



Polarisation in online social networks
Detecting polarisation using cross-community social ties

Călin-Mihai Olaru¹
Supervisor(s): Dr. Anna L.D. Latour¹

¹**EEMCS, Delft University of Technology, The Netherlands**

A Thesis Submitted to EEMCS Faculty Delft University of Technology,
In Partial Fulfilment of the Requirements
For the Bachelor of Computer Science and Engineering
June 21, 2026

Name of the student: Călin-Mihai Olaru
Final project course: CSE3000 Research Project
Thesis committee: Dr. Anna L. D. Latour, Dr. Megha Khosla

An electronic version of this thesis is available at <http://repository.tudelft.nl/>.

Abstract

The last decade has seen an uptick in political polarisation on online social networks. This has led to increasing distrust and antagonism, sometimes resulting in conflict. As such, studying the emergence of political polarisation is increasingly important in the current political landscape. The field of network science provides ways to model the underlying structure of social networks. This facilitates the measurement and prediction of polarisation. We propose a way of predicting the emergence of political polarisation by looking at the community structure of a network. Specifically, we examine the extent to which social ties between members of different communities influence how political polarisation takes form. We measure polarisation on two levels: ideological and relational. We experiment on synthetic networks, generated using two different models. One of the models produces networks with homogeneous community structures. The other generates them according to power law distributions. Our findings show a negative correlation between the fraction of cross-community social ties and the level of polarisation in the network.

1 Introduction

Political polarisation has become an important topic in the last decade, especially with the prevalence of social media platforms. Narratives, opinions, and ideologies spread at an unprecedented rate through online social networks, fueling antagonism and social division. Events such as the Cambridge Analytica scandal of 2016 [Faris *et al.*, 2017] and the Romanian elections of 2024 have demonstrated the increasing influence of social networks on politics. Polarisation in the online medium can lead to consequences in the real world, such as distrust in democratic institutions and conflicts between polarised groups of individuals. As such, the study of political polarisation in online social networks has become a matter of societal urgency.

Recently, researchers have turned to the field of network science as a way to model and analyse interactions in online social networks. The complex structures of online networks can be abstracted into graphs, where nodes represent individuals, and edges represent interactions between them. Focusing on the underlying structure of the network offers insights into how influence and opinions spread from node to node and allows researchers to measure properties that may be correlated to political polarisation in a network. In our study, the property of particular interest is the community structure of a network. This represents the degree to which nodes in the network can be grouped into highly-connected clusters [Girvan and Newman, 2002]. We are also interested in the social ties connecting individuals in different communities. These may act as a bridge between groups that have opposing opinions. A low amount of cross-community relations could lead to groups segregating themselves and becoming echo chambers. As these ties become more abundant, people are exposed to

different views and ideas, which could result in less antagonism. Understanding how these ties influence the spread of political attitudes across a network could be a step towards understanding how polarisation develops in online networks.

In this paper, we systematically evaluate the correlation between social ties connecting different communities and the emergence of political polarisation. Specifically, we ask the following research question: *To what extent does the fraction of cross-community social ties influence the level of ideological and relational polarisation in a social network?* Our work is inspired by a recently proposed framework which models networks as two-level graphs of social and attitudinal ties [Mephram *et al.*, 2025]. Measuring polarisation across two dimensions, ideological and relational, gives a more complete view on how polarisation takes shape in a network. We answer our research question by running experiments on synthetic networks. Our findings show that a higher fraction of cross-community social ties is negatively correlated with both dimensions of polarisation.

The remainder of the paper is structured as follows. In Section 2, we discuss previous work on the topic of network polarisation and community structure, as well as how it relates to our work. We introduce notations and define terms we use throughout the paper in Section 3. We detail our approach to measuring polarisation, including the algorithms we use to generate networks and spread opinions in Section 4. In Section 5, we describe our experimental setup and present our findings. We discuss responsible research aspects, including ethical and reproducibility considerations, in Section 6. Section 7 serves as a conclusion, as we summarise the paper, discuss limitations, and give recommendations for future work.

2 Related Work

In this section, we discuss existing methods for measuring polarisation. We show their limitations and their relevance for our work.

A common way of measuring polarisation is through *modularity*. Modularity represents the strength of division of a network into clusters, compared to a random expectation [Coscia, 2025]. There are multiple ways of calculating the modularity of a network. The most common one involves calculating the difference between the actual number of edges and the expected number of edges between each pair of nodes in the graph, averaged out over the whole network [Newman, 2006]. Modularity can be used for community detection purposes and is correlated to the level of polarisation in a network.

A different method uses the nodes along the boundary in the community [Guerra *et al.*, 2021]. The authors argue that modularity is not a direct indicator of polarisation, as even networks with small or negative values of modularity show some level of polarisation. A better way of measuring it is to calculate the popularity of nodes along the boundary. Networks with a high concentration of popular nodes along the boundary show a lower value of polarisation than those with a high concentration of popular nodes inside the communities. While this is a more accurate indication of polarisation, it has only been proven to work for networks that can be divided

into two well-defined communities.

A new framework for computing the level of polarisation in a network has recently been proposed [Mephram *et al.*, 2025]. The network is modeled as a multi-level graph, containing interactions between individuals as well as attitudes towards political issues. Thus, polarisation is measured on a *relational* and *ideological* level. Ideological polarisation is computed as the probability that two individuals show consistent attitudes (both agree or disagree) towards two given political issues. Relational polarisation represents the probability that two individuals are tied given agreement on an issue or not tied given disagreement. Averaging out these metrics results in the overall macro-level network polarisation.

3 Preliminaries

In this section, we introduce our notation and describe the polarisation measurement framework.

3.1 Network Representation

We represent the network as a two-level graph, following the recently introduced framework we previously discussed [Mephram *et al.*, 2025].

The first level is the relational layer, denoted as $G_R = (V_R, E_R)$ where $V_R = \{1, \dots, n\}$ is a set of n nodes and $E_R \subseteq V_R \times V_R$ is a set of undirected edges. The nodes represent the individuals who make up the network. An edge $(u, v) \in E_R$ denotes the existence of a social tie between individuals u and v .

The second level is the attitudinal layer, denoted as $G_O = (V_R \cup V_I, E_O)$, where $V_I = \{i_1, \dots, i_m\}$ is a set of m nodes representing political issues and $E_O \subseteq V_R \times V_I \times \{1, -1\}$ is a set of edges representing political opinions. An edge (u, i_j, o) indicates that individual u holds opinion $o \in \{1, -1\}$ on political issue i_j . A value of 1 represents support, while -1 indicates opposition. An individual u may hold no opinion on a particular political issue i_j , in which case there is no edge between them.

3.2 Polarisation Measurement

Two complementary metrics are defined: ideological polarisation and relational polarisation. Both metrics are made up of two components: *attraction* and *repulsion*.

Ideological polarisation. Ideological polarisation represents the extent to which pairs of individuals hold consistent opinions on political issues. It is measured using *four-cycles*, made up of a pair of individuals and a pair of political issues. The idea is that individuals tend to hold internally consistent opinions regarding political issues. This means that, if two individuals agree on one issue, they are likely to agree on another (attraction), while the opposite is true if they disagree (repulsion). An *agreeing four-cycle* (\hat{A}) occurs when two individuals share the same opinion (1 or -1) on two political issues. A *disagreeing four-cycle* (\hat{D}) occurs when two individuals have different opinions on two political issues. An *inconsistent four-cycle* (\hat{I}) occurs when two individuals agree on one political issue, but disagree on another. We use these structural motifs to calculate the ideological attraction and repulsion:

$$I_{attr} = \frac{\#\hat{A}}{\#\hat{A} + \#\hat{I}} \quad (1)$$

$$I_{rep} = \frac{\#\hat{D}}{\#\hat{D} + \#\hat{I}} \quad (2)$$

Relational polarisation. Relational polarisation captures the extent to which social ties are aligned with opinion similarity. It is measured using *triads*, involving two individuals and their opinions on one political issue. The idea is that in a polarised network, individuals tend to be socially tied to people they agree with, while having few to no relations to people holding opposite opinions. Triads are categorised as closed if the individuals are socially tied, or as open otherwise. Additionally, they are categorised as agreeing or disagreeing based on the opinions of both individuals on a political issue. An overview of the different types of triads can be seen in Table 1. We use these triads to measure relational attraction and repulsion:

$$R_{attr} = \frac{A\Delta}{A\Delta + A\wedge} \quad (3)$$

$$R_{rep} = \frac{D\Delta}{D\Delta + D\wedge} \quad (4)$$

The polarisation metrics are then computed as follows:

$$I_{pol} = \frac{I_{attr} + I_{rep}}{2} \quad (5)$$

$$R_{pol} = \frac{R_{attr} + R_{rep}}{2} \quad (6)$$

Triad	Notation
Agreeing opinion closed	$A\Delta$
Agreeing opinion open	$A\wedge$
Disagreeing opinion closed	$D\Delta$
Disagreeing opinion open	$D\wedge$

Table 1: The four types of triads, classified based on the existence or absence of a social tie between individuals, as well as agreement or disagreement on a political issue.

3.3 Toy Example

We present a toy example to demonstrate how polarisation is measured. Consider the small network shown in Figure 1, with 4 actors divided into 2 communities, and 2 issues.

Members of different communities disagree on both issues, while members of the same community agree on both. Since there are no inconsistent four-cycles, then $I_{attr} = 1.0$, $I_{rep} = 1.0$. All individuals who agree are socially tied, so $R_{attr} = 1.0$. Since nodes B and C are socially tied, but disagree on both issues, then $\#D\Delta = 2$. There are 2 disagreeing open triads between A and C , and 2 between A and D , so $\#D\wedge = 4$. Then $R_{rep} = \frac{4}{6} \approx 0.67$.

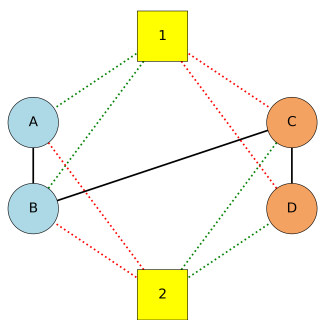


Figure 1: Toy network with 4 actors and 2 issues. Circles are nodes, squares are issues. Red-colored edges represent opposition to an issue, while green-colored edges show support. Black edges represent social ties.

4 Approach

In this section, we discuss our approach to measuring polarisation using social ties. We detail the synthetic network models we use. We then describe how opinions are initialised and how the network evolves.

4.1 Synthetic Network Generation

We perform this study on synthetic network graphs. The main reason for this approach over using real datasets is that synthetic networks allow for full control over the structural properties, particularly the degree to which members of different communities are socially tied. We use two different models for generating synthetic graphs.

Stochastic Block Model. A stochastic block model (SBM) is used to generate a random graph with a known community structure [Holland *et al.*, 1983]. The nodes in the relational layer V_R are partitioned into blocks of roughly equal size. These blocks represent different communities. The edges are generated according to a preference matrix. The preference matrix determines the probability of connecting to nodes in the same community ($p_{in} \in [0, 1]$) and to nodes in another community ($p_{out} \in [0, 1]$). In this study, instead of directly setting p_{in} and p_{out} , we use an expected internal degree deg_{in} . Real social networks tend to get sparser as they get larger, so setting the probabilities directly leads to an unrealistic density. We then calculate the probabilities p_{in} and p_{out} as follows:

$$p_{in} = \frac{deg_{in}}{\text{community size}} \quad (7)$$

$$p_{out} = \frac{\alpha}{1 - \alpha} \cdot p_{in} \quad (8)$$

The variable α represents our independent variable, which is the cross-community tie ratio. Since the stochastic block model produces networks with a homogeneous community structure, it can be used to test the model on particular cases, such as a two-party system. However, real datasets are rarely homogeneous in their degree distributions and community sizes. Thus, testing only on networks generated using SBM may lead to inaccurate findings.

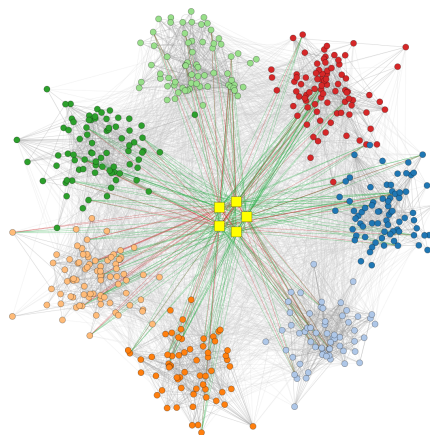


Figure 2: Network of 500 nodes, generated using LFR ($\mu = 0.4$), after opinion initialization.

Lancichinetti–Fortunato–Radicchi Benchmark Model.

The LFR benchmark model is used to generate networks with heterogeneous distributions for node degrees and community sizes [Lancichinetti *et al.*, 2008]. These follow power laws with exponents γ and β respectively. The edges representing social ties are generated according to the mixing parameter $\mu \in [0, 1]$. This parameter represents the fraction of edges that connect nodes from different communities. This is our independent variable in this case, as it directly models the percentage of cross-community ties. Networks built using the LFR model match the heterogeneity of empirical data [Clauset *et al.*, 2009]. In this work, we use LFR benchmark networks to verify that our findings are robust and not artifacts of the SBM homogeneity assumptions.

4.2 Opinion Initialisation

The attitudinal layer is made up of opinions on M independent political issues. Political attitudes are split into two competing platforms. For the first platform, we randomly assign an opinion as 1 or -1 for each issue. The second platform is made to have an ideological overlap fraction of p_{agree} with the first. This is meant to add a level of inconsistency that might be present in real networks. Every community in the network is randomly and uniformly assigned a political platform to align with.

After we assign the platforms, we select individuals in the network as initial “seeds” to begin the evolution of the network (which we discuss in Section 4.3). To prevent high-degree nodes and larger communities from having a disproportionate influence, we randomly select roughly 10% of nodes in each network as seeds. These nodes are then assigned the political platforms of the community they belong to. The other nodes do not have a connection with the attitudinal layer initially. An example of a network with initialised opinions can be seen in Figure 2.

4.3 Network Evolution

We describe the algorithm used to spread opinions among individuals and change the structure of the network. A formal overview of the algorithm can be seen in Algorithm 1.

We model the evolution of the network as an asynchronous process, running for a maximum of T steps. At each discrete step t , we randomly shuffle the set of nodes V_R . We do this to mitigate the influence of initial ordering on how opinions spread. Since the algorithm is asynchronous, any change to the network is reflected in the global state before the next node is processed. When a node i is processed, two things happen: *homophilous rewiring* and *opinion spread*.

Homophilous rewiring (lines 5-12). Node i scans its neighbourhood N_i . The nodes of interest here are all nodes that disagree with i on an issue $x \in V_I$ (i.e. $O_{j,x} = -O_{i,x}$). For each node j fitting this criterion, the social tie (i, j) is severed with probability p_{rewire} . If the tie is indeed severed, we sample up to 50 random non-neighbours k with the same opinion on issue x (i.e. $O_{k,x} = O_{i,x}$). If such a node is found, a new social tie (i, k) is formed. We use this phase of the algorithm to model the *homophilous selection* process in social networks [McPherson *et al.*, 2001].

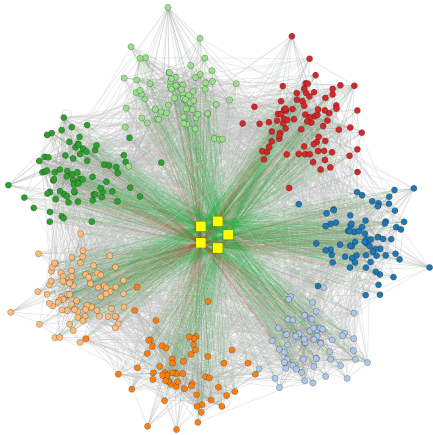


Figure 3: Network of 500 nodes, generated using LFR ($\mu = 0.4$), after running the evolution algorithm.

Opinion spread (lines 13-19). After rewiring social ties, the network evolution continues on the ideological layer. We simulate the spread of opinions according to the Linear Threshold model [Kempe *et al.*, 2003]. Each actor node has a randomly assigned threshold $\theta_i \sim U(0.05, 0.33)$. This threshold represents the resistance of each individual to peer pressure. In social networks, opinions typically require multiple reinforcements before being adopted. This phenomenon is known as a complex contagion [Centola and Macy, 2007]. Node i aggregates the opinions of its updated neighbourhood N'_i on issue x . Node i can adopt an opinion $O \in \{1, -1\}$ if the fraction of neighbours holding it represents a majority. If this fraction exceeds θ_i , the node adopts the majority opinion. If the threshold is not met, i remains neutral on issue x , i.e. the attitudinal tie (i, x) is not formed.

Algorithm 1 Network Evolution

Require: V_R : set of actors, V_I : set of issues, T : maximum number of steps

- 1: **for** $t \leftarrow 1$ **to** T **do**
- 2: Shuffle V_R
- 3: $converged \leftarrow true$
- 4: **for each** $i \in V_R$, **each** $x \in V_I$ **do**
- 5: **for each** $j \in N_i$ s.t. $O_{i,x} \neq 0$ and j disagrees with i on x **do**
- 6: **if** $\text{random}(0, 1) < p_{rewire}$ **then**
- 7: sever tie (i, j) ; sample up to 50 random non-neighbours k with $O_{k,x} = O_{i,x}$
- 8: **if** such k exists **then** establish tie (i, k)
- 9: **end if**
- 10: $converged \leftarrow false$
- 11: **end if**
- 12: **end for**
- 13: let F^+, F^- be the fraction of N_i holding $+1, -1$ on x
- 14: **if** $F^+ \geq \theta_i$ **and** $F^+ > F^-$ **then** $O_{i,x} \leftarrow 1$
- 15: **else if** $F^- \geq \theta_i$ **and** $F^- > F^+$ **then** $O_{i,x} \leftarrow -1$
- 16: **end if**
- 17: **if** $O_{i,x}$ changed **then** $converged \leftarrow false$
- 18: **end if**
- 19: **end for**
- 20: **if** $converged$ **then break**
- 21: **end if**
- 22: **end for**
- 23: **return** G'

Convergence. The algorithm converges if, after a full step, no rewiring or opinion change takes place. The algorithm terminates within T steps, even if the convergence criterion is not met. The modified network G' is returned after termination. Figure 3 shows an example of a network after running the evolution algorithm.

5 Experiments

In this section, we detail the experiments we conducted to answer our research questions, as well as our findings.

5.1 Research Questions

As stated in Section 1, our main research question is the following:

- **RQ:** *To what extent does the fraction of cross-community social ties influence the level of ideological and relational polarisation in a social network?*

We split our main research question into two sub-questions. Each sub-question isolates a different network structure and corresponds to a single synthetic network generation model.

- **RQ1:** *How does the fraction of cross-community ties affect ideological and relational polarisation in networks of various sizes, with a homogeneous community structure?*

We answer this sub-question by experimenting on networks generated using the Stochastic Block model.

- **RQ2:** *How does the fraction of cross-community ties affect ideological and relational polarisation in networks of various sizes, with heterogeneous degree and community size distributions?*

We answer this sub-question by experimenting on networks generated using the LFR Benchmark model.

5.2 Experimental Setup

We explain the parameters and the software used in the experiments. We run experiments on networks of various configurations to make sure we answer our research questions thoroughly. A full overview of the experimental parameters can be seen in Table 2.

Global parameters. We test our model on networks of different sizes and structures to get a complete view on the effect of cross-community social ties. We work on networks of $N \in \{500, 1000, 5000\}$ individuals. We initialise political platforms of $M = 5$ issues. This number models a sufficiently complex political landscape, while not making the four-cycle calculations too computationally expensive. The two platforms have an ideological overlap of $p_{agree} = 0.2$. This offers a small degree of inconsistency that might be present between real political platforms. A fraction of $r_{init} = 0.1$ of the nodes in each community are selected as initial seeds to start the network evolution. Since both models involve some stochasticity, we run 20 iterations for each parameter configuration. An iteration consists of generating a network, initialising opinions, simulating the network evolution algorithm, and finally computing the polarisation metrics.

SBM parameters. For the stochastic block model, we generate graphs containing $c \in \{2, 4, 8\}$ communities to test our model under different structures. For the cross-community social tie fraction, we set the values $\alpha \in \{0.1, 0.15, \dots, 0.6\}$. We set the average internal degree to $deg_{in} = 20$. This helps produce somewhat sparse graphs, matching real social networks, while allowing the influence to spread through the network. Due to homogeneity assumptions, SBM graphs are not accurate representations of larger, real datasets. As such, we only generate SBM graphs of sizes $N \in \{500, 1000\}$.

LFR parameters. The LFR benchmark model generates graphs according to power law distributions. The bounds of the distribution exponents are as follows: $2 \leq \gamma \leq 3$ for the degree distribution and $1 < \beta \leq 2$ for the community size distribution. As such, we choose $\gamma = 2.5$ and $\beta = 1.5$. For the mixing parameter, we set the values $\mu \in \{0.1, 0.15, \dots, 0.6\}$. It might seem unintuitive for more than half of the social ties to be external, however it has been shown that community structure is preserved for values of μ up to 0.75 [Lancichinetti and Fortunato, 2009].

Software. We implemented the model in Python 3.12.1. The two main libraries that we used were NetworkX 3.6.1 and igraph 1.0.0, for network processing and graph visualizations. We used pandas 3.0.3 and matplotlib 3.10.9 for data analysis and for plotting the results. In addition, we used numpy 2.4.4 for mathematical functions.

Global Parameter	Notation	Value(s)
Network size	N	$\{500, 1000, 5000\}$
Issues	M	5
Ideological overlap	p_{agree}	0.2
Initial seed ratio	r_{init}	0.1
Iterations per config	it	20
Rewiring probability	p_{rewire}	0.05
Maximum steps	T	30
SBM Parameter	Notation	Value(s)
Communities	c	$\{2, 4, 8\}$
Avg. internal degree	deg_{in}	20
Cross-community ratio	α	$\{0.1, 0.15, \dots, 0.6\}$
LFR Parameter	Notation	Value(s)
Degree dist. exp.	γ	2.5
Community size exp.	β	1.5
Mixing parameter	μ	$\{0.1, 0.15, \dots, 0.6\}$
Minimum degree	deg_{min}	5
Maximum degree	deg_{max}	50

Table 2: The parameters used for the experiment. Parameters in bold represent the independent variables that we are measuring polarisation for.

5.3 Results

We present our experimental results and answer the research question. We conduct our experiments using the parameters we outline in Section 5.2. For clarity purposes, we only plot the average values obtained for each parameter configuration. An overview of the full data, including the standard deviations, can be seen in Appendix A.

RQ1. Figures 4 and 5 illustrate the results we obtained for networks generated using the stochastic block model, under various community structures. Overall, both metrics decrease as the fraction of cross-community ties increases across configurations. However, there are some interesting results to highlight. For networks of 1000 nodes, relational polarisation seems to spike at various points. This could be caused by the random nature of stochastic models. As the number of communities increases, polarisation starts decreasing faster. For $c = 8$ communities, the ideological polarisation starts lower, no matter the number of nodes. This is worth investigating.

RQ2. Figure 6 illustrates the consistent decrease of polarisation as the mixing parameter μ increases. The decrease is smoother and more consistent than in the case of SBM graphs. However, there are still some spikes for the relational polarisation between $\mu = 0.45$ and $\mu = 0.55$. This is an area where the community structure tends to become less clear. Ideological polarisation seems to mostly plateau around 0.5 after $\mu = 0.35$.

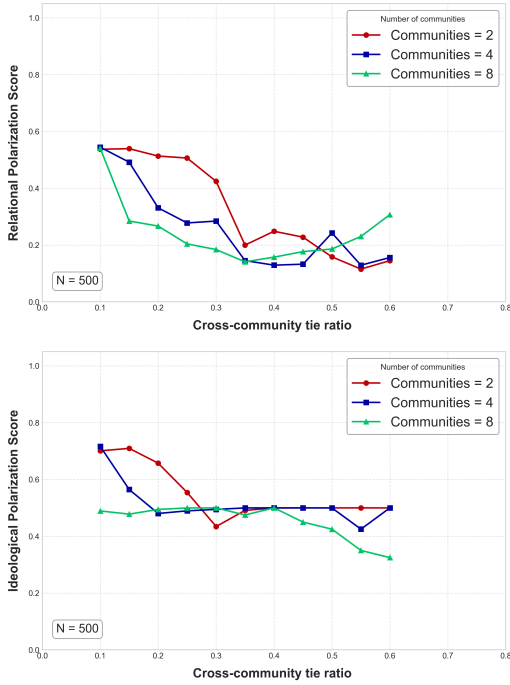


Figure 4: The influence of cross-community ties on relational (top) and ideological (bottom) polarisation, generated using SBM, with 500 nodes.

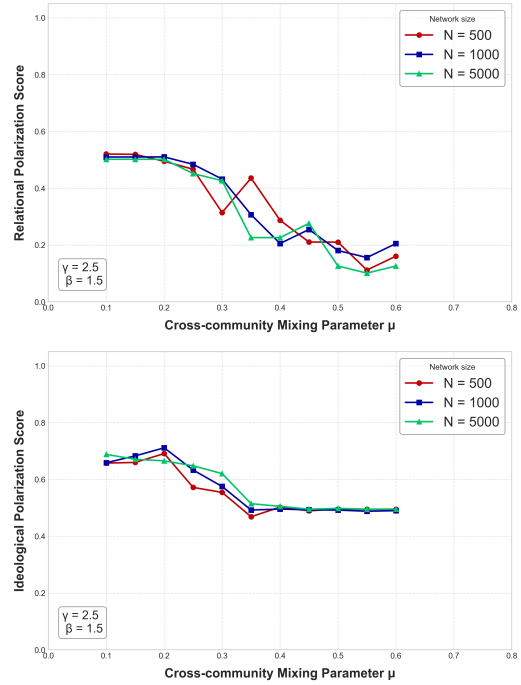


Figure 6: The influence of cross-community ties on relational (top) and ideological (bottom) polarisation, on networks generated using LFR, with $\gamma = 2.5$ and $\beta = 1.5$

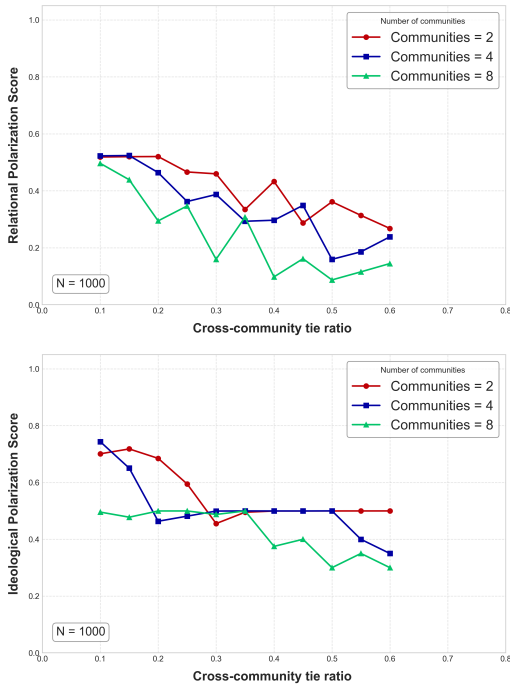


Figure 5: The influence of cross-community ties on relational (top) and ideological (bottom) polarisation, generated using SBM, with 1000 nodes.

6 Responsible Research

Political polarisation is a sensitive topic. Our intent in this paper is not to express, imply, or support any particular political views. Our contribution is purely a systematic investigation of the correlation between social ties and network polarisation. We caution against over-interpreting our results as causal claims about real-world political behaviour in social networks. We recognise that models of polarisation could be repurposed to increase it rather than understanding or mitigating it. We explicitly discourage such use.

To ensure the reproducibility of the results, we include all parameter values we used for the experiments in Section 5.2. We also describe the software used to implement and run our model. For transparency purposes, the source code of our model is publicly available to allow verification, replication, and extension of our work ¹. We include the limitations of our research in Section 7.2.

For the purpose of this study, we did not use Generative AI tools to generate or interpret content or results. During the research, AI was only used to clarify certain concepts. For writing the paper, AI tools were only used for spell-checking and formatting certain figures and tables. In terms of coding, AI tools were only used to help with plotting and visualizing graphs, and as a last resort to help with bug-fixing. The models we used were Google Gemini (3.5 Flash and 3.1 Pro), as well as Anthropic's Claude (Sonnet 4.6). An overview of the tasks we used LLMs for and the prompts we used is given in Appendix B.

¹<https://github.com/KhalyN1/Network-Polarisation-RP>

7 Conclusion

In this section, we summarise the findings of the paper, and then discuss limitations and suggestions for future work.

7.1 Discussion

In this paper, we study the impact of cross-community social ties on the emergence of relational and ideological polarisation. We look at how the initial fraction of social ties between different communities affects the way opinions spread and networks evolve. We measure two dimensions of polarisation: ideological and relational. We test our model on synthetic networks, generated using the Stochastic Block model and LFR benchmark.

Our experiments show that the fraction of cross-community social ties is negatively correlated with both forms of polarisation, across all parameter configurations. This is consistent across both SBM and LFR networks. The LFR results show a smoother decline, suggesting that heterogeneous networks are less volatile than homogeneous ones. Our findings indicate that encouraging individuals to interact with people outside their circle in online social networks could lead to a lower level of polarisation.

7.2 Limitations & Future Work

Real datasets. Our current work has only been carried out on synthetic graphs. The model could be worth testing on real networks to check its robustness and the credibility of the research. The Stanford Large Network Dataset Collection is a good place to check out such datasets, particularly for social networks [Leskovec and Krevl, 2014].

Weighted networks. Currently, our model works on graphs with unweighted social edges and attitudinal edges with values $O \in \{-1, 1\}$. This is a very simplistic framing that does not capture the full complexity of real social interactions. A possible next step is to work on graphs where edges have weights $W_{social} \in (0, 1]$ and $W_{attitude} \in [-1, 1]$. For social edges, the weight determines the strength of a tie between two individuals. In the case of attitudinal edges, the sign denotes the opinion on a certain issue, while the absolute value determines the strength of the opinion.

Opinion dynamics. On networks with weighted edges, a possible approach would be to use more complex opinion dynamics models, rather than the Linear Threshold diffusion model. Some examples include the Deffuant Bounded Confidence model [Deffuant, 2002] and the Dandekar model [Dandekar *et al.*, 2013].

Acknowledgements

I would like to thank my supervisor, Dr. Anna L.D. Latour, for the continuous feedback, guidance, and support within the CSE3000 Research Project course. I would also like to thank PhD students Fatemeh Jafari and Bruna Bazaluk Machado Videira for offering valuable information and advice throughout the project.

References

- [Centola and Macy, 2007] Damon Centola and Michael Macy. Complex contagions and the weakness of long ties. *American Journal of Sociology*, 113(3):702–734, November 2007.
- [Clauset *et al.*, 2009] Aaron Clauset, Cosma Rohilla Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. *SIAM Review*, 51(4):661–703, 2009.
- [Coscia, 2025] Michele Coscia. *The Atlas for the Aspiring Network Scientist*. IT University of Copenhagen, 2 edition, 2025. Accessed: Feb. 11, 2025.
- [Dandekar *et al.*, 2013] Pranav Dandekar, Ashish Goel, and David T. Lee. Biased assimilation, homophily, and the dynamics of polarization. *Proceedings of the National Academy of Sciences*, 110(15):5791–5796, April 2013.
- [Deffuant, 2002] Guillaume; Amblard Deffuant. How can extremism prevail? a study based on the relative agreement interaction model, October 2002.
- [Faris *et al.*, 2017] Robert Faris, Hal Roberts, Bruce Etling, Nikki Bourassa, Ethan Zuckerman, and Yochai Benkler. Partisanship, Propaganda, and Disinformation: Online Media and the 2016 U.S. Presidential Election, August 2017.
- [Girvan and Newman, 2002] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences*, 99(12):7821–7826, 2002.
- [Guerra *et al.*, 2021] Pedro Guerra, Wagner Meira Jr., Claire Cardie, and Robert Kleinberg. A Measure of Polarization on Social Media Networks Based on Community Boundaries. *Proceedings of the International AAAI Conference on Web and Social Media*, 7(1):215–224, August 2021.
- [Holland *et al.*, 1983] Paul W. Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt. Stochastic blockmodels: First steps. *Social Networks*, 5(2):109–137, 1983.
- [Kempe *et al.*, 2003] David Kempe, Jon Kleinberg, and Éva Tardos. Maximizing the spread of influence through a social network. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '03*, pages 137–146, New York, NY, USA, August 2003. Association for Computing Machinery.
- [Lancichinetti and Fortunato, 2009] Andrea Lancichinetti and Santo Fortunato. Community detection algorithms: A comparative analysis. *Physical Review E*, 80(5):056117, November 2009.
- [Lancichinetti *et al.*, 2008] Andrea Lancichinetti, Santo Fortunato, and Filippo Radicchi. Benchmark graphs for testing community detection algorithms. *Physical Review E*, 78(4):046110, October 2008.
- [Leskovec and Krevl, 2014] Jure Leskovec and Andrej Krevl. SNAP Datasets: Stanford large network dataset collection. <http://snap.stanford.edu/data>, June 2014.

- [McPherson *et al.*, 2001] Miller McPherson, Lynn Smith-Lovin, and James M. Cook. Birds of a feather: Homophily in social networks. *Annual Review of Sociology*, 27(Volume 27, 2001):415–444, August 2001.
- [Mepham *et al.*, 2025] K. Mepham, A. Vörös, and C. Stadtfeld. Network polarization: The study of political attitudes and social ties as dynamic multilevel networks. *Network Science*, 13:e7, January 2025.
- [Newman, 2006] M. E. J. Newman. Modularity and community structure in networks. *Proceedings of the National Academy of Sciences*, 103(23):8577–8582, June 2006.
- [Shakarian *et al.*, 2015] Paulo Shakarian, Abhinav Bhatnagar, Ashkan Aleali, Elham Shaabani, and Ruocheng Guo. *The Independent Cascade and Linear Threshold Models*, page 35–48. Springer International Publishing, Cham, 2015.

A Full measurement data

We show the full data gathered from the measurements for both the relational polarisation (**r_pol**) and ideological polarisation (**i_pol**). We include means and standard deviations (**std**). Standard deviations are rounded to one significant digit. Means are then rounded to that same digit.

α	r_pol mean	r_pol std	i_pol mean	i_pol std
0.1	0.537	0.005	0.7	0.2
0.15	0.539	0.006	0.7	0.2
0.2	0.5	0.1	0.6	0.2
0.25	0.5	0.1	0.6	0.2
0.3	0.4	0.2	0.43	0.04
0.35	0.2	0.2	0.49	0.03
0.4	0.2	0.2	0.4994	0.0009
0.45	0.2	0.2	0.4997	0.0004
0.5	0.2	0.2	0.4997	0.0003
0.55	0.115	0.174	0.4999	0.0003
0.6	0.145	0.196	0.4998	0.0004

Table 3: Polarisation means and standard deviations for SBM graphs with $N = 500$ nodes and $c = 2$ communities.

α	r_pol mean	r_pol std	i_pol mean	i_pol std
0.1	0.545	0.006	0.7	0.2
0.15	0.5	0.2	0.6	0.2
0.2	0.3	0.2	0.48	0.06
0.25	0.3	0.2	0.49	0.03
0.3	0.3	0.2	0.49	0.02
0.35	0.1	0.2	0.500	0.001
0.4	0.1	0.2	0.4999	0.0002
0.45	0.1	0.2	0.4999	0.0002
0.5	0.2	0.2	0.4997	0.0005
0.55	0.1	0.1	0.4	0.2
0.6	0.2	0.1	0.500	0.001

Table 4: Polarisation means and standard deviations for SBM graphs with $N = 500$ nodes and $c = 4$ communities.

α	r_pol mean	r_pol std	i_pol mean	i_pol std
0.1	0.54	0.01	0.5	0.1
0.15	0.3	0.2	0.48	0.04
0.2	0.3	0.2	0.49	0.02
0.25	0.2	0.2	0.500	0.001
0.3	0.2	0.2	0.4998	0.0004
0.35	0.1	0.1	0.5	0.1
0.4	0.2	0.1	0.4999	0.0002
0.45	0.2	0.1	0.5	0.2
0.5	0.2	0.1	0.4	0.2
0.55	0.2	0.2	0.4	0.2
0.6	0.3	0.2	0.3	0.2

Table 5: Polarisation means and standard deviations for SBM graphs with $N = 500$ nodes and $c = 8$ communities.

α	r_pol mean	r_pol std	i_pol mean	i_pol std
0.1	0.519	0.002	0.7	0.2
0.15	0.520	0.003	0.7	0.2
0.2	0.520	0.003	0.7	0.2
0.25	0.5	0.2	0.6	0.2
0.3	0.5	0.2	0.46	0.05
0.35	0.3	0.2	0.50	0.02
0.4	0.4	0.2	0.499	0.001
0.45	0.3	0.3	0.499	0.001
0.5	0.4	0.2	0.4995	0.0003
0.55	0.3	0.2	0.4996	0.0003
0.6	0.3	0.2	0.4997	0.0003

Table 6: Polarisation means and standard deviations for SBM graphs with $N = 1000$ nodes and $c = 2$ communities.

α	r_pol mean	r_pol std	i_pol mean	i_pol std
0.1	0.523	0.003	0.7	0.2
0.15	0.524	0.004	0.6	0.2
0.2	0.5	0.2	0.46	0.07
0.25	0.4	0.2	0.48	0.04
0.3	0.4	0.2	0.499	0.001
0.35	0.3	0.2	0.4997	0.0003
0.4	0.3	0.2	0.4996	0.0004
0.45	0.3	0.2	0.4995	0.001
0.5	0.2	0.2	0.4999	0.0002
0.55	0.2	0.2	0.4	0.2
0.6	0.2	0.2	0.3	0.2

Table 7: Polarisation means and standard deviations for SBM graphs with $N = 1000$ nodes and $c = 4$ communities.

α	r_pol mean	r_pol std	i_pol mean	i_pol std
0.1	0.5	0.1	0.5	0.1
0.15	0.4	0.2	0.48	0.04
0.2	0.3	0.2	0.4996	0.0004
0.25	0.3	0.2	0.4996	0.0004
0.3	0.2	0.2	0.49	0.06
0.35	0.3	0.2	0.500	0.001
0.4	0.1	0.1	0.4	0.2
0.45	0.2	0.2	0.4	0.2
0.5	0.1	0.1	0.3	0.2
0.55	0.1	0.1	0.3	0.2
0.6	0.1	0.2	0.3	0.2

Table 8: Polarisation means and standard deviations for SBM graphs with $N = 1000$ nodes and $c = 8$ communities.

μ	r_pol mean	r_pol std	i_pol mean	i_pol std
0.1	0.520	0.003	0.7	0.2
0.15	0.519	0.003	0.7	0.2
0.2	0.5	0.1	0.7	0.2
0.25	0.5	0.2	0.6	0.1
0.3	0.3	0.2	0.6	0.1
0.35	0.4	0.2	0.47	0.04
0.4	0.3	0.3	0.50	0.06
0.45	0.2	0.2	0.49	0.03
0.5	0.2	0.2	0.50	0.02
0.55	0.1	0.2	0.49	0.02
0.6	0.2	0.2	0.49	0.02

Table 9: Polarisation means and standard deviations for LFR graphs with $N = 500$ nodes.

μ	r_pol mean	r_pol std	i_pol mean	i_pol std
0.1	0.510	0.001	0.7	0.2
0.15	0.510	0.002	0.7	0.2
0.2	0.510	0.002	0.7	0.2
0.25	0.5	0.1	0.6	0.2
0.3	0.4	0.2	0.57	0.2
0.35	0.3	0.2	0.49	0.07
0.4	0.2	0.2	0.49	0.01
0.45	0.3	0.3	0.49	0.02
0.5	0.2	0.2	0.49	0.02
0.55	0.1	0.2	0.49	0.03
0.6	0.2	0.2	0.49	0.03

Table 10: Polarisation means and standard deviations for LFR graphs with $N = 1000$ nodes.

μ	r_pol mean	r_pol std	i_pol mean	i_pol std
0.1	0.502	0.001	0.7	0.2
0.15	0.501	0.001	0.7	0.2
0.2	0.501	0.001	0.7	0.2
0.25	0.4	0.2	0.6	0.2
0.3	0.4	0.2	0.6	0.2
0.35	0.2	0.3	0.51	0.03
0.4	0.2	0.3	0.50	0.05
0.45	0.3	0.2	0.49	0.06
0.5	0.1	0.2	0.50	0.06
0.55	0.1	0.2	0.49	0.02
0.6	0.1	0.2	0.49	0.02

Table 11: Polarisation means and standard deviations for LFR graphs with $N = 5000$ nodes.

B Use of Generative AI

As previously mentioned, we did not use Generative AI to generate or interpret results. All the data present in plots and tables has been obtained through experimentation. We also did not use Large Language Models to outright generate new content. We used LLMs for the following tasks:

- Clarifying concepts (3.5 Flash)
- Proofreading and improving writing (3.5 Flash, Sonnet 4.6)
- Comparing approaches (3.5 Flash, Sonnet 4.6)
- Formatting .csv data into \LaTeX tables (3.1 Pro)
- Generating plotting and visualization code (3.1 Pro)
- Last resort bug-fixing (3.1 Pro)

Below we give some examples of the prompts we gave and the answers we received. For clarity and simplicity, some of the prompts were bundled together. Some answers were summarised or truncated if they were too convoluted.

Visualizing networks

Prompt: *I have an igraph network generated using the Stochastic Block model. The network has two layers. The social layer is made up of nodes belonging to different communities, with edges being social interactions. The attitudinal layer is made up of individuals and issues, with edge values representing opinions (1 or -1). I would like some code to visualize this graph. Individuals are circles. Each community has a different color, and each member of a community has the same color. Communities are positioned in a large circle around the center. Members of the same community are more closely grouped together. Issues are yellow squares, positioned in a smaller circle around the center. Social edges are grey. Opinion edges are green if the value is 1, red if the value is -1.*

Answer: To achieve exactly what you want, we can manually compute a custom layout based on trigonometry (polar coordinates) to place the communities and issues in their respective concentric circles. Here is a complete, self-contained Python script using igraph (and matplotlib for the rendering backend) to accomplish this. [Code for visualizing a graph using igraph and matplotlib]

Note: We used this code to visualize the networks in Figure 2 and Figure 3. We did not have to make many changes to it, only changing the positioning of the nodes a little.

Comparing SBM and LFR

Prompt: *For now, I have only used the Stochastic Block model to generate graphs. However, it might not be very realistic. How does the Lancichinetti–Fortunato–Radicchi benchmark model compare in terms of random graph generation with community structure?*

Answer: You've hit on a major limitation of the Stochastic Block Model (SBM). While SBM is mathematically clean and a great baseline, it generates networks that look highly artificial.

The Lancichinetti–Fortunato–Radicchi (LFR) benchmark was specifically designed to address these unrealistic features. It is widely considered one of the gold standards for testing community detection algorithms because it mimics the statistical properties found in real-world social and biological networks.

Here is how the LFR benchmark compares to the SBM across the key dimensions of realism.

- **In SBM:** The degree distribution within a community is typically Poisson or binomial. This means most individuals have roughly the same number of friends, and it is exceedingly rare to have "super-hubs" (individuals with massive amounts of connections). You must explicitly define the sizes of the blocks (e.g., three communities of 50 nodes). In standard formulations, communities are often treated as roughly uniform or manually scaled.
- **In LFR:** It incorporates a power-law (scale-free) degree distribution, which matches real-world social networks. You will naturally get a few highly connected "influencers" and a long tail of less-connected individuals. Community sizes themselves follow a power-law distribution. In a large population, you will naturally get many small, tight-knit communities and a few massive, sprawling ones. This heterogeneous community size is an inherent property of real-world networks. One of LFR's greatest strengths is how it handles the "fuzziness" of communities via a mixing parameter, denoted as μ (mu). Each node allocates a fraction $(1 - \mu)$ of its links to members of its own community, and a fraction μ of its links to the rest of the network. While SBM can mimic this by carefully tuning its probability matrix (P_{ij}), LFR makes this explicitly controllable as a global topology dial.

Note: We fact-checked this response using research papers and articles on both graph models [Holland *et al.*, 1983; Lancichinetti *et al.*, 2008; Lancichinetti and Fortunato, 2009]. This response proved to be quite a good comparison. We ended up using both models in the study. This comparison also helped us split the main research question into sub-questions, as there were some differences in the results. The structural differences proved to have an effect on how polarisation evolves in the network.

Comparing Linear Threshold and Independent Cascade

Prompt: *Right now, I am only looking at a final "snapshot" of the network. I want to initialise networks with only a few individuals having opinions, and let the network evolve. For a social/friendship network, which would work better: Linear Threshold or Independent Cascade? Give a short comparison of the two diffusion models.*

Answer: For a social/friendship network where you are modeling opinions and attitudes rather than a virus or a viral marketing fad, the **Linear Threshold (LT) model** is generally a much better fit than the **Independent Cascade (IC) model**. Here is a short comparison of how they behave and why they suit different types of diffusion.

1. Linear Threshold (LT) Model

The LT model is based on peer pressure and reinforce-

ment. Every individual has a personal threshold (e.g., 30%). They will only adopt the new opinion if the percentage of their neighbors holding that opinion exceeds their threshold.

- **Mechanism:** Deterministic once thresholds are set. It aggregates influence from all neighbors.
- **Social Realism:** High for opinions. People rarely change their political stance or core beliefs just because one friend mentions it. It requires a "critical mass" of friend groups to trigger a change.
- **Best Used For:** Complex contagions—behaviors, social movements, opinions, and high-risk choices that require validation from multiple sources.

2. Independent Cascade (IC) Model

The IC model behaves like a disease outbreak. If a friend adopts an opinion, they get exactly one chance to infect each of their neighbors with a certain probability (p). If they fail, they don't try again.

- **Mechanism:** Probabilistic. Each edge acts independently, regardless of what the rest of the neighborhood is doing.
- **Social Realism:** Low for deep opinions, but high for information. You can hear a piece of gossip or a news headline from just one person and instantly "catch" the information.
- **Best Used For:** Simple contagions—the spread of viral memes, breaking news, or information awareness where peer pressure isn't required.

Note: We used this mostly to get a clear, short comparison. We fact-checked this response using literature on both diffusion models [Kempe *et al.*, 2003; Shakarian *et al.*, 2015]. This comparison helped us make a decision on which model to use. We ended up using the Linear Threshold model in our research, as it models behaviour that is more common in friendship networks when dealing with the spread of opinions.

Improving Experimental Setup

Prompt: *This is my current experimental setup [Pasted experimental parameters]. Critique it and give me some suggestions for improving it.*

Answer: Here is a structural and methodological critique of your experimental setup. The critique addresses parameter choices, scaling rules, and experimental soundness, followed by concrete recommendations.

Methodological Critique

- **Asymmetric Network Size Ranges (N):** The LFR experiments test network scales across several orders of magnitude ($N \in \{100, 500, 1000, 5000\}$). However, the SBM experiments omit the largest configuration ($N = 5000$). To support consistent comparative claims regarding how network topology scales with diffusion dynamics, both models should ideally span identical sizes.
- **Rigid p_{in} Assignment in SBM:** The SBM script holds the internal community edge probability fixed at $p_{in} = 0.15$ while scaling N . In random graphs, keeping edge

probabilities constant while N increases causes the average degree to grow linearly ($\langle k \rangle \approx p_{in} \cdot \frac{N}{C}$, where C is the number of communities). As a result, the SBM graphs at $N = 1000$ will be dramatically denser than those at $N = 100$. This breaks structural parity with the LFR setup, where degrees are strictly bounded by `min_degree` and `max_degree`.

- **Misaligned Independent Variables:** In LFR, the mixing parameter μ measures the fraction of *edges* an individual sends outside their community. In your SBM, the independent variable is p_{out} (absolute probability). While related, a direct cross-model comparison is difficult because p_{out} does not represent a normalized mixing ratio.
- **Static Attitudinal Layer ($M = 10$):** The number of issues is fixed at 10 across all simulations. While this isolates social structural effects, it ignores how opinion dynamics scale when the issue space expands or contracts relative to the population size N .

Proposed Technical Improvements

To achieve rigorous structural alignment between both generative layers, the following mathematical alignments and enhancements are recommended:

Structural Alignment via Average Degree

To make SBM and LFR comparable, the SBM probabilities should be derived from target expected degrees rather than hardcoded constants. Let $\langle k_{in} \rangle$ be the expected internal degrees and $\langle k_{out} \rangle$ be the expected external degrees.

For an SBM with C equally sized communities of size $n_c = N/C$, the probabilities p_{in} and p_{out} should scale dynamically with N according to:

$$p_{in} = \frac{\langle k_{in} \rangle}{n_c - 1} = \frac{(1 - \mu)\langle k \rangle}{\frac{N}{C} - 1} \quad (9)$$

$$p_{out} = \frac{\langle k_{out} \rangle}{N - n_c} = \frac{\mu\langle k \rangle}{N - \frac{N}{C}} \quad (10)$$

Using this mapping allows μ to act as the unified independent variable for both the LFR and SBM experiments.

Note: Most of these proposed improvements ended up being implemented in the model. Deriving the SBM parameters from expected degrees and ratios made it so we had similar parameters to LFR. Thus, there was more consistency in the experiment and we had to tune fewer parameters every time we changed the network size. We did not end up testing on networks of size $N = 5000$ for SBM, as the model does not produce realistic networks, so testing it on larger sizes did not really add anything to the experiment. We also kept a fixed number of issues, but reduced it from 10 to 5, to reduce the runtime of the experiments.

LFR parameters

Prompt: *I am using the LFR model to generate graphs. For the mu variable, i am using the following values: [0.1, 0.15,...0.6] For the other parameters in LFR, I would like to keep them fixed. What values should i use for min degree, max degree, gamma, beta?*

Answer: Good framing — these are your nuisance parameters: you want them set to plausible, stable values that don't interact badly with your μ sweep or distort your polarization signal.

The Parameters & What They Do

gamma — exponent of the degree distribution (power law). Controls degree heterogeneity.

beta — exponent of the community size distribution (power law). Controls community size heterogeneity.

k / k_min — minimum (and mean) degree. Sets how connected nodes are on average. k_max — maximum degree. Caps the hubs.

Recommended Values γ (degree distribution exponent)

- Use: 2.5 – 3.0, with 2.5 being the standard default
- Real social networks empirically fall in [2, 3]; 2.5 is the most cited in LFR literature
- Avoid $\gamma < 2$ (non-normalizable distribution, unstable graphs) or $\gamma > 3.5$ (too homogeneous, loses realistic hub structure)
- Since you're not varying it, fix at 2.5 and mention it as a controlled parameter

β (community size exponent)

- Use: 1.5, the conventional default
- Less studied than γ but values in [1, 2] are standard
- Higher $\beta \rightarrow$ more similarly-sized communities; lower \rightarrow more size variance
- For polarization research, overly unequal community sizes can confound your metric, so 1.5 is the safe, neutral choice

k_min (minimum/mean degree)

- Use: 10–20 for most network sizes
- Too low (< 5) and sparse nodes destabilize community structure at high μ , which will introduce noise into your polarization measure exactly where you don't want it
- Too high and the graph becomes dense, compressing the dynamic range of your polarization metric
- k = 15 is the most common benchmark default — a good anchor

k_max (maximum degree)

- Use: $k_{\max} = \sqrt{n}$ as a rule of thumb, where n is your network size
- For n = 1000 $\rightarrow k_{\max} \approx 50$; for n = 5000 $\rightarrow k_{\max} \approx 70$
- This is the standard LFR convention and prevents unrealistic super-hubs
- Make sure k_max is meaningfully larger than k_min or the degree distribution collapses

Note: We mostly used this to confirm the parameter values that were also used in the literature [Lancichinetti and Fortunato, 2009; Lancichinetti *et al.*, 2008]. We tried some other values for β and γ , but in the end we used $\beta = 1.5$ and $\gamma = 2.5$ for the experiment. The answers for the values of the degrees were inconsistent and not very similar to the literature, so we did not really take them into consideration. We used a configuration using min and max degrees, not a fixed mean degree in the end.

Plotting results

Prompt: *I save my results to a pandas dataframe, structured like this: [Structure of the dataframe]. I would like some code to plot these results. I want separate plotting functions for the relational and ideological polarisation metrics, as well as for SBM and LFR results. For the SBM pipeline, I have different values for the number of communities (2, 4, and 8). I would like to have different colors for the plots for each number of communities.*

Answer: Here is the Python code using matplotlib to plot your polarization results. It contains two separate functions as requested, loops through each unique number of communities to assign a distinct color, and plots the mean value with a shaded area representing the standard deviation. [Code for plotting the results]

Note: There is not much to discuss about this. The code was mostly fine for plotting results. We had to make some minor corrections to the axes. We also changed the font sizes and colors, but otherwise we used this code for plotting results.

Formatting data into \LaTeX tables

Prompt: [CSV files pasted] *I need to split these into 3 LaTeX tables for each file. One for communities = 2, one for communities = 4, one for communities = 8. I need the following columns: ratio, relational_polarization_mean, relational_polarization_std, ideological_polarization_mean, ideological_polarization_std. Do not include trailing 0's, only keep the first 3 digits after 0 for any number after rounding. Make the table booktabs style, caption on the bottom.*

Answer: Here are the three LaTeX tables generated from the [name of csv file] dataset, split by the number of communities (2, 4, and 8). These tables follow the booktabs package styling rules, place the caption at the bottom, and format all floating-point numbers to at most 3 decimal places with any trailing zeros removed. [LaTeX code for tables 3- 8 as seen in Appendix A]

Note: We checked that the data in the generated tables matches the .csv file before introducing them in the paper. We used this for SBM networks with both N = 500 and N = 1000.

Generating \LaTeX pseudocode

Prompt: [Pasted code] *I would like to turn this code snippet into LaTeX pseudocode*

Answer: Here is the formal LaTeX pseudocode for your network evolution algorithm. I have structured it using the standard algorithm and algpseudocode packages, which are highly preferred by computational and network science journals.

The logic has been abstracted slightly from Python-specific syntax (like dictionary lookups and igraph specifics) to standard algorithmic notation for better readability in academic or formal contexts.

Note: This pseudocode was too low-level, to the point where it was too similar to actual code. We ended up having to make it more high-level, but some parts of this were still kept in. The code itself also changed a bit, so this pseudocode is not fully representative of the final version.

Algorithm 2 Evolution of Opinions and Social Structure

```

1: for  $t \leftarrow 1$  to  $T_{max}$  do
2:    $converged \leftarrow true$   $\triangleright$  Flag to track systemic equilibrium
3:   shuffle array  $A$   $\triangleright$  Randomize asynchronous activation order
4:   for each  $i \in A$  do
5:     for each  $x \in X$  do
6:       let  $N(i) = \{j \in A \mid (i, j) \in E\}$   $\triangleright$  Current social neighborhood
7:       let  $W_{ix} \in \{-1, 0, 1\}$  be the opinion of  $i$  on  $x$ 
8:       for each  $j \in N(i)$  do
9:         if  $W_{ix} \neq 0$  and  $W_{jx} = -W_{ix}$  then
10:          if  $\text{random}(0, 1) < p$  then
11:            sever social tie  $(i, j)$ 
12:            let  $C_{valid} = \{k \in A \mid k \notin N(i) \text{ and } W_{kx} \neq -W_{ix}\}$ 
13:            establish social tie  $(i, \text{random}(C_{valid}))$ 
14:             $converged \leftarrow false$ 
15:          end if
16:        end if
17:      end for
18:      recalculate  $N(i)$  to reflect any structural rewiring
19:      let  $F^+, F^-$  be the fraction of  $N(i)$  holding  $+1, -1$  respectively on  $x$ 
20:      if  $F^+ \geq \theta_i$  and  $F^+ > F^-$  then
21:        if  $W_{ix} \neq 1$  then
22:           $W_{ix} \leftarrow 1$ 
23:           $converged \leftarrow false$ 
24:        end if
25:      else if  $F^- \geq \theta_i$  and  $F^- > F^+$  then
26:        if  $W_{ix} \neq -1$  then
27:           $W_{ix} \leftarrow -1$ 
28:           $converged \leftarrow false$ 
29:        end if
30:      end if
31:    end for
32:  end for
33:  if  $converged = true$  then break  $\triangleright$  Halt if no topology or states changed
34:  end if
35: end for
36: return  $G$ 

```
