Research

Authors for correspondence:
Tucker Evans
e-mail: tucker.e.evans.19@dartmouth.edu
Feng Fu
e-mail: fufeng@gmail.com

Opinion formation on dynamic networks: identifying conditions for the emergence of partisan echo chambers

Tucker Evans[1] and Feng Fu[1,2]

[1]Department of Mathematics, Dartmouth College, 27 North Main Street, 6188 Kemeny Hall, Hanover, NH 03755, USA
[2]Department of Biomedical Data Science, Geisel School of Medicine at Dartmouth, Lebanon, NH 03756, USA

TE, 0000-0001-7191-7252; FF, 0000-0001-8252-1990

Modern political interaction is characterized by strong partisanship and a lack of interest in information sharing and agreement across party lines. It remains largely unclear how such partisan echo chambers arise and how they coevolve with opinion formation. Here, we explore the emergence of these structures through the lens of coevolutionary games. In our model, the payoff of an individual is determined jointly by the magnitude of their opinion, their degree of conformity with their social neighbours and the benefit of having social connections. Each individual can simultaneously adjust their opinion and the weights of their social connections. We present and validate the conditions for the emergence of partisan echo chambers, characterizing the transition from cohesive communities with a consensus to divisive networks with splitting opinions. Moreover, we apply our model to voting records of the US House of Representatives over a timespan of decades to understand the influence of underlying psychological and social factors on increasing partisanship in recent years. Our work helps elucidate how the division of today has come to be and how cohesion and unity could otherwise be attained on a variety of political and social issues.

1. Introduction

In a political context of growing partisanship and echo chamber formation, a natural question is how such a culture of division arose and what might be to come in the future if the trend continues. To address this question using an opinion formation

model, it is necessary to take into account the underlying motives of an individual's actions and also his or her position in the socio-political milieu. This environment consists of all the influences on that individual: his or her sources of information and the social pressures that he or she feels to conform to one view or another. Network science offers the theoretical basis for the quantification and description of these factors and provides metrics for the modular structure implied by 'echo chambers' and partisan politics.

Over the past decades, making use of the extensive available data, researchers have been able to identify some common topological characteristics of social networks. For example, social networks are frequently found to be highly heterogeneous with the small-world property [1,2]. In the broader biological context, high-degree nodes have been shown to play an essential role in promoting network robustness and increasing the efficiency of communication and flow through the network [1,3]. Furthermore, biological networks frequently exhibit organization into clustered communities of nodes (often referred to as modules), the recognition of which is in itself a topic of study [4–6]. This line of research is in parallel with more theoretical considerations of how large networks can be categorized and described [7–12].

The discovery of these network topological features begs the question of how they influence the dynamics taking place on the network. Concerning opinion dynamics specifically, how does the network structure of a group of individuals change the way in which those individuals develop their own beliefs? Previous studies of opinion dynamics have explored population structures with the general characteristics of social networks with an emphasis on refining the opinion updating procedures employed by individual nodes in the network [13–22].

Most recently, the coevolution of opinion dynamics and individual social connections has received increasing attention [23–26]. In these models, individuals modify their interactions with those who chronically disagree with them, usually with the purpose of avoiding discussion of the topics which arouse contention. Building on this prior work, we investigate opinion formation on dynamic social network through the lens of a coevolutionary game [27]. The relevant structural information of the network includes the opinions of the individuals in the network and the relative influence that each individual exercises on his or her neighbours, concepts which can be understood as values attached to the nodes and to the edges, respectively. This paper will consider how individuals' opinions change alongside the structure of the network in which they are embedded, taking as a basic assumption that the two are fundamentally linked and that they mutually influence each others' evolution. Development of this kind can lead to complex structure within the network, an excellent example of which are the idealistically divided sub-communities that develop in the political sphere, from democrats and republicans [28,29] to pro-vaccination and anti-vaccination groups [30].

Improving our understanding of opinion evolution is essential in the context of modern politics. Partisanship has come to dominate the political sphere, among the political elite and the general population [28,29,31–33], stalling political consensus on issues that demand rapid decision-making. It has become a major research concern to effectively understand circumstances that will lead to convergence of opinion and others that will lead to divergence of opinion and a weakening of information transfer [34]. The model and methods provided in this work allow for the analysis of the current state of politics and the partial elucidation of how the division of today has come to be.

2. Model

Let us denote by $\mathbf{X}$ the vector of opinions of all nodes in the network, where X_i refers to the belief of the ith node. The values of the X_i are allowed to range in the interval $[-1, 1]$, where negative values indicate disagreement and positive values denote agreement with a given proposition. Greater magnitudes indicate stronger opinions. Let $\mathbf{C}$ represent the weighted adjacency matrix of the network, composed of the weights on the connections between any pair of two nodes in the network, with C_{ij} referring to the connection between the ith node and the jth node. For the purpose of our model, the connections between nodes are considered to be symmetrical, meaning that $C_{ij} = C_{ji}$. The value of C_{ij} can range between zero and one, where zero indicates that there is no connection between i and j and one indicates the strongest possible tie between the two.

The payoff (i.e. fitness) of a node in the network is determined as the sum of three different factors: the strength of the node's opinion, the node's level of agreement with neighbouring nodes in the network (nodes with which the node in question has a non-zero connection) and the degree of the node (the total sum of the weights of the node's connections). The importance of each of these three components is described by the parameter α, β and γ, respectively.

For a network of size N, the fitness of an individual node i is given by

$$f_i(X, C) = \alpha|X_i| + \beta \sum_{i \neq j} C_{i,j} X_i X_j + \gamma \sum_{i \neq j} C_{i,j}. \tag{2.1}$$

The first and third terms of the fitness function are the benefits for a greater magnitude of opinion and more connection with neighbouring nodes, respectively, and their forms are self-explanatory. For a given node, the second term sums over all other nodes in the network, determining the cognitive dissonance accrued by interaction with each of them [35,36]. Note that where X_i and X_j are of opposite sign, meaning that the nodes i and j disagree, their contribution to the second term of the fitness function will be negative. Thus disagreeing nodes incur a penalty. The penalty is scaled by the strength of the connection between the two nodes, as the effects of cognitive dissonance will be most pronounced when there is a strong influence between individuals, thus making contradictory opinions harder to ignore.

The purpose of our model is to understand the behaviour of an opinion network given a set of initial conditions including the starting connections between nodes of the network, the initial beliefs of the nodes in the network and the parameters of the fitness function. The coevolution of network connections and opinions is achieved through an iterative process of mutation (perturbation) that amounts to a simplified genetic algorithm: changes are made to both connection and opinion values in the network during each iteration, with predetermined step-sizes. At each step, only those changes that improve the fitness of the network are kept (see details in the Methods section). Over many iterations, these random mutations improve the network, leading ultimately to a resting stable state.

3. Results

Let us start with qualitatively understanding the long-term coevolutionary dynamics of network connections and individual opinions. We make the assumption that the nodes' initial beliefs are distributed randomly over the full range of beliefs and that there exists no initial bias on the connections between nodes (nodes of similar opinions are not already clustered for instance). As shown in equation (2.1), the fitness function of an individual i can be improved unilaterally by increasing the magnitude of its opinion X_i, regardless of their network connections. This suggests that individuals tend to adopt extreme opinions, either $+1$ or -1 in the long term, and the direction will be determined by the initial configurations within their local neighbourhoods.

Furthermore, taking the partial derivatives of f_i with respect to C_{ij}, we obtain

$$\frac{\partial f_i(X, C)}{\partial C_{i,j}} = \beta X_i X_j + \gamma. \tag{3.1}$$

We find that node i is likely to increase the strength of its connection to node j, if $X_i X_j > -\gamma/\beta$. Similarly, we find that node i tends to increase its magnitude of opinion, if $\sum X_j C_{i,j} > -\text{sign}(X_i)\alpha/\beta$. In particular, for $\beta > \gamma$, nodes with opposite views are much more likely to become disconnected, and hence the resulting network tends to divide into two echo chambers with splitting views. To illustrate these effects, we turn to computer simulations to study how the long-term behaviour of the network depends on the parameters β and γ (see Methods).

Figure 1 demonstrates that the parameter choices of β and γ mainly drive the coevolutionary dynamics of opinion and networks. Figure 1a,b reflects what can happen when β is dominant rather than γ. There is an initial stage of divergence, but in this case it reduces over time. As the value of maintaining connections is lower than it was for the convergent network, the relationships between nodes that disagree quickly deteriorate, leaving entirely separate populations. In this final state, the nodes in the network have no cognitive dissonance. Their opinions are entirely aligned with those of the other nodes present in their local community. We can relate this scenario to the real-world phenomena of political echo chambers. People with similar opinions group together to the extent that they never hear opinions differing from their own. These results complement other research into decision-making and economic games on evolving networks [37,38].

Figure 1c,d shows that cohesive communities with a consensus of opinion can be obtained for fixed α and $\beta < \gamma$. In this case, there is often an initial phase in which opinions become more extreme on both sides of the argument, but then are turned around such that the network converges to consensus (figure 1d). This apparent change in opinion is a result of the coevolving network structure, which at the turn around point has created sufficient connections between individuals that do not agree, such that there is greater benefit to moving toward the same opinion than otherwise (figure 1c). With γ higher,
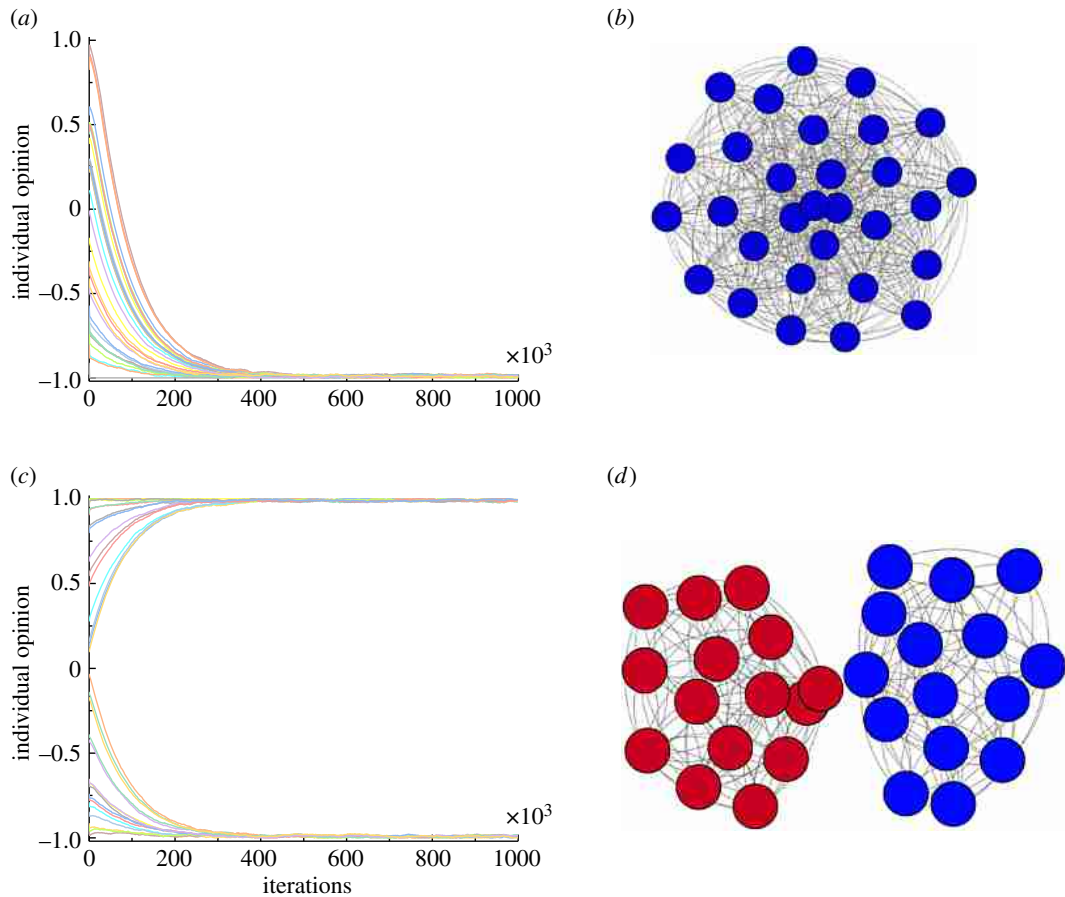
Figure 1. Coevolution of opinions and networks, where individuals can adjust their opinions and their network connections. Panels (*a*) and (*b*) show an example of cohesive community with consensus in opinion (Parameters: $\alpha = 0.5$, $\beta = 0.5$, $\gamma = 0.7$). Panels (*c*) and (*d*) show the formation of echo chambers (Parameters: $\alpha = 0.5$, $\beta = 0.05$, $\gamma = 0.7$).

however, the likelihood of connection overwhelming initial extremism is higher, as early on, more connections are formed. We might expect such an arrangement in the real world in the case of family members; here, it is difficult to rapidly change the level of communication with other individuals on a given topic, but there is still benefit to maintaining cognitive consistency. We would expect this to promote the ultimate convergence of opinions in the local network at least.

To help gain a complete picture of these results, we perform comprehensive simulations across the parameter space (γ, β), as shown in figure 2. This illustrates some of the intuitions that might be gained from the analytical solutions for the behaviour of a single node. Where β exceeds γ, it is generally expected that there will be division for example, and this is indeed what we see in the plot (figure 2*a,b*). Where β is small, it is expected that there will be division of opinion, but with a reasonable level of connections in the network. These are the cases located near the bottom of the plot (figure 2*a*). Interestingly, the divisions between these critical regions are not perfect; there are sprays of random model runs on both sides of the various divides in the plot which defy the expectation given the parameter values. Over many runs, we use these values to form a plot of the variance across the parameter space as in figure 2*b*.

As revealed by our sensitivity analysis (figure 2*a,b*), besides entirely split networks and fully connected networks (figure 1), it is possible to have sub-community formation with multiple levels of inter-group connections (figure 2*c,d*). Indeed, this is what we find in the case of the US House of Representatives voting records [39]: there is rarely complete consensus or complete division.

We can understand these simulation results intuitively as follows. The α and β terms tend to drive the network opinions apart and the connections between them to zero, whereas the γ term tends to pull the network together. The specific parameter choices determine the balance between these two forces, and therefore can lead to a variety of typical behaviours of the coevolutionary dynamics of opinions and networks (figures 1 and 2). It should be noted that the influence of the initial structure of the network on its final state may be drastic. There exist cases in which the initial structure of the network, due to the random initialization used, biases the behaviour of the network to such a great extent that it
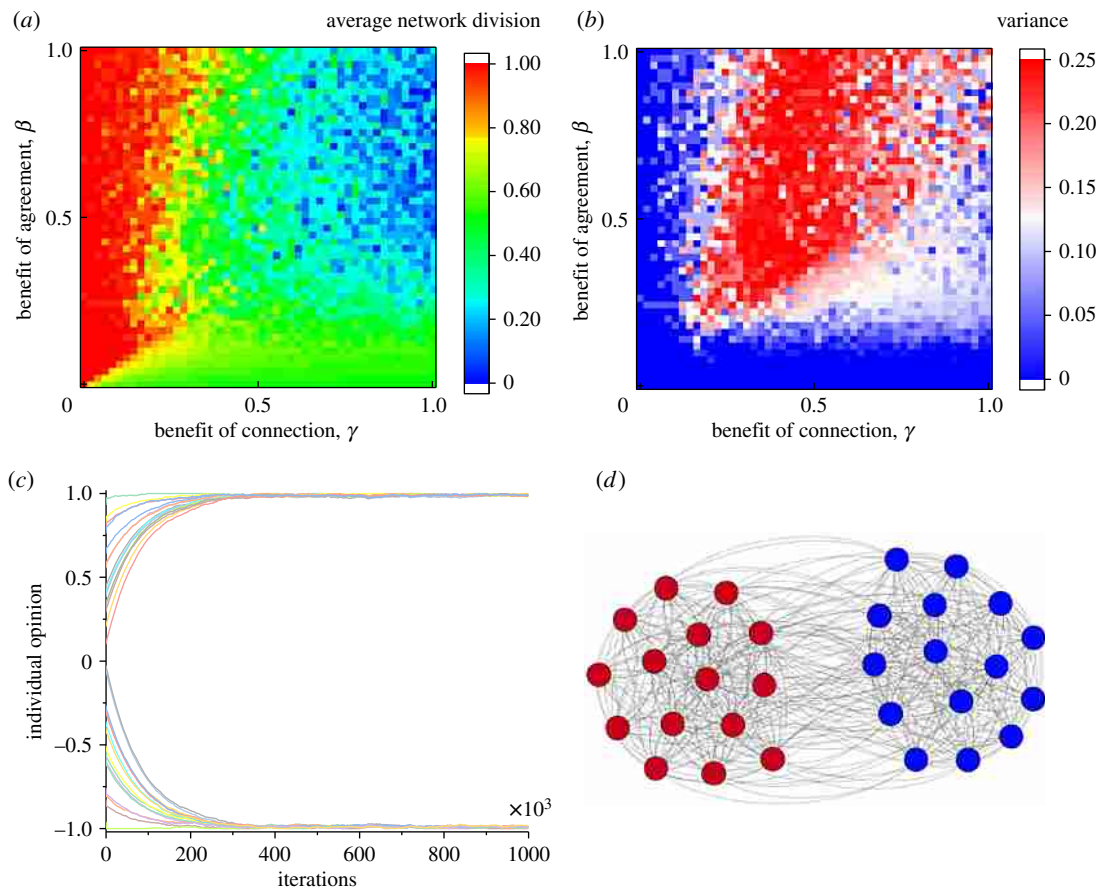
Figure 2. Conditions for the emergence of echo chambers (parameters: $\alpha = 0.5$, $\beta = 0.7$, $\gamma = 0.05$).

reverses what would have happened in a graph without any such bias. In particular, these might be cases where by chance a random division in the connections of the network is aligned with the divisions in the opinions of the nodes. This might lead to division in the network regardless of the ratio between the parameters β and γ. Alternatively, there may be a case in which there is a strong initial connection between nodes of different opinions, promoting the development of consensus where there might normally have been division.

There may also be initial biases, arising from the random initialization by chance, in the distributions of the opinions of the nodes. If the belief values chosen are particularly skewed to one side, even where the values are within their prescribed positive or negative ranges, it may be the case that the nodes will reach consensus when the parameters would normally have induced division. Similarly, a distribution of beliefs that contain only extreme opinions to begin with may stay that way even if the parameters might have predicted consensus. It is notable that the model does not allow people to hold on to weak opinion values for very long. Even for weak α values, once groups are somewhat separated, the β term leads to a mutual reinforcement of opinion between nodes. The greater the magnitudes of the beliefs of the nodes in question, the larger the β term can become, benefiting both of the nodes in question.

As a real-world application of our model, we use voting records of the United States House of Representatives over a timespan of decades [39], to understand the influence of underlying psychological and social factors on increasing partisanship in recent years. This dataset has been previously considered as an excellent example for the measurement of network division over time [39]. House members were considered to have a positive belief level if they identified with the democrats and negative belief level if they were republican. While party affiliation clearly encompasses a broad range of beliefs, we take it as a proxy for belief on a particular topic.

Figure 3 shows the calculated division values for different trial years, which span over the last five decades (each year indicating a session of the House of Representatives), along with the inferred model parameter ratio β/γ. Figure 4 shows network snapshots over time that are generated by our model with best-fitted parameter values. We found evidence to support that the relative emphases of opinion conformity (represented by β) and general connection (represented by γ) have been changing over time.
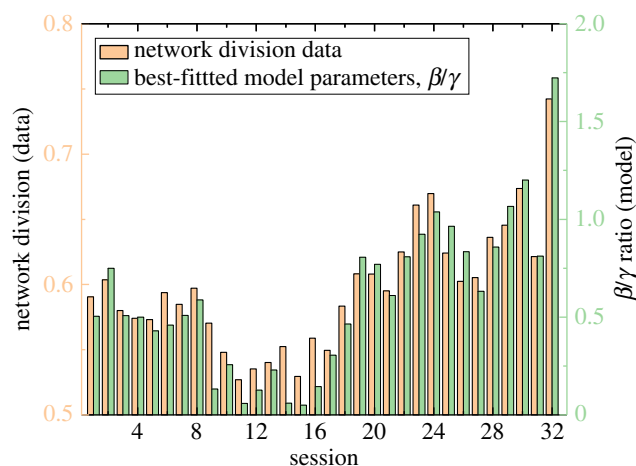
Figure 3. Inferred model parameters for given network division. Parameters: $\alpha = 0.5$.
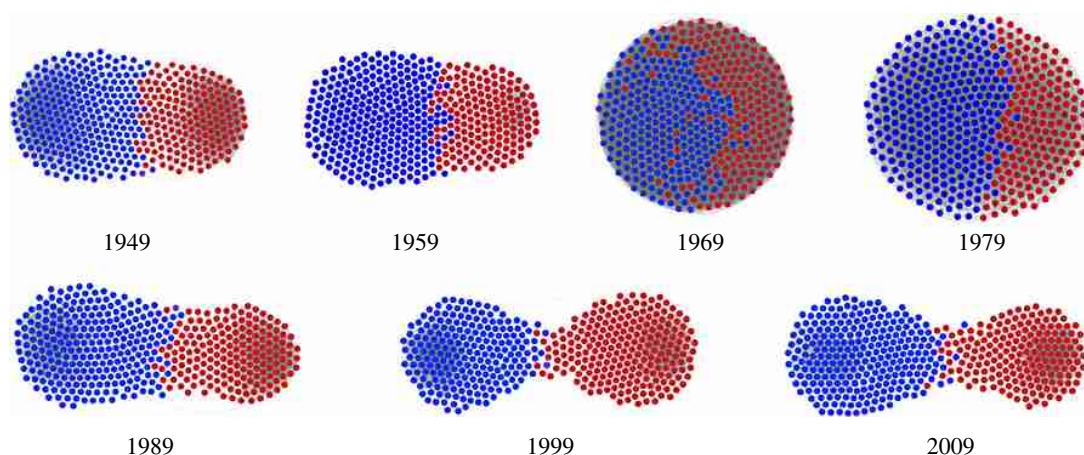

1949   1959   1969   1979

1989   1999   2009

Figure 4. Snapshots of House of Representatives networks generated using our model with inferred model parameter values β and γ ($\alpha = 0.5$).

Over this time period, we can see that both network division values and the ratio of β and γ have first experienced decreases followed by dramatic increases. In particular, recent years see a clear growth in the level of division, which along with a commensurate growth in β/γ, suggests a shift in priority from connection to congruity of opinion. These results demonstrate that our model and specifically the parameters of the fitness function can be used as a way to understand changes in opinion networks over time. Greater network division can be modelled as a growth in the value of local homogeneity of opinion when compared to the general value placed on connections with other nodes.

4. Discussion and conclusion

Our model provides a method for the simulation of changing opinions on a network, allowing for the coevolution of opinions and the connections between nodes. We have shown that it is possible to influence the final shape of a network through the modification of a fitness function based on the minimization of cognitive dissonance. Example runs of the model have demonstrated that the initial distribution of the nodes in the network and of their opinions can tip the balance of cases where the parameters of the fitness function do not strongly promote division or cohesion in the final network, but that the behaviour of the networks can be connected to those parameter values in intuitive ways.

The current model allows for the simultaneous evolution of both the connections between the nodes and their individual opinions. The purpose of the evolution is to maximize the fitness of the network, where the definition of fitness includes the effects of cognitive dissonance, node degree and opinion magnitude. This is in line with other models that have considered the interplay between individual forces acting on a given

node and the combined forces of other nodes in the network [40]. The resulting formulation is reminiscent of a spin-glass model, but is not equivalent (see Castellano *et al*. for a review of spin-glass models and similar constructions [41]). Cognitive dissonance is defined as the discomfort felt by an individual who holds two competing opinions at the same time. The discomfort it produces prompts individuals to avoid cases where they are confronted with ideas that do not match their own [22,35,36,42]. This brings out the primary assumption of this paper: people do not form their opinions based on evidence, but rather choose the opinions that they hold and the individuals with whom they surround themselves for the purpose of maximizing fitness as we have defined it.

An application of our model to partisanship in the US House of Representatives suggests that it is possible to estimate the fitness function parameters for a real-world network and that the results are logically connected to the meaning of the different parameters. The case study points to a growth in the importance of homophily at the expense of having fewer connections in American politics. The result is greater partisanship and polarization, and has been widely documented [28]. The fitness function forces changes in the network that ultimately lead to clustering of those nodes who share the same opinions and a division of the network into mutually disaffected populations. This structure mimics that of 'echo chambers' in the media and in politics, and is self-reinforcing. If this trend continues, we are likely to see only greater political division in the future rather than a return toward greater cohesion. While some variety is necessary for effective problem-solving, it is concerning to think that politics may become so divided that no compromises can be reached on even the most important questions. Datasets pertaining to changing social networks and the flow of information through them have become significantly easier to access in recent decades with the advent of the Internet and large-scale social networking sites [4,43,44]. The modelling of Twitter or Facebook networks as they change in time can be, for instance, another natural application of this research.

The importance of the underlying substrate topology has been demonstrated clearly for the Prisoner's Dilemma and similar games [45–47]. We find that the same is true of opinion dynamics. Initial conditions can heavily influence the final behaviour of a given network. One of the fundamental assumptions of this paper was that individuals choose their opinions and their connections only to avoid disagreement with their neighbours in the social network. This is indeed enough to create structure in the network similar to the structures that we see played out in real-world social networks, but it does also beg the question of where the validity of a particular opinion actually comes into play. When does the evidence start to matter in the formation of opinions? Further research should consider the interplay between logical evidence and the influence of other individuals through the medium of cognitive dissonance. A model able to combine these two components effectively would be much more complete and more accurate in the prediction of future states of a given network of opinion. Future research will also consider the application of multiple levels of opinion, delving into the world of belief 'systems' in contrast to single beliefs [21].

5. Methods

5.1. Simulation procedure

Summing over all of the nodes, we have $f(\mathbf{X}, \mathbf{C}) = \sum_{i=1}^{N} f_i$ as the total fitness function for the entire network. Expanding this quantity we have:

$$f(\mathbf{X}, \mathbf{C}) = \sum_{i=1}^{N} \left(\alpha |X_i| + \beta \sum_{i \neq j} C_{i,j} X_i X_j + \gamma \sum_{i \neq j} C_{i,j} \right). \quad (5.1)$$

The quantity $f(\mathbf{X}, \mathbf{C})$ is used as a measure of the progress of the network as a whole toward the steady-state solution. Any possible improvements to $f(\mathbf{X}, \mathbf{C})$ eventually become more and more rare as the network approaches a steady state, and so the rate of actual change of f with respect to each iteration of the model decreases.

Specifically, the update step of the model consists of two rounds of random changes to the network, first to the connections and then to the beliefs of the nodes. The number of connections that are changed and the number of beliefs that are changed can be manipulated. The case for which only one node is changed at a time is similar to a Moran process [48], and a change of all of the nodes during each step would be akin to a Wright–Fisher update process [49]. The changes in this model are not made with either of these update processes explicitly, however. Each round includes a validation stage

which ensures that the changes made have improved the fitness of the nodes in the network. Where there is no change or negative change to the fitness of a node, the corresponding mutation is removed and the state of that node variable is reset to its value at the end of the last iteration. The maximum value of a mutation can also be manipulated, but remains fixed throughout the course of the model run. This is particularly useful for the exploration of edge cases where mutations of either the connections or the beliefs of the nodes approach zero, at which point the connections or beliefs are fixed.

We take $X_i(t)$ to denote the belief of the ith node at time t and $X'_i(t)$ as the state after mutation in time step t. By the same token, C_{ij} and C'_{ij} are the values of the connection between the ith and jth nodes at time t and the their mutated values. Let ΔX_i indicate the mutation in the belief of the ith node while ΔC_{ij} indicates the mutation of the connection between the ith and jth nodes. The value of this mutation is pulled from a uniform distribution in the range $[-0.001, 0.001]$. We then have the following as the update rule for each of the nodes in the network:

$$X'_i(t) = X_i(t) + \Delta X_i \tag{5.2}$$

and

$$X_{t+1} := \begin{cases} X'_i(t) & f_i(X'(t), C(t)) > f_i(X(t), C(t)) \\ X_i(t) & f_i(X'(t), C(t)) \leq f_i(X(t), C(t)). \end{cases} \tag{5.3}$$

Similarly for the connections between nodes, we have the following update rule (with mutations in the range $[-0.0001, 0.0001]$):

$$C'_{ij}(t) = C_{ij}(t) + \Delta C_{ij} \tag{5.4}$$

and

$$C_{ij}(t+1) := \begin{cases} C'_{ij}(t) & f_i(X(t), C'(t)) > f_i(X(t), C(t)) \\ C_{ij}(t) & f_i(X(t), C'(t)) \leq f_i(X(t), C(t)). \end{cases} \tag{5.5}$$

In Matlab, the above update rules are enacted as matrix operations, with a vector of mutations being added to $\mathbf{X}$ and then validated, while a matrix of changes is added to $\mathbf{C}$ before validation.

Eventually, this process brings the network to a steady-state solution where neither the opinions of the nodes nor the connections between them can be changed in a way that is beneficial to the nodes of the network.

5.2. Network division

To compare the final states of the networks produced by the model, we introduce a simple metric for network division, which we define as the degree to which two groups are disconnected in a network. The two groups in the case of opinion dynamics will be those who agree and those who disagree. The sums of the weights on connections between nodes of the same type (denoted $\mathbf{I}$) and the weight on connections between nodes of different types (denoted $\mathbf{O}$) are calculated. The value of network division is then calculated to be $\mathbf{I}/(\mathbf{I} + \mathbf{O})$. For convenience, we define networks where all nodes converge to the same belief value as having network division values of 0. Where there are only connections in the network between nodes of the same type, the value of network division will be 1. Partially split networks will have values less than 1, but greater than 0 where there are any in-group connections at all. Network division will be equal to 0 in the case of a bipartite graph, where there are only intergroup connections.

5.3. Model-data integration

In applying this model to data concerning voting in the US House of Representatives, we calculate the division value using the average number of times that individuals voted together. The in-group connections for democrats becomes the number of democrats for a given year multiplied by the average number of democrat–democrat voting pairs, with a similar construction applied to the republican members. The total in-group connections $\mathbf{I}$ is defined as the sum of these two products. The out-group connection $\mathbf{O}$ is defined as the average number of democrat–republican pairs multiplied with the total number of members. A network division value for each session of the House of Representatives can then be calculated using the formula described above.

As a simplifying assumption, we take the beliefs of the individuals of Congress to be steady, and fix the value of α at 0.5. Theoretically, the parameter α should have no effect in this scenario, as taking the partial derivative of the fitness function with respect to the connections between nodes entirely removes the first term.

We wished to find the β/γ values that would produce these same values of network division in a similar network. We begin with a population of nodes that had the same ratio of 'democrats' to 'republicans', meaning nodes with positive or negative belief values, as was found in a given session of the House of Representatives. We then ran the model on this network for 40 different ratios of β/γ for each session. Performing a linear regression on the resultant set of data points, we gain a hypothesized relationship between the ratio β/γ and network division for that particular arrangement of nodes. We can then use our hypothesized function to find what value of β/γ would produce the network division observed. This same process is applied to all of the sessions for which there are data.

References

1. Albert R, Jeong H, Barabasi A-L. 2000 Error and attack tolerance of complex networks. *Nature* **406**, 378–382. (doi:10.1038/35019019)

2. Milgram S. 1967 The small world problem. *Psychol. Today* **2**, 60–67. (doi:10.1037/e400002009-005)

3. Borgatti SP. 2003 The key player problem. In *Dynamic social network modeling and analysis: workshop summary and papers* (eds R Breiger, K Carley, P Pattison), pp. 241–252. Washington, DC: The National Academies Press.

4. Newman MEJ. 2003 The structure and function of complex networks. *SIAM Rev.* **45**, 167–256. (doi:10.1137/S003614450342480)

5. Newman MEJ. 2006 Modularity and community structure in networks. *Proc. Natl Acad. Sci. USA* **103**, 8577–8582. (doi:10.1073/pnas.0601602103)

6. Mäs M, Flache A, Helbing D. 2010 Individualization as driving force of clustering phenomena in humans. *Comput. Biol.* **6**, e1000959. (doi:10.1371/journal.pcbi.1000959)

7. Zadorozhnyi V, Yudin E. 2012 Structural properties of the scale-free Barabasi-Albert graph. *Autom. Control* **73**, 702–716. (doi:10.1134/s0005117912040091)

8. Jácome S, da Silva L, Moreira A, Andrade Jr J, Herrmann H. 2010 Iterative decomposition of Barabasi–Albert scale-free networks. *Physica A* **389**, 3674–3677. (doi:10.1016/j.physa.2010.03.052)

9. Dorogovtsev SN, Goltsev AV, Mendes JFF. 2008 Critical phenomena in complex networks. *Rev. Mod. Phys.* **80**, 1275–1335. (doi:10.1103/RevModPhys.80.1275)

10. Dorogovtsev SN, Mendes JFF. 2002 Evolution of networks. *Adv. Phys.* **51**, 1079–1187. (doi:10.1080/00018730110112519)

11. Erdös P, Renyi A. 1959 On random graphs. *Publicationes Mathimaticae* **6**, 290–297.

12. Erdös P, Renyi A. 1960 On the evolution of random graphs. *Mag. Tud. Akad. Mat. Kutató Int. Közl.* **5**, 17–61.

13. Sznajd-Weron K, Sznajd J. 2000 Opinion evolution in closed community. *Int. J. Mod. Phys. C* **11**, 1157–1165. (doi:10.1142/S0129183100000936)

14. Deffuant G, Amblard F, Weisbuch G, Faure T. 2002 How can extremism prevail? A study based on the relative agreement interaction model. *J. Artif. Soc. Soc. Simul.* **5**. (http://jasss.soc.surrey.ac.uk/5/4/1.html)

15. Sznajd-Weron K. 2005 Sznajd model and its applications. (http://arxiv.org/abs/physics/0503239)

16. Watts DJ, Dodds PS. 2007 Influentials, networks, and public opinion formation. *J. Consumer Res.* **34**, 441–458. (doi:10.1086/518527)

17. Szabó G, Szolnoki A. 2002 Three-state cyclic voter model extended with Potts energy. *Phys. Rev. E* **65**, 036115. (doi:10.1103/PhysRevE.65.036115)

18. Sood V, Antal T, Redner S. 2008 Voter models on heterogeneous networks. *Phys. Rev. E* **77**, 041121. (doi:10.1103/PhysRevE.77.041121)

19. Redner S. 2017 Dynamics of voter models on simple and complex networks. (http://arxiv.org/abs/1705.02249)

20. Jalili M, Perc M. 2017 Information cascades in complex networks. *J. Complex Networks* **5**, 665–693. (doi:10.1093/comnet/cnx019)

21. Friedkin NE, Proskurnikov AV, Tempo R, Parsegov SE. 2016 Network science on belief system dynamics under logic constraints. *Science* **354**, 321–326. (doi:10.1126/science.aag2624)

22. Axelrod R. 1997 The dissemination of culture. *J. Conflict Resolut.* **41**, 203–226. (doi:10.1177/0022002797041002001)

23. Holme P, Newman ME. 2006 Nonequilibrium phase transition in the coevolution of networks and opinions. *Phys. Rev. E* **74**, 056108. (doi:10.1103/PhysRevE.74.056108)

24. Zanette DH, Gil S. 2006 Opinion spreading and agent segregation on evolving networks. *Physica D* **224**, 156–165. (doi:10.1016/j.physd.2006.09.010)

25. Nardini C, Kozma B, Barrat A. 2008 Who's talking first? Consensus or lack thereof in coevolving opinion formation models. *Phys. Rev. Lett.* **100**, 158701. (doi:10.1103/PhysRevLett.100.158701)

26. Fu F, Wang L. 2008 Coevolutionary dynamics of opinions and networks: from diversity to uniformity. *Phys. Rev. E* **78**, 016104. (doi:10.1103/PhysRevE.78.016104)

27. Perc M, Szolnoki A. 2010 Coevolutionary games—a mini review. *BioSystems* **99**, 109–125. (doi:10.1016/j.biosystems.2009.10.003)

28. Pew Research Center. 2014 Political polarization in the american public: how increasing ideological uniformity and partisan antipathy affect politics, compromise and everyday life. Technical Report. See http://www.people-press.org/2014/06/12/political-polarization-in-the-american-public/.

29. Pew Research Center. 2016 Partisanship and political animosity in 2016: highly negative views of the opposing party—and its members. Technical Report. See http://www.people-press.org/2016/06/22/partisanship-and-political-animosity-in-2016/.

30. Schmidt AL, Zollo F, Scala A, Betsch C, Quattrociocchi W. 2018 Polarization of the vaccination debate on facebook. (http://arxiv.org/abs/1801.02903)

31. Andris C, Lee D, Hamilton M, Martino M, Gunning C, Selden J. 2015 The rise of partisanship and

super-cooperators in the U.S. House of representatives. *PLoS ONE* **10**, e0123507. (doi:10. 1371/journal.pone.0123507)

32. Pew Research Center. 2017 Partisan conflict and congressional outreach: partisan criticism generates most engagement in social media; most liberal and conservative legislators, party leadership more likely to 'go negative'. Technical Report. See http://www.people-press. org/2017/02/23/partisan-conflict-and-congressional-outreach/.

33. Pew Research Center. 2016 Few Clinton or Trump supporters have close friends in the other camp: how voters are talking—or arguing—about the campaign. Technical Report. See http://www.people-press.org/2016/08/03/few-clinton-or-trump-supporters-have-close-friends-in-the-other-camp/.

34. Antonopoulos CG, Shang Y. 2018 Opinion formation in multiplex networks with general initial distributions. *Sci. Rep.* **8**, 2852. (doi:10. 1038/s41598-018-21054-0)

35. Akerlof GA, Dickens WT. 1982 The economic consequences of cognitive dissonance. *Am. Econ. Rev.* **72**, 307–319.

36. Acharya A, Blackwell M, Sen M. 2016 Explaining preferences from behavior: a cognitive dissonance approach. *J. Pol.* **80**, 400–411. (doi:10.1086/694541)

37. Szolnoki A, Perc M. 2009 Resolving social dilemmas on evolving random networks. *Europhys. Lett.* **86**, 30007. (doi:10.1209/0295-5075/86/30007)

38. Szolnoki A, Perc M. 2009 Emergence of multilevel selection in the prisoner's dilemma game on coevolving random networks. *New J. Phys.* **11**, 093033. (doi:10.1088/1367-2630/11/9/093033)

39. Johnson NF et al. 2017 Population polarization dynamics and next-generation social media algorithms. (http://arxiv.org/abs/1712.06009)

40. Moon H, Lu T-C. 2015 Network catastrophe: self-organized patterns reveal both the instability and the structure of complex networks. *Sci. Rep.* **5**, 9450. (doi:10.1038/srep09450)

41. Castellano C, Fortunato S, Loreto V. 2009 Statistical physics of social dynamics. *Rev. Mod. Phys.* **81**, 591–646. (doi:10.1103/RevModPhys. 81.591)

42. Gawronski B, Brannon SM. 2017 What is cognitive consistency and why does it matter? In *Cognitive dissonance: progress on a pivotal theory in social psychology* (ed. E Harmon-Jones), 2nd edn. Washington, DC: American Psychological Association.

43. Adamic LA, Adar E. 2003 Friends and neighbors on the Web. *Social Networks* **25**, 211–230. (doi:10.1016/S0378-8733(03)00009-1)

44. Borner K, Maru JT, Goldstone RL. 2004 The simultaneous evolution of author and paper networks. *Proc. Natl Acad. Sci. USA* **101**, 5266–5273. (doi:10.1073/pnas.0307625100)

45. Guillermo Abramson MK. 2001 Social games in a social network. *Phys. Rev. E* **63**, 030901.

46. Helbing D, Lozano S. 2010 Phase transitions to cooperation in the prisoner's dilemma. *Phys. Rev. E* **81**, 057102. (doi:10.1103/PhysRevE.81. 057102)

47. Roca CP, Cuesta JA, Sánchez A. 2009 Evolutionary game theory: temporal and spatial effects beyond replicator dynamics. *Phys. Life Rev.* **6**, 208–249. (doi:10.1016/j.plrev.2009.08.001)

48. Moran P. 1958 Random processes in genetics. *Math. Proc. Camb. Philos. Soc.* **54**, 60–71. (doi:1017/S0305004100033193)

49. Tran TD, Hofrichter J, Jost J. 2012 An introduction to the mathematical structure of the Wright–Fisher model of population genetics. *Theory Biosci.* **132**, 73–82. (doi:10. 1007/s12064-012-0170-3)