. First, this allows more than two clusters of opinions. Second, it allows for the clusters of opinion to be centered at other values than the extreme of the opinion spectrum (Sasahara et al., 2021). These are definite positives, and similar features can also be generated by the model of this paper. However, the behavior of the system changes drastically when values of very low levels of boundary thresholds are studied, which makes a comprehensive review of the system as a whole much more complex to study. This echoes previous findings from two different literatures. First, facilitating cooperation via tie deletion recapitulates findings from computational biology, which states that positive assortment leads to increased levels of cooperation (Boyd et al., 2010; Dakin & Ryder, 2018; Pepper & Smuts, 2002; Santos et al., 2006). Second, more random new connections leading to more polarization is a stable finding in opinion dynamics models which include negative influence (Flache & Macy, 2006, 2011; Turner & Smaldino, 2018). Here we show that these processes are interlinked.

## 12 Short comings and future work

### 12.1 Including Empirical Data

Previous review articles have pointed to the fact that the field needs to include empirical data more in their studies of these dynamical systems (Flache et al., 2017; Noorazar et al., 2020; Sîrbu et al., 2017). Although I make the attempt to integrate and test the model with real empirical networks in this paper, we still need more data. One of the main problems with integrating the data into the models seems to be that these models require very specific data to be properly evaluated. What we really need is dynamical networks of social networks of opinions over time. Previous results have focused on measuring dynamical networks (Kossinets & Watts, 2009), but including the opinions of the agents would be critical information to have for these types of models. If this data was available, it could be possible to identify how the patterns of the networks evolve over time. Although this paper shows that co-evolutionary models can better explain empirical opinionated networks, it does so by comparing the network to a snapshot of a real-world network. In other words, the fact that the co-evolutionary model fits better does not necessarily mean that it does so by implementing the real underlying mechanisms that gives rise to the patterns in the data. By having access to dynamical networks, it would be possible to identify the realism of the proposed mechanisms for both social influence and network formation. This data would also allow researchers to test more specific findings within this paper. Specifically, it would be possible to check whether increases in tie-deletion to dissimilar people decreases polarization.

### 12.2 Accurate perceptions of opinions

An important limitation to this study is that interactions are assumed to be symmetrical and accurate. By interactions being symmetrical, I mean that agents push and pull on each other's opinions by the same factor. This is most likely not an accurate reflection of reality. Individuals of an interacting dyad will probably move more or less than its counterpart. This has been modelled previously in related models to the one proposed here as either "stubbornness", or situations where more extreme agents will be less likely to move (Flache et al., 2017; Ghaderi & Srikant, 2014; Yildiz, Ozdaglar, Acemoglu, Saberi, & Scaglione, 2013). Although I believe the inclusion of this mechanism will increase the realism of this model, the mechanism of stubbornness is not a key variable of interest to this study. Stubbornness is nonetheless worth mentioning because it could interact with some of the variables of interest to this paper. If extreme agents are more stubborn, the system could become even more path dependent than it already is (Turner & Smaldino, 2018). To see this, initial increases of polarization by chance encounters would lead to agents being more unlikely to change, which would make reaching consensus more implausible for the system. This could decrease the effect of tie-deletion on facilitating consensus, as agents would be more unlikely to overcome initial increases in polarization.

A perhaps more critical flaw in the assumptions of the underlying mechanisms of this model is the lack of noise in interactions. It is highly unlikely that humans perceive or share opinions accurately. Instead, it is more reasonable to assume that people will sometimes misunderstand or misrepresent opinions (Jussim & Osgood, 1989). Opinion

formation is likely skewed by the same cognitive heuristics and biases that shape other parts of how we include information into our decisions (Arceneaux, 2012). This includes effects such as confirmation bias, which might play a vital role in how we update our opinions (Allahverdyan & Galstyan, 2014). Typically, models include noise in the interactions between agents to emulate imperfect communication (Sîrbu et al., 2017; Su, Chen, & Hong, 2017). Noisy interactions could potentially play a vital role in facilitating cooperation, as it can allow agents that would not normally interact positively to perceive each other as more similar by chance (Allahverdyan & Galstyan, 2014; Su et al., 2017). Including more realistic cognitive mechanisms might change the dynamics of the system in important ways which are not considered in this model. One way of implementing this would be to reimagine updating one's opinion as Bayesian updating considering other's opinions as new evidence (Allahverdyan & Galstyan, 2014). This can be done by letting the opinion of an agent be described by a distribution instead of a single number. The distribution could be a normal distribution, where the mean gives the general opinion of an agent, and the standard deviation describes the certainty of the agent's opinion.

### 12.3 Shrinking Variance

Because no new agents are added to the network and because interactions are non-noisy, the variance in opinions shrinks over the course of time. At first glance, this might not seem very problematic. People tend to conform to the majority of their social groups (McPherson et al., 2001), but we wouldn't expect near perfect consensus on most issues facing us. Introducing mechanisms that keep a sufficient amount of variance would be more realistic but might also decrease the path dependency of the system. SAY SOMETHING ABOUT MODELS HAVING AGENTS TRYING TO BE DIFFERENT

### 12.4 Social influence exclusively

The model investigated here assumes that one's opinion is only changed by local interactions. Although this is in line with most models of opinion dynamics, there are good reasons to believe that global trends and tendencies are powerful forces that shapes opinion (Bener et al., 2016). These global tendencies might shape our opinions more and more, as media consumption rises. Global perturbations to the system could be powerful forces that potentially could have large impacts on whether the system can reach consensus. These forces would be especially counterproductive for reaching consensus, if they only influence certain parts of the system. The power of such effects is not hard to imagine – partisan news media is a vital piece of the puzzle of how opinions are formed (Pennycook & Rand, 2019). As media consumption rises and becomes a more and more important force for what information we receive, these force might already be a more important variable to consider than local interactions of agents (Bener et al., 2016; Strömbäck, Djerf-Pierre, & Shehata, 2013).

## 13 A broader perspective

The model investigated in this paper points to a possible mechanism for cooperation. As we have seen that mechanism is the coevolution of networks and opinions, which leads to distance in opinion space being closely linked to distance in network space. Whenever there is a possibility for negative interactions in the system, the speed of co-evolution becomes tantamount for whether the system reaches consensus or not. It is therefore interesting to see whether how connected we are and how much we disagree corresponds in empirical networks. As was pointed out in the introductory paragraph, polarization has been on the rise in many western democracies over the last two decades (Boxell et al., 2020; Center, 2014a). On the other hand, estimations of the average path length between users on Facebook has decreased from 2011 to 2016 (Bhagat, Burke, Diuk, Edunov, & Filiz, 2016). The coevolutionary model investigated in this paper suggests that the increase in connectivity and the simultaneous increase in polarization could be a bad omen. If anything like negative influence characterizes interactions on social media (i.e. outrage), the decrease in average path length will make negative encounters more likely, and most likely increase rather than decrease polarization. Although being more connected might sound good on paper, it might be detrimental to our ability to cooperate in the long run. This effect is congruent with previous literature, where the introduction of long-range ties sharply increases polarization of opinions in the population (Flache & Macy, 2011; Turner & Smaldino, 2018).

## 14 Conclusion

## References

- Akiba, T., Sano, S., Yanase, T., Ohta, T., & Koyama, M. (2019, July). Optuna: A Next-generation Hyperparameter Optimization Framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining* (pp. 2623–2631). New York, NY, USA: Association for Computing Machinery. Retrieved 2022-02-24, from <https://doi.org/10.1145/3292500.3330701> doi: 10.1145/3292500.3330701
- Allahverdyan, A. E., & Galstyan, A. (2014, July). Opinion Dynamics with Confirmation Bias. *PLOS ONE*, 9(7), e99557. Retrieved 2022-03-31, from <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0099557> (Publisher: Public Library of Science) doi: 10.1371/journal.pone.0099557
- Andris, C., Lee, D., Hamilton, M. J., Martino, M., Gunning, C. E., & Selden, J. A. (2015, April). The Rise of Partisanship and Super-Cooperators in the U.S. House of Representatives. *PLOS ONE*, 10(4), e0123507. Retrieved 2022-02-24, from <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0123507> (Publisher: Public Library of Science) doi: 10.1371/journal.pone.0123507
- Arceneaux, K. (2012). Cognitive Biases and the Strength of Political Arguments. *American Journal of Political Science*, 56(2), 271–285. Retrieved 2022-03-31, from <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1540-5907.2011.00573.x> (.eprint:

- <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-5907.2011.00573.x> doi: 10.1111/j.1540-5907.2011.00573.x
- Asikainen, A., Iñiguez, G., Ureña-Carrión, J., Kaski, K., & Kivelä, M. (2020, May). Cumulative effects of triadic closure and homophily in social networks. *Science Advances*. Retrieved 2022-01-26, from <https://www.science.org/doi/abs/10.1126/sciadv.aax7310> (Publisher: American Association for the Advancement of Science) doi: 10.1126/sciadv.aax7310
- Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., ... Volfovsky, A. (2018, September). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences*, 115(37), 9216–9221. Retrieved 2022-03-19, from <https://www.pnas.org/doi/10.1073/pnas.1804840115> (Publisher: Proceedings of the National Academy of Sciences) doi: 10.1073/pnas.1804840115
- Barabási, A.-L., & Bonabeau, E. (2003, May). Scale-free networks. *Scientific American*, 288(5), 60–69. Retrieved 2022-03-17, from <https://doi.org/10.1038/scientificamerican0503-60> doi: 10.1038/scientificamerican0503-60
- Baumann, F., Lorenz-Spreen, P., Sokolov, I. M., & Starnini, M. (2020, January). Modeling Echo Chambers and Polarization Dynamics in Social Networks. *Physical Review Letters*, 124(4), 048301. Retrieved 2022-01-26, from <https://link.aps.org/doi/10.1103/PhysRevLett.124.048301> (Publisher: American Physical Society) doi: 10.1103/PhysRevLett.124.048301
- Baumann, F. T. W. (2021). *Modeling Opinion Dynamics on Networks: How Social Influence Shapes the Formation of Consensus and Polarization* (Ph.D., Humboldt Universitaet zu Berlin (Germany), Germany). Retrieved 2022-01-26, from <https://www.proquest.com/docview/2578182483/abstract/A3837AF8EFD40DAPQ/1> (ISBN: 9798544276395)
- Bener, A. B., Çağlayan, B., Henry, A. D., & Prałat, P. (2016, October). Empirical Models of Social Learning in a Large, Evolving Network. *PLOS ONE*, 11(10), e0160307. Retrieved 2022-04-01, from <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0160307> (Publisher: Public Library of Science) doi: 10.1371/journal.pone.0160307
- Bhagat, S., Burke, M., Diuk, C., Edunov, S., & Filiz, I. O. (2016, February). *Three and a half degrees of separation - Meta Research*. Retrieved 2022-04-01, from <https://research.facebook.com/blog/2016/2/three-and-a-half-degrees-of-separation/>
- Bianconi, G., Darst, R. K., Iacoviacci, J., & Fortunato, S. (2014). Triadic closure as a basic generating mechanism of the structure of complex networks. *Physical review. E, Statistical, nonlinear, and soft matter physics*. doi: 10.1103/PhysRevE.90.042806
- Boxell, L., Gentzkow, M., & Shapiro, J. M. (2020, January). *Cross-Country Trends in Affective Polarization* (Working Paper No. 26669). National Bureau of Economic Research. Retrieved 2022-02-24, from <https://www.nber.org/papers/w26669> (Series: Working Paper Series) doi: 10.3386/w26669
- Boyd, R., Gintis, H., & Bowles, S. (2010, April). Coordinated Punishment of Defectors Sustains Cooperation and Can Proliferate When Rare. *Science*, 328(5978), 617–620. Retrieved 2022-03-30, from <https://www.science.org/doi/full/10.1126/science.1183665> (Publisher: American Association

- for the Advancement of Science) doi: 10.1126/science.1183665
- Broido, A. D., & Clauset, A. (2019, March). Scale-free networks are rare. *Nature Communications*, 10(1), 1017. Retrieved 2022-03-17, from <https://www.nature.com/articles/s41467-019-08746-5> (Number: 1 Publisher: Nature Publishing Group) doi: 10.1038/s41467-019-08746-5
- Bruch, E., & Atwell, J. (2015, May). Agent-Based Models in Empirical Social Research. *Sociological Methods & Research*, 44(2), 186–221. Retrieved 2022-04-05, from <https://doi.org/10.1177/0049124113506405> (Publisher: SAGE Publications Inc) doi: 10.1177/0049124113506405
- Center, P. R. (2014a, June). *Political Polarization and Growing Partisan Antipathy*. Retrieved 2022-03-11, from <https://www.pewresearch.org/politics/2014/06/12/section-2-growing-partisan-antipathy/>
- Center, P. R. (2014b, June). *Political Polarization in the American Public*. Retrieved 2022-03-11, from <https://www.pewresearch.org/politics/2014/06/12/political-polarization-in-the-american-public/>
- Chacoma, A., & Zanette, D. H. (2015, October). Opinion Formation by Social Influence: From Experiments to Modeling. *PLOS ONE*, 10(10), e0140406. Retrieved 2022-02-24, from <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0140406> (Publisher: Public Library of Science) doi: 10.1371/journal.pone.0140406
- Crandall, D., Cosley, D., Huttenlocher, D., Kleinberg, J., & Suri, S. (2008, August). Feedback effects between similarity and social influence in online communities. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 160–168). New York, NY, USA: Association for Computing Machinery. Retrieved 2022-04-08, from <https://doi.org/10.1145/1401890.1401914> doi: 10.1145/1401890.1401914
- Crooks, A. T., & Heppenstall, A. J. (2012). Introduction to agent-based modelling. In *Agent-based models of geographical systems* (pp. 85–105). Springer.
- Dakin, R., & Ryder, T. B. (2018, December). Dynamic network partnerships and social contagion drive cooperation. *Proceedings of the Royal Society B: Biological Sciences*, 285(1893), 20181973. Retrieved 2022-03-30, from <https://royalsocietypublishing.org/doi/full/10.1098/rspb.2018.1973> (Publisher: Royal Society) doi: 10.1098/rspb.2018.1973
- Del Vicario, M., Vivaldo, G., Bessi, A., Zollo, F., Scala, A., Caldarelli, G., & Quattrociocchi, W. (2016, December). Echo Chambers: Emotional Contagion and Group Polarization on Facebook. *Scientific Reports*, 6(1), 37825. Retrieved 2022-04-09, from <https://www.nature.com/articles/srep37825> (Number: 1 Publisher: Nature Publishing Group) doi: 10.1038/srep37825
- Dimock, M., & Wike, R. (2020, November). *America is exceptional in the nature of its political divide*. Retrieved 2022-03-11, from <https://www.pewresearch.org/fact-tank/2020/11/13/america-is-exceptional-in-the-nature-of-its-politi>
- Epstein, J. M. (1999). Agent-based computational models and generative social science. *Complexity*, 4(5), 41–60.
- Ferraz de Arruda, H., Maciel Cardoso, F., Ferraz de Arruda, G., R. Hernández, A., da Fon-

- toura Costa, L., & Moreno, Y. (2022, April). Modelling how social network algorithms can influence opinion polarization. *Information Sciences*, 588, 265–278. Retrieved 2022-03-11, from <https://www.sciencedirect.com/science/article/pii/S0020025521012901> doi: 10.1016/j.ins.2021.12.069
- Flache, A. (2018, December). Between Monoculture and Cultural Polarization: Agent-based Models of the Interplay of Social Influence and Cultural Diversity. *Journal of Archaeological Method and Theory*, 25(4), 996–1023. Retrieved 2022-01-26, from <https://doi.org/10.1007/s10816-018-9391-1> doi: 10.1007/s10816-018-9391-1
- Flache, A., & Macy, M. W. (2006, April). Why more contact may increase cultural polarization. *arXiv:physics/0604196*. Retrieved 2022-03-20, from <http://arxiv.org/abs/physics/0604196> (arXiv: physics/0604196)
- Flache, A., & Macy, M. W. (2011, January). Small Worlds and Cultural Polarization. *Journal of Mathematical Sociology*, 35(1-2), 146–176. doi: 10.1080/0022250X.2010.532261
- Flache, A., Mäs, M., Feliciani, T., Chattoe-Brown, E., Deffuant, G., Huet, S., & Lorenz, J. (2017). Models of Social Influence: Towards the Next Frontiers. *Journal of Artificial Societies and Social Simulation*, 20(4), 2.
- Fogarty, L., Ammar, M., Holding, T., Powell, A., & Kandler, A. (2022, March). Ten simple rules for principled simulation modelling. *PLOS Computational Biology*, 18(3), e1009917. Retrieved 2022-04-05, from <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1009917> (Publisher: Public Library of Science) doi: 10.1371/journal.pcbi.1009917
- Friedkin, N. E., & Johnsen, E. C. (1990, January). Social influence and opinions. *The Journal of Mathematical Sociology*, 15(3-4), 193–206. Retrieved 2022-01-26, from <https://doi.org/10.1080/0022250X.1990.9990069> (Publisher: Routledge eprint: <https://doi.org/10.1080/0022250X.1990.9990069>) doi: 10.1080/0022250X.1990.9990069
- Fuglede, B., & Topsøe, F. (2004, June). Jensen-Shannon divergence and Hilbert space embedding. In *International Symposium on Information Theory, 2004. ISIT 2004. Proceedings*. (pp. 31–). doi: 10.1109/ISIT.2004.1365067
- Galesic, M., Olsson, H., Dalege, J., van der Does, T., & Stein, D. L. (2021, March). Integrating social and cognitive aspects of belief dynamics: towards a unifying framework. *Journal of The Royal Society Interface*, 18(176), 20200857. Retrieved 2022-03-11, from <https://royalsocietypublishing.org/doi/10.1098/rsif.2020.0857> (Publisher: Royal Society) doi: 10.1098/rsif.2020.0857
- Ghaderi, J., & Srikant, R. (2014, December). Opinion dynamics in social networks with stubborn agents: Equilibrium and convergence rate. *Automatica*, 50(12), 3209–3215. Retrieved 2022-03-31, from <https://www.sciencedirect.com/science/article/pii/S0005109814004154> doi: 10.1016/j.automatica.2014.10.034
- Holme, P., & Newman, M. E. J. (2006, November). Nonequilibrium phase transition in the coevolution of networks and opinions. *Physical Review E*, 74(5), 056108. Retrieved 2022-03-17, from <https://link.aps.org/doi/10.1103/PhysRevE.74.056108> (Publisher: American Physical Society) doi: 10.1103/PhysRevE.74.056108

- Hutter, F., Hoos, H., & Leyton-Brown, K. (2014). An efficient approach for assessing hyperparameter importance. In *International conference on machine learning* (pp. 754–762).
- Ilany, A., & Akçay, E. (2016, June). Social inheritance can explain the structure of animal social networks. *Nature Communications*, 7(1), 12084. Retrieved 2022-01-26, from <https://www.nature.com/articles/ncomms12084> (Bandiera.abtest: a Cc\_license\_type: cc\_by Cg\_type: Nature Research Journals Number: 1 Primary\_atype: Research Publisher: Nature Publishing Group Subject\_term: Social evolution;Theoretical ecology Subject\_term\_id: social-evolution;theoretical-ecology) doi: 10.1038/ncomms12084
- Iyengar, S., Leikes, Y., Levendusky, M., Malhotra, N., & Westwood, S. (2019). The Origins and Consequences of Affective Polarization in the United States. *Annual Review of Political Science*. doi: 10.1146/ANNUREV-POLISCI-051117-073034
- Jackson, M., & Rogers, B. (2004, June). Search and the Strategic Formation of Large Networks: When and Why do We See Power Laws and Small Worlds?
- Jackson, M., & Rogers, B. W. (2007). Meeting Strangers and Friends of Friends: How Random are Social Networks? doi: 10.1257/AER.97.3.890
- Jalili, M. (2015, November). Coevolution of opinion formation and network dynamics in complex networked systems. In *2015 International Conference on Information Society (i-Society)* (pp. 79–83). doi: 10.1109/i-Society.2015.7366863
- Jussim, L., & Osgood, D. W. (1989). Influence and Similarity Among Friends: An Integrative Model Applied to Incarcerated Adolescents. *Social Psychology Quarterly*, 52(2), 98–112. Retrieved 2022-04-08, from <https://www.jstor.org/stable/2786910> (Publisher: [Sage Publications, Inc., American Sociological Association]) doi: 10.2307/2786910
- Kossinets, G., & Watts, D. (2009, September). Origins of Homophily in an Evolving Social Network. *American Journal of Sociology*, 115(2), 405–450. Retrieved 2022-01-29, from <https://www.journals.uchicago.edu/doi/abs/10.1086/599247> (Publisher: The University of Chicago Press) doi: 10.1086/599247
- Levin, S. A., Milner, H. V., & Perrings, C. (2021, December). The dynamics of political polarization. *Proceedings of the National Academy of Sciences*, 118(50), e2116950118. Retrieved 2022-04-08, from <https://www.pnas.org/doi/10.1073/pnas.2116950118> (Publisher: Proceedings of the National Academy of Sciences) doi: 10.1073/pnas.2116950118
- Lev-On, A., & Manin, B. (2009, October). *Happy Accidents: Deliberation and Online Exposure to Opposing Views* (SSRN Scholarly Paper No. 1481869). Rochester, NY: Social Science Research Network. Retrieved 2022-04-09, from <https://papers.ssrn.com/abstract=1481869>
- Levy, R. (2021, March). Social Media, News Consumption, and Polarization: Evidence from a Field Experiment. *American Economic Review*, 111(3), 831–870. Retrieved 2022-04-10, from <https://www.aeaweb.org/articles?id=10.1257/aer.20191777> doi: 10.1257/aer.20191777
- Mäs, M., Flache, A., & Kitts, J. A. (2014). Cultural integration and differentiation in groups and organizations. In *Perspectives on culture and agent-based simulations* (pp. 71–90). Springer.

- McCoy, J., Rahman, T., & Somer, M. (2018, January). Polarization and the Global Crisis of Democracy: Common Patterns, Dynamics, and Pernicious Consequences for Democratic Polities. *American Behavioral Scientist*, 62(1), 16–42. Retrieved 2022-04-06, from <https://doi.org/10.1177/0002764218759576> (Publisher: SAGE Publications Inc) doi: 10.1177/0002764218759576
- McPherson, M., Smith-Lovin, L., & Cook, J. M. (2001). Birds of a Feather: Homophily in Social Networks. *Annual Review of Sociology*, 27(1), 415–444. Retrieved 2022-03-14, from <https://doi.org/10.1146/annurev.soc.27.1.415> (eprint: <https://doi.org/10.1146/annurev.soc.27.1.415>) doi: 10.1146/annurev.soc.27.1.415
- Melamed, D., & Simpson, B. (2016, March). Strong ties promote the evolution of cooperation in dynamic networks. *Social Networks*, 45, 32–44. Retrieved 2022-03-31, from <https://www.sciencedirect.com/science/article/pii/S0378873315300605> doi: 10.1016/j.socnet.2015.11.001
- Noorazar, H., Vixie, K. R., Talebanpour, A., & Hu, Y. (2020, July). From classical to modern opinion dynamics. *International Journal of Modern Physics C*, 31(07), 2050101. Retrieved 2022-02-24, from <https://www.worldscientific.com/doi/abs/10.1142/S0129183120501016> (Publisher: World Scientific Publishing Co.) doi: 10.1142/S0129183120501016
- Page, S. (2010). Diversity and complexity. In *Diversity and complexity*. Princeton University Press.
- Peixoto, T. P. (2022, January). Disentangling homophily, community structure and triadic closure in networks. *Physical Review X*, 12(1), 011004. Retrieved 2022-01-26, from <http://arxiv.org/abs/2101.02510> (arXiv: 2101.02510) doi: 10.1103/PhysRevX.12.011004
- Pennycook, G., & Rand, D. G. (2019, July). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, 188, 39–50. Retrieved 2022-04-01, from <https://www.sciencedirect.com/science/article/pii/S001002771830163X> doi: 10.1016/j.cognition.2018.06.011
- Pepper, J. W., & Smuts, B. B. (2002, August). A Mechanism for the Evolution of Altruism among Nonkin: Positive Assortment through Environmental Feedback. *The American Naturalist*, 160(2), 205–213. Retrieved 2022-03-30, from <https://www.journals.uchicago.edu/doi/full/10.1086/341018> (Publisher: The University of Chicago Press) doi: 10.1086/341018
- Santos, F. C., Pacheco, J. M., & Lenaerts, T. (2006, October). Cooperation Prevails When Individuals Adjust Their Social Ties. *PLOS Computational Biology*, 2(10), e140. Retrieved 2022-02-24, from <https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.0020140> (Publisher: Public Library of Science) doi: 10.1371/journal.pcbi.0020140
- Sasahara, K., Chen, W., Peng, H., Ciampaglia, G. L., Flammini, A., & Menczer, F. (2021, May). Social influence and unfollowing accelerate the emergence of echo chambers. *Journal of Computational Social Science*, 4(1), 381–402. Retrieved 2022-01-26, from <https://doi.org/10.1007/s42001-020-00084-7> doi: 10.1007/s42001-020-00084-7
- Sîrbu, A., Loreto, V., Servedio, V. D., & Tria, F. (2017). Opinion dynamics: models, extensions and external effects. In *Participatory sensing, opinions and collective awareness* (pp. 363–401). Springer.

- Smaldino, P. (2016, January). Models Are Stupid, and We Need More of Them.. doi: 10.4324/9781315173726-14
- Smaldino, P. E. (2020, July). How to Translate a Verbal Theory Into a Formal Model. *Social Psychology*, 51(4), 207–218. Retrieved 2022-04-06, from <https://econtent.hogrefe.com/doi/10.1027/1864-9335/a000425> (Publisher: Hogrefe Publishing) doi: 10.1027/1864-9335/a000425
- Somer, M., & McCoy, J. (2018, January). Déjà vu? Polarization and Endangered Democracies in the 21st Century. *American Behavioral Scientist*, 62(1), 3–15. Retrieved 2022-04-06, from <https://doi.org/10.1177/0002764218760371> (Publisher: SAGE Publications Inc) doi: 10.1177/0002764218760371
- Spears, R. (2021). Social Influence and Group Identity. *Annual Review of Psychology*, 72(1), 367–390. Retrieved 2022-01-26, from <https://doi.org/10.1146/annurev-psych-070620-111818> (.eprint: <https://doi.org/10.1146/annurev-psych-070620-111818>) doi: 10.1146/annurev-psych-070620-111818
- Strömbäck, J., Djerf-Pierre, M., & Shehata, A. (2013, December). The Dynamics of Political Interest and News Media Consumption: A Longitudinal Perspective. *International Journal of Public Opinion Research*, 25(4), 414–435. Retrieved 2022-04-01, from <https://doi.org/10.1093/ijpor/eds018> doi: 10.1093/ijpor/eds018
- Su, W., Chen, G., & Hong, Y. (2017, November). Noise leads to quasi-consensus of Hegselmann-Krause opinion dynamics. *Automatica*, 85, 448–454. Retrieved 2022-03-31, from <https://www.sciencedirect.com/science/article/pii/S0005109817304296> doi: 10.1016/j.automatica.2017.08.008
- Takács, K., Flache, A., & Maes, M. (2014, June). *Is There Negative Social Influence? Disentangling Effects of Dissimilarity and Disliking on Opinion Shifts* (SSRN Scholarly Paper No. ID 2445649). Rochester, NY: Social Science Research Network. Retrieved 2022-02-24, from <https://papers.ssrn.com/abstract=2445649> doi: 10.2139/ssrn.2445649
- Taylor, C. E., Mantzaris, A. V., & Garibay, I. (2018, December). Exploring How Homophily and Accessibility Can Facilitate Polarization in Social Networks. *Information*, 9(12), 325. Retrieved 2022-01-26, from <https://www.mdpi.com/2078-2489/9/12/325> (Number: 12 Publisher: Multidisciplinary Digital Publishing Institute) doi: 10.3390/info9120325
- Tsai, W.-H. S., Tao, W., Chuan, C.-H., & Hong, C. (2020, March). Echo chambers and social mediators in public advocacy issue networks. *Public Relations Review*, 46(1), 101882. Retrieved 2022-03-29, from <https://www.sciencedirect.com/science/article/pii/S036381120300035> doi: 10.1016/j.pubrev.2020.101882
- Turner, M. A., & Smaldino, P. E. (2018, November). Paths to Polarization: How Extreme Views, Miscommunication, and Random Chance Drive Opinion Dynamics. *Complexity*, 2018, e2740959. Retrieved 2022-03-10, from <https://www.hindawi.com/journals/complexity/2018/2740959/> (Publisher: Hindawi) doi: 10.1155/2018/2740959
- Watts, D. J. (1999, September). Networks, Dynamics, and the Small-World Phenomenon. *American Journal of Sociology*, 105(2), 493–527. Retrieved 2022-03-17, from <https://www.journals.uchicago.edu/doi/abs/10.1086/210318> (Publisher: The University of Chicago Press) doi: 10.1086/210318

- Watts, D. J., & Strogatz, S. H. (1998, June). Collective dynamics of ‘small-world’ networks. *Nature*, 393(6684), 440–442. Retrieved 2022-02-02, from <https://www.nature.com/articles/30918>. (Number: 6684 Publisher: Nature Publishing Group) doi: 10.1038/30918
- Wilson, A. E., Parker, V. A., & Feinberg, M. (2020, August). Polarization in the contemporary political and media landscape. *Current Opinion in Behavioral Sciences*, 34, 223–228. Retrieved 2022-04-06, from <https://www.sciencedirect.com/science/article/pii/S2352154620301078> doi: 10.1016/j.cobeha.2020.07.005
- Winter, F., & Kataria, M. (2020, May). You are who your friends are?: An experiment on homophily in trustworthiness among friends. *Rationality and Society*, 32(2), 223–251. Retrieved 2022-01-26, from <https://doi.org/10.1177/1043463120919380> (Publisher: SAGE Publications Ltd) doi: 10.1177/1043463120919380
- Wu, J., Chen, X.-Y., Zhang, H., Xiong, L.-D., Lei, H., & Deng, S.-H. (2019, March). Hyperparameter Optimization for Machine Learning Models Based on Bayesian Optimizationb. *Journal of Electronic Science and Technology*, 17(1), 26–40. Retrieved 2022-04-04, from <https://www.sciencedirect.com/science/article/pii/S1674862X19300047> doi: 10.11989/JEST.1674-862X.80904120
- Yildiz, E., Ozdaglar, A., Acemoglu, D., Saberi, A., & Scaglione, A. (2013, December). Binary Opinion Dynamics with Stubborn Agents. *ACM Transactions on Economics and Computation*, 1(4), 19:1–19:30. Retrieved 2022-03-31, from <https://doi.org/10.1145/2538508> doi: 10.1145/2538508