The process of polarisation as a loss of dimensionality: measuring changes in polarisation using Singular Value Decomposition of network graphs

Sage Anastasi[a], Giulio Valentino Dalla Riva[b]

[a]*Department of Mathematics and Statistics, Canterbury University, New Zealand*
[b]*Department of Mathematics and Statistics, Canterbury University, New Zealand*

Abstract

The increasing polarisation in our societies has been the focus of much research and is a major international concern. Current approaches to defining and detecting polarisation largely rely on finding evidence of bimodality in social networks or voter opinion surveys. Much of this research is based on the USA's Republican and Democrat parties, which can be difficult to apply to political systems with more than two parties or blocs. This approach also makes it hard to detect temporal trends in polarization, as the results usually fall into a binary of polarised or non-polarised; it is difficult to robustly show that subsequent increases in the bimodality of a polarised distribution are meaningful changes.

Our work is aligned with Baldassarri and Gelman's theory that polarisation should be defined as increasing correlation between positions in the ideological field, which reduces political pluralism. We also draw from post-structuralist work which argues that polarisation is the process of both the ideological fiend and the material society being segregated into two poles, as in cases of apartheid. This means that in order to measure the polarisation occurring in a society, it would be beneficial to be able to assess social networks directly.

To measure polarisation in the social network, we use Random Dot Product Graphs to embed social networks in metric spaces. In the case of a social network, the embedded dimensionality corresponds to the number of reasons that two people may form a social connection. A decrease in the optimal dimensionality for the embedding of the network graph, as measured using truncated Singular Value Decomposition of the graph adjacency matrix, is indicative of increasing polarisation in the network.

We apply this method to the communication interactions among New Zealand Twitter users discussing climate change issues, from 2017 to 2023. We find that the discussion is becoming more polarised over time, as shown by a decrease in the dimensionality of the communication network. Second, we apply this method to discussions of the COP climate change conferences, showing that our methods agree with other researchers' detection of polarisation in this space. Finally, we use networks generated by stochastic block models to explore how an increase of the isolation between distinct communities, or the increase of the predominance of one community over the other, in the communication networks are identifiable as polarisation processes.

Corresponding author: Sage Anastasi, sage.anastasi@pg.canterbury.ac.nz, University of Canterbury, Private Bag 4800, Christchurch 8140

Keywords: polarization, ideological polarization, political polarization, social complexity, climate change, random dot product graphs, graph dimensionality, singular value decomposition
2000 MSC: 62P25

---

1. Introduction

Social and political polarisation is an issue of increasing concern in the international community, since it is associated with severe social divisions and weakening of the democratic consensus [1]. Much of the research on polarisation is directed at the USA, particularly in the wake of the Trump presidency, but concern about polarisation is also present in a wide range of countries [2] [3] [4] [5].

Data-based definitions of polarisation primarily focus on the presence of a bimodal division in the distribution of ideological or social affiliations, such as the spectrum of left- to right-wing politics; bimodality is assessed by inspection (e.g. no overlap between the ideological groups)[6] or tested for significance using tests such as Hartigan's Dip test [7][8]. However, this approach can struggle when assessing multi-party democracies that do not have a clear left- and right-wing split [9]. When investigating social networks, bimodality is expressed as strong in-group/out-group divisions [10], and is often characterised by hostility between the two groups; it is common to consider this hostility as a sign of polarsation, not just the differences in policy positions between the two poles (usually the Republican and Democrat parties in the USA) [11] [12]. On social media, these hostile divisions usually take the form of "echo chambers" that focus on one set of political views and exclude all others [13] [14].

Another common axiom in current polarisation studies is that the polarisation is maximised when the groups are of equal size, as well as strongly divided [15]. This is intended to help tell polarisation conflicts apart from other major social conflicts, such as conflicts over wealth inequality. However, a consequence of this axiom is that polarisation cannot become severe if one of the groups is a small minority (e.g. refugees); given that some major social divisions are based on opposition to small minority groups, we are interested in whether there are methods for measuring polarisation that do not require this axiom.

Computational data science approaches to measuring polarisation are popular, especially for their speed when analysing millions of records from social media services. There are fast and simple algorithms for detecting communities that can then be investigated for polarisation, such as the Louvain method [16]. Network data is usually processed using a dimension reduction algorithm, such as Principal Component Analysis or Canonical Correspondence Analysis, in order to create a single dimension that can be evaluated for bimodality. Measuring latent positions empirically means projecting them in lower dimensional spaces, and then assess the resulting first dimension for bimodality [17] In other cases, clustering algorithms are more appropriate than bimodality tests; there are a number of ways of measuring distances between the clusters, which can indicate polarisation between the groups if the clusters also have low internal variation [18].

There are criticisms of the focus on bimodality in polarisation research. Baldassarri and Gelman [19] argue that focus on radicalisation relies on correctly selecting the ideological dimension that is the centre of the polarisation; incorrect selection can result in failing to find polarisation when it is present. They instead propose a definition of polarisation as increasing correlation in the ideological space, in comparison to an integrated society which "is not a society in which conflict is absent, but rather one in which conflict expresses itself through nonencompassing interests and identities". The process of increasing correlation between ideological positions reduces political pluralism and restricts possible ideological opinions, until maximum correlation is reached and an oppositional binary is created. They did not find that there was increasing correlation in the ideological field of US-American voters pre-2004. However, subsequent research by Kozlowski and Murphy [20] using the same methods found that correlation rapidly increased between

2004 and 2016. They noted that the increase in polarisation was strongest in the domains of economics and civil rights issues, rather than in the domain of moral issues that the "culture war" framing of polarisation may suggest. Similarly, DellaPosta [21] conceptualises polarisation as similar to an oil spill, with the increasing correlation in ideological positions spreading polarisation to previously apolitical members of society. The article analyses how the "belief network" of US-American politics has changed over time, concluding that the network has developed clusters which have reduced the prevalence of cross-cutting ideological positions; this means that pluralism has decreased and polarisation has increased.

Incorporating post-structuralist political theory allows us to expand on this understanding of polarisation as correlation. Ernesto Laclau & Chantal Mouffe [22] begin from the same understanding of pluralism and polarisation as scholars such as Baldassarri and Gelman, but they develop this concept beyond just they ideological field. They argue that "In a colonized country, the presence of the dominant power is every day made evident through a variety of contents: differences of dress, of language, of skin colour, of customs [...] the colonizer is discursively constructed as the anti-colonized." (p. 128). Two poles are constructed that are mutually exclusive and have nothing in common, sustained by segregation in all layers of society (e.g. the South African regime of racial apartheid). Notably, the polarisations that they examine do not occur primarily in the division between political parties, but along fault lines such as race and ethnicity, economic class, the urban-rural divide, and the division between coloniser and colonised.

In this paper we present a novel method of measuring polarisation that follows from both Laclau & Mouffe and Baldassarri & Gelman, namely using Singular Value Decomposition to determine the correlation structure of social networks. Using social networks allows us to capture interactions between people that are not explicitly political — being neighbours, sharing a workplace, etc — but which become politicised and segregated during extreme polarisation. As such, we are able to determine whether correlation is increasing not just among possible political positions, but whether it is increasing among social determinants of interaction as well. Our method gives a value that corresponds to the network's capacity for complexity. Incorporating these additional layers of a society into the analysis should make it easier to detect whether polarisation is occurring.

2. Methods

We represent the conversation happening on a social platform (Twitter, Facebook, Instagram, etc.) as a network. Each user that took part in the conversation is mapped to a node. We add an edge between two nodes if the two respective users have communicated in the time window of the observation. Depending on the chosen social platform considered and the specific research question, a communication can be given by a reply, a mention, a quote, a repost/retweet/share, or a set of these. These networks can be directed (as is more common) or undirected. Here, we consider communication networks only as unweighted graphs, although the generalization to weighted graphs doesn't present fundamental challenges.

2.1. Network modelling

Having established these networks, we model them as Random Dot Product Graphs (RDPGs) [23]. RDPGs are used instead of other graph embeddings because their optimal dimensionality is established a priori to the analysis, so it is independent of the network's size.

In the most general, directed, case under the RDPG model, each node $i \in \{1, ..., N\}$ in a graph G is associated with two vectors of traits, L_i and R_i, that give the node position in a pair of metric spaces, L and R. L_i and R_i are in general not directly observable. Then, the probability that an edge from node i to node j exists is given by the proximity of L_i and R_j, namely by the dot product

$$L_i \cdot R_j = \mathbb{P}(i \to j)\,.$$

In other words, the position of a node in L describe its outgoing edge topology, and the position of a node in R describe its incoming edge topology. In general, given the two matrices L and R, the edges are drawn with independent probabilities given by LR. We call the couple (L, R) the *RDPG embedding* of G.

In inference tasks, starting from an observed graph, the goal is to estimate the position of the nodes in the latent spaces, given the interaction structure of the network. We do not parametrise the network for this analysis. For a fixed dimension d of the two latent spaces, this is achieved by a d-truncated Singular Value Decomposition as follows.

Let A be the adjacency matrix of G. Let $A = U\Sigma V'$ be a singular value decomposition of A, so that U and V' are orthogonal matrices, and Σ is the diagonal matrix which i-th entry is the i-th singular value of A (sorted in decreasing order). Notice that in general U and V' are only identifiable up to orthogonal transformations (any rotation of them would keep the dot product constant, so they would determine the same graph). Denoting $M|_k$ the truncation of a matrix M to its first k columns, for any d, the two matrices $\hat{L} = U|_d\sqrt{\Sigma}$ and $\hat{R} = \sqrt{\Sigma}\,(V|_d)'$ determine a rank-d optimal approximations of A. That is, $\hat{L}\hat{R} = \hat{A}$ minimizes the Frobenius distance to A between all the rank-d matrices.

In the undirected case, $\hat{L} = \hat{R}$ so that $\hat{L}_i \cdot \hat{R}_j = \hat{R}_i \cdot \hat{L}_j$ and the probabilities of interaction are symmetric.

Embedding dimension

We define the *dimension* of a communication network as the *optimal* choice of d for the RDPG embedding of the network and denote it $\hat{d}$.

An a-priori optimal choice for $\hat{d}$ can be obtained from, Σ, the sorted sequence of singular values of the network's adjacency matrix A. Various methods exist. Here we adopt the elbow method presented in [24]. The elbow method identifies the most likely change point in the sequence of values of Σ by sequentially fitting two Gaussian distributions with independent mean, and equal variance. One Gaussian distribution is fitted to the largest d singular values, and the other to the smallest $K - d$. Then, $\hat{d}$ is the value of d that maximise the sum of the log-likelihoods of the two distributions.

Notice that $\hat{d}$ is robust to network size.

SVD Entropy

Given a network, we can assess its graph complexity by computing its SVD entropy, which is again based on Σ. A network has higher SVD entropy when many of its singular vectors are highly important for its structure, meaning that the network can not be efficiently compressed. This commonly read as an indication of high network complexity [25], and it is related (although not in a linear nor straightforward way) to its dimension. We normalise the SVD entropy using Pielou's evenness [26], so that the results do not depend on the network size.

In particular, let Σ be the sequence of singular values of a network's adjacency matrix A. The *nuclear norm* of A is given by the sum total of Σ (that is, the sum of all singular values). We define the normalized values $s_i = \frac{\sigma_i}{\|A\|_*}$, where $i \in \{1, ...N\}$ and σ_i is the i-th singular value.

Then, the (Pielou normalised) SVD entropy of a graph G is given by

$$J = -ln(N)^{-1} \times \sum_{i=1}^{N} s_i ln(s_i)$$

where the sum term in the definition is, indeed, an entropy.

2.2. Polarization

We define a *process of polarization* as the loss of dimensionality of a graph observed in time. Namely, we find the optimal RDPG embedding dimension $\hat{d}$ at multiple time points. If $\hat{d}$ decreases over time then we argue that the network has become more polarised during that time. This is based on the same principles as the view that the process of polarisation is one of increasing correlation. The dimensions of $\hat{d}$ are all uncorrelated; as such reduction in $\hat{d}$ corresponds to a reduction in uncorrelation in the social network graph, which is an increase in correlation.

We complement this definition by also observing the complexity of the graph, as determined by its SVD entropy, and notice whether it corresponds to an increase or decrease of polarisation.

2.3. Code

All the Social Network analysis discussed above have been performed in `Julia`[27], in particular using the packages `Graphs.jl`[28] for network manipulation, `PROPACK.jl`[29] for computing the (truncated) singular values, and DotProductGraphs.jl for computing the embedding dimension and SVD entropy.

All scripts are available at https://github.com/gvdr/Sage_data.

3. Data

We apply our computational framework for polarisation to three different data sets: two from Twitter, and one consisting of simulated interaction networks.

3.1. Climate discussion in New Zealand Twitter

We obtained 12939 tweets by querying Twitter's Academic API v2.0 for keywords related to climate change: climate, pollution, age, CO2, and carbon. We restricted our search to tweets geographically tagged as being from New Zealand, and tweets published after 2017. We then considered two time windows that corresponded to roughly equally sized networks: between 2017 and 2020, and 2020 to 2023. For each time frame, we built a network by considering each user (identified by their unique IDs) as a node, and any mention, reply, or quote tweet between two users as an edge. There were 6767 tweets in the 2017-2020 network and 6172 in the 2020-2023 network. We analyzed the two networks independently.

3.2. COP discussion in Twitter

Falkenberg et al. collected a large corpus of tweets by querying Twitter's Academic API v2.0 for tweets mentioning "COP2x" where x was an integer between 0 and 6 (inclusive) [17]. They restricted their search to tweets in English, and covered the COP from 20 to 26 (years 2014 to 2022, with 2020 and 2021 skipped due to the Covid-19 pandemic). Their adjacency matrix was constructed based on whether a user i retweeted tweets from a political influencer j. Their focus was whether there was noticeable division among users based on whether they were spreading true information about climate change or disinformation from climate change denialist influencers.

To test for polarisation using our framework, we built a network for each COP using the same data by considering each user (identified by their unique IDs) as a node, and any mention, reply, or quote tweet between two users as an edge. The network for each year was analysed independently.

3.3. Synthetic Data

We wanted to explore what happens when common axioms about polarisation are evaluated using our approach. To do this, we generated two networks using stochastic block models. The first network is used to evaluate the how changing the chances of between-groups engagement affects dimensionality of the network. The second dataset is used to evaluate what happens when the two groups in a polarisation are different sizes, to test the concept that polarisation is maximised when the two groups are equal in size and then decreases as one group becomes predominant.

Table 1: Pointwise estimates of network dimensionality and entropy.

Year	Dimension	Dimension GC	Entropy	Entropy GC
2017-2020	39.0	39.0	0.980229	0.979954
2020-2023	27.0	24.0	0.97439	0.97372

Between-groups engagement We simulated a stochastic block network of 1000 nodes, split into two equally sized groups. We varied the probability of each node forming a connection within its group and varying probabilities of in-group and between-groups linking. In particular, we simulated a networks with in-group link probability of 0.3 to 0.45 with steps of 0.05. and between-group probabilities of 0.1, 0.05, and 0.01.

Predominance of a group We simulated a stochastic block network of 1000 nodes, split into two groups. We fixed probabilities of in-group linking between 0.3 and 0.45 with steps of 0.05, and fixed the probability of between-groups linking at 0.05. We varied the sizes of the two groups, in particular, we simulated a process in which one of the two groups becomes predominant, progressively increasing its size from 0.5 of the full network to 0.2, 0.1, and finally 0.01.

4. Results

4.1. NZ Twitter Climate Change Data

We find that the dimensionality of the NZ Twitter discussion of climate change has decreased over time, indicating that it is becoming more polarised. Similarly, the von Neumann entropy of the network is lower in the 2020-2023 period than in the 2017-2020 period. This suggests that the complexity of the network is decreasing and the political positions that it is possible for users to hold are becoming narrower over time, which matches Baldassarri and Gelman's view of how polarisation works politically.

4.2. COP Data

We found that the dimensionality of the network of Twitter users discussing the COP conference has been decreasing over time, though the decrease was not linear. Unexpectedly, the von Neumann entropy of the network did not decrease in this way, and instead it was at its highest in 2022 even though the dimensionality of the network was at its lowest.

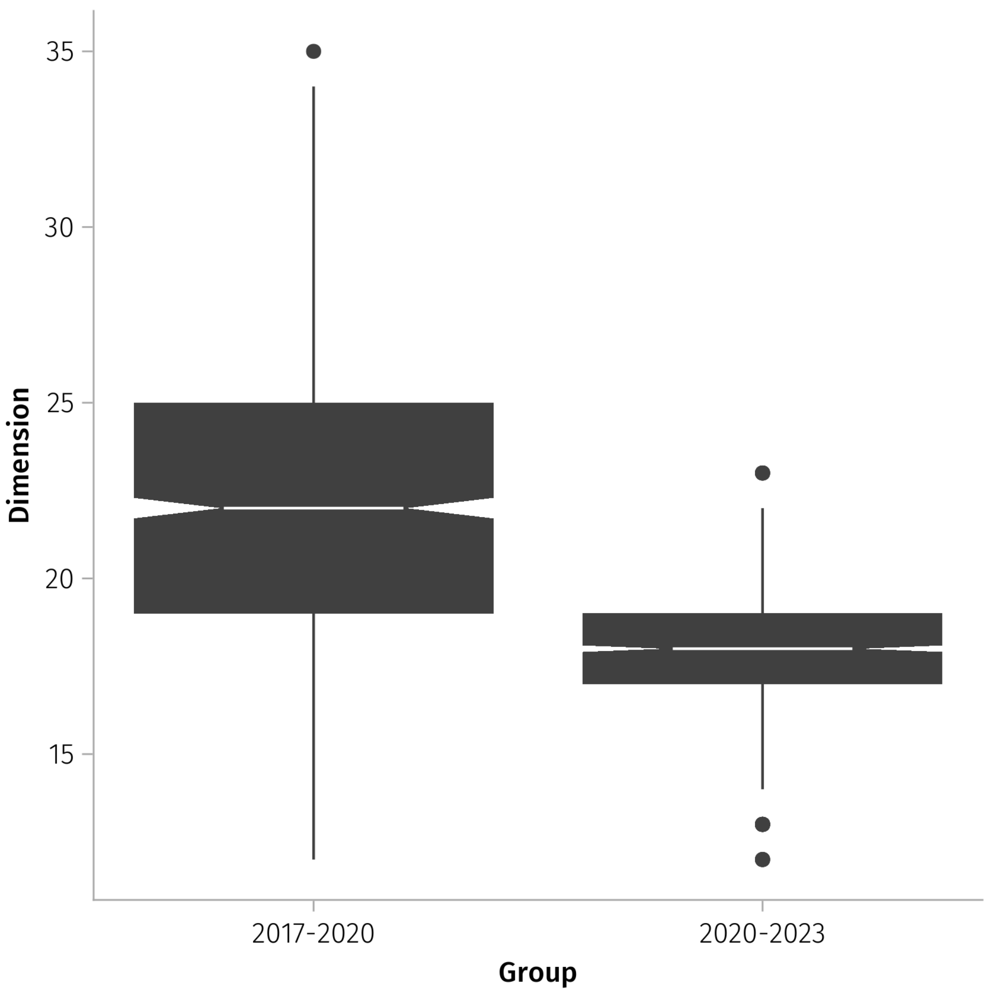


Figure 1: Plot comparing the network dimensionality of NZ climate change tweets in 2017-2020 to the dimensionality in 2020-2023.

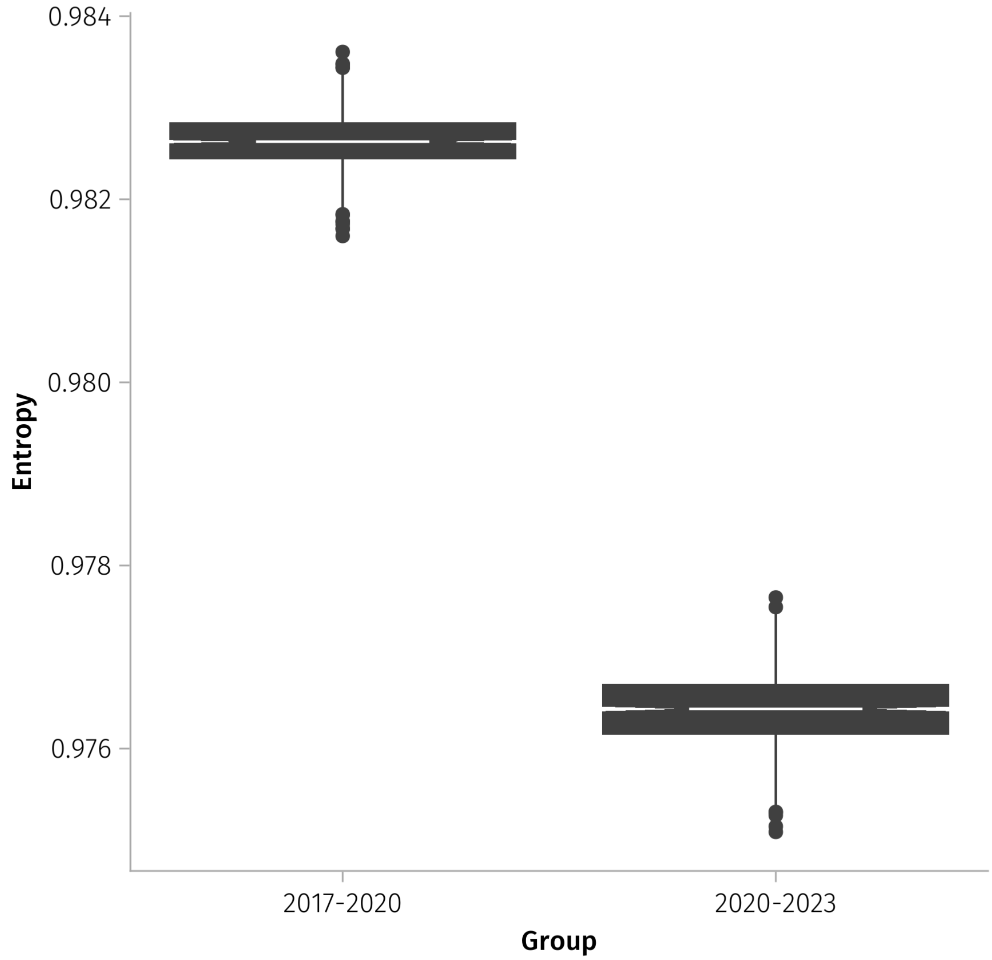
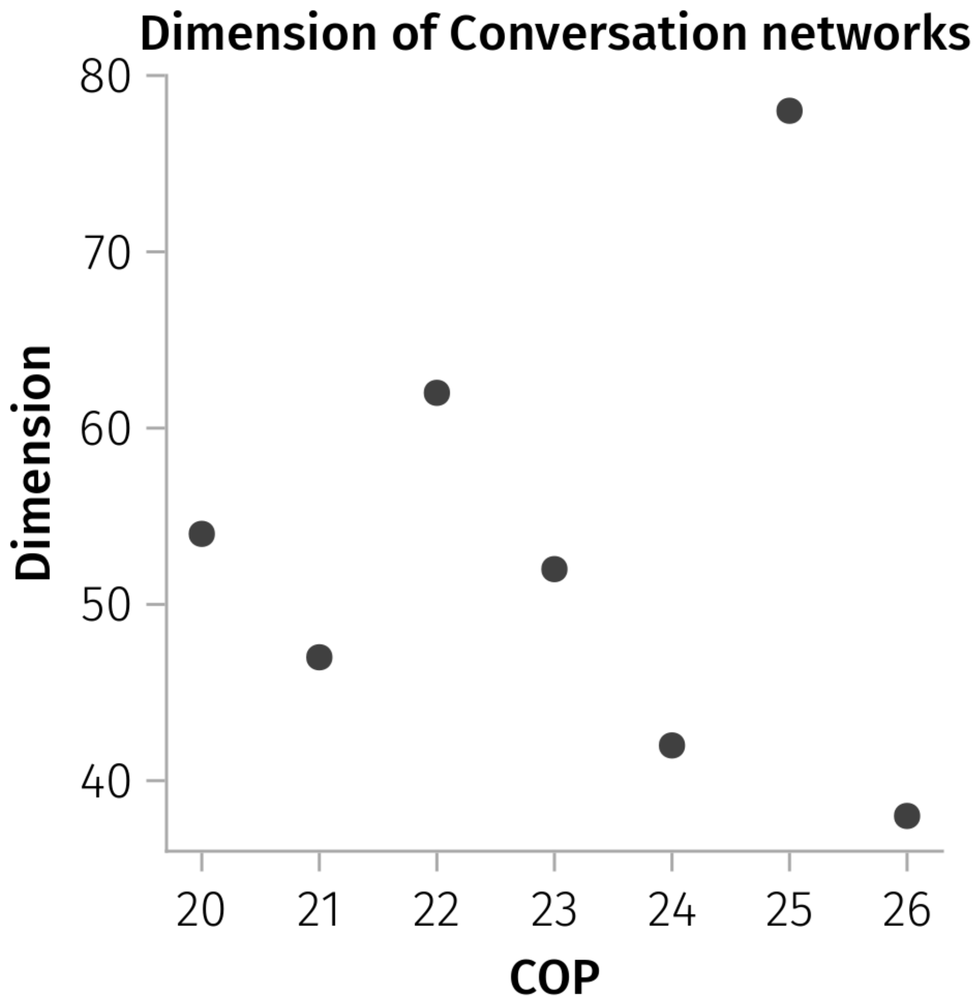


Figure 2: Plot comparing the von Neumann entropy of NZ climate change tweets in 2017-2020 to the entropy in 2020-2023.



[h]

Figure 3: Plot comparing the network dimensionality of tweets for COP20-COP26.

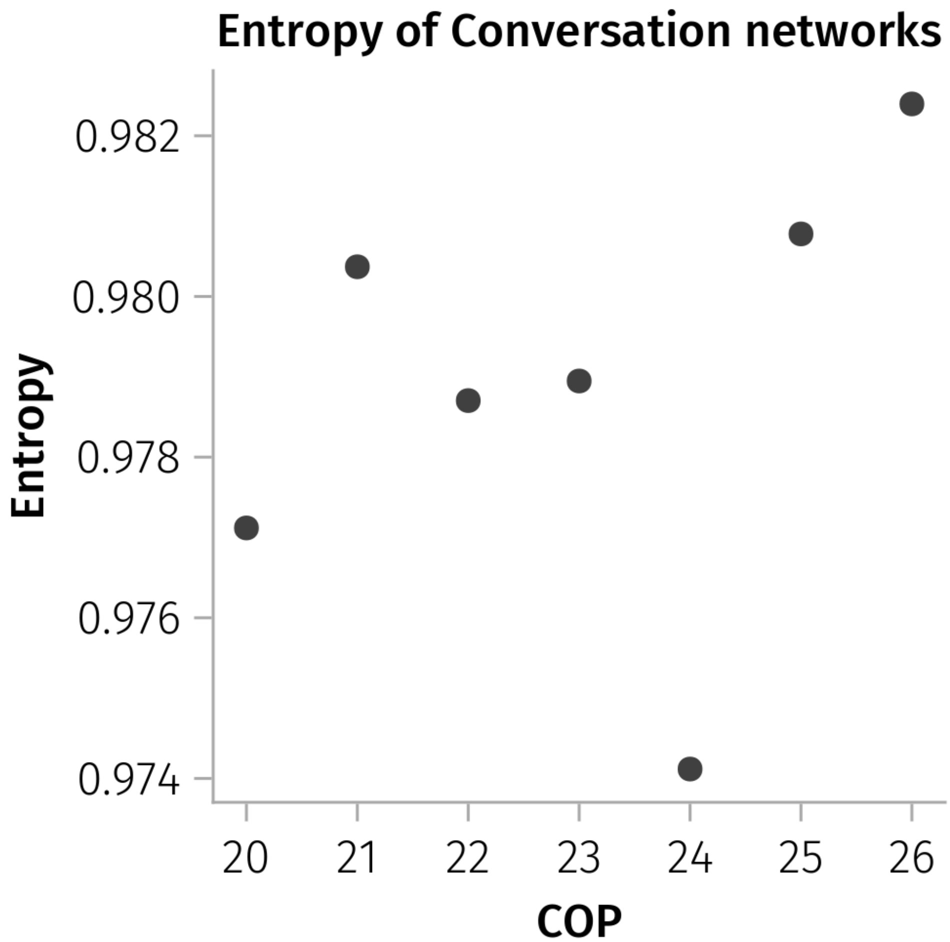


Figure 4: Plot comparing the network von Neumann entropy of tweets for COP20-COP26.

Table 2: Pointwise estimates of network dimensionality and entropy based on the first 100 SVD values.

COP	Dimension	Entropy
20	14.0	0.9802473
21	7.0	0.97777313
22	2.0	0.97607696
23	3.0	0.97573924
24	9.0	0.9748669
25	3.0	0.9791759
26	2.0	0.9754415

Table 3: Pointwise estimates of network dimensionality and entropy based on the first 1000 SVD values.

COP	Dimension	Entropy
20	54.0	0.97711855
21	47.0	0.9803695
22	62.0	0.9787037
23	52.0	0.9789476
24	42.0	0.9741172
25	78.0	0.9807784
26	38.0	0.9823967

Interestingly, Falkenberg et al. expected to find polarisation during COP21, due to the signing of the Paris Agreement at COP21. Their Hartigan's Dip Test for COP21 returned a significant result ($p = 0.003$), but they go on to claim that COP21 was not polarised despite this result. In our data, COP21 has lower dimensionality than the years before or after. It may be possible that the network as a whole became more polarised, which is captured by our data, but this effect had not yet occurred among the "influencers" that Falkenberg et al selected. Our results support Falkenberg et al's suggestion that the increase in polarisation they observed was due to an increase in the prominence of anti-climate and generally far-right influencers on Twitter, since COP26 was the conference with the lowest associated network dimensionality among Twitter users discussing it.

4.3. Simulated Data

As expected, increasing the chance of connection between the two blocks increases the dimensionality of the network (and therefore decreases the polarisation). The effect was consistent across all in-group link probabilities tested. This indicates that a potential social strategy to decrease polarsation could include facilitating the creation of connections between different groups.

For cases where the in-group link probability was lower, the von Neumann entropy decreased as the out-group link probability increased. In cases where the in-group link probability was higher, the entropy remained consistent or increased as the out-group link probability increased. As such, von Neumann entropy may be a less reliable indicator of polarisation than $\hat{d}$.

We found that the dimensionality of the network increases slightly as one group becomes predominant in the network, but decreases strongly when one group is much larger than the other. This effect was consistent across all in-group link probabilities tested. The von Neumann entropy of the network also strongly decreased when one group was much larger than the other (99 to 1), but did not exhibit the same behaviour as the dimensionality when the group was only starting to become predominant (80 to 20, and 90 to 10). At low in-group link probabilities, the von Neumann entropy decreased as one group became predominant; at higher in-group link probabilities, the entropy either remained stable or increased slowly as one group became predominant.

It is possible that our experiment did not decrease the group size far enough to trigger the effect expected by Esteban and Ray. However, it does

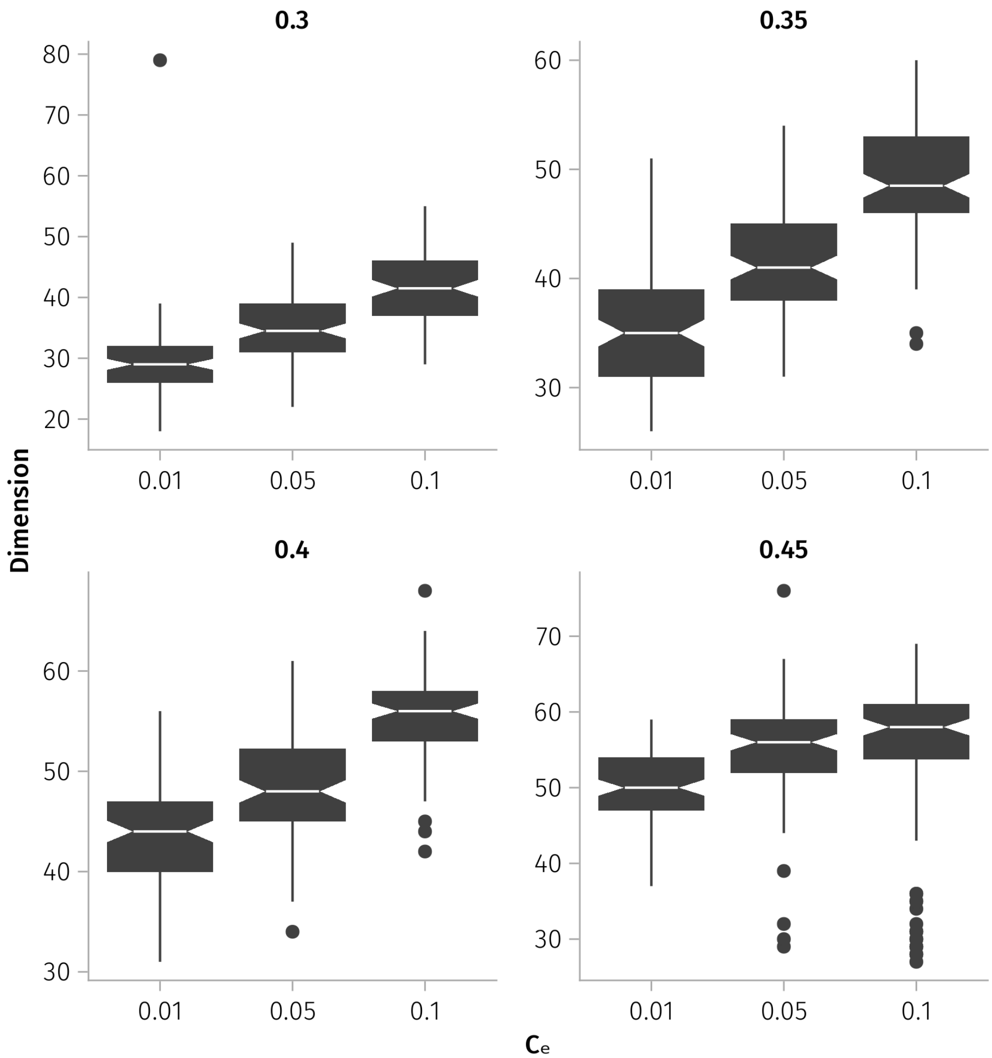


Figure 5: Network dimensionality of the stochastic blockmodel as the link probabilities are changed.

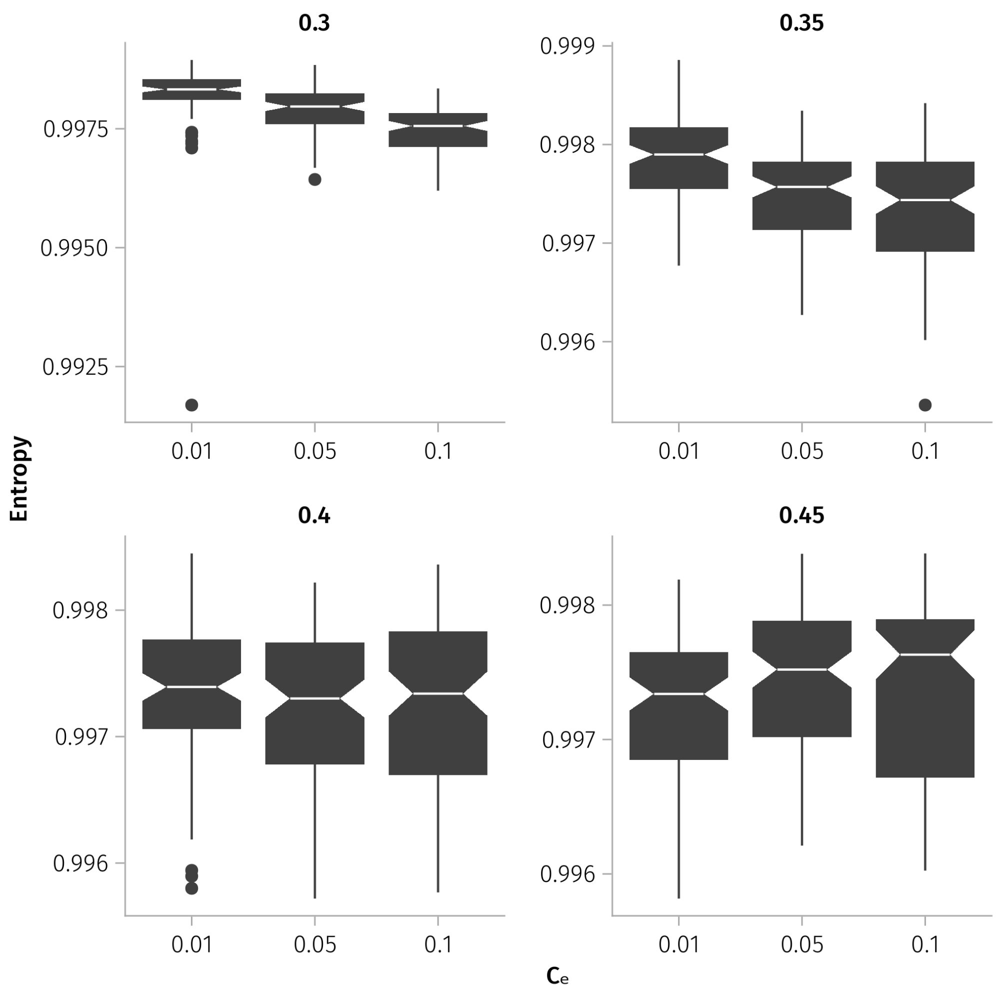


Figure 6: von Neumann entropy of the stochastic blockmodel as the link probabilities are changed.

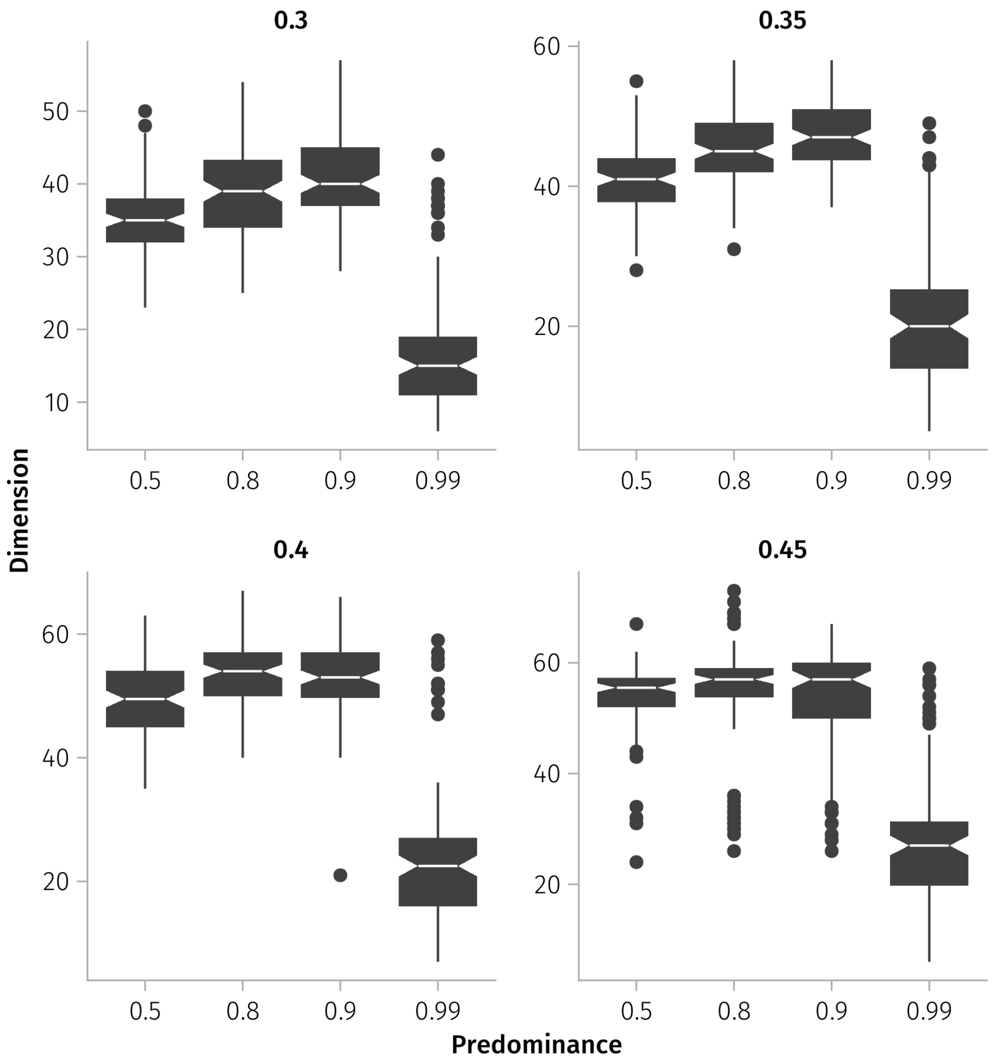


Figure 7: Network dimensionality of the stochastic blockmodel as the sizes of the blocks are changed.

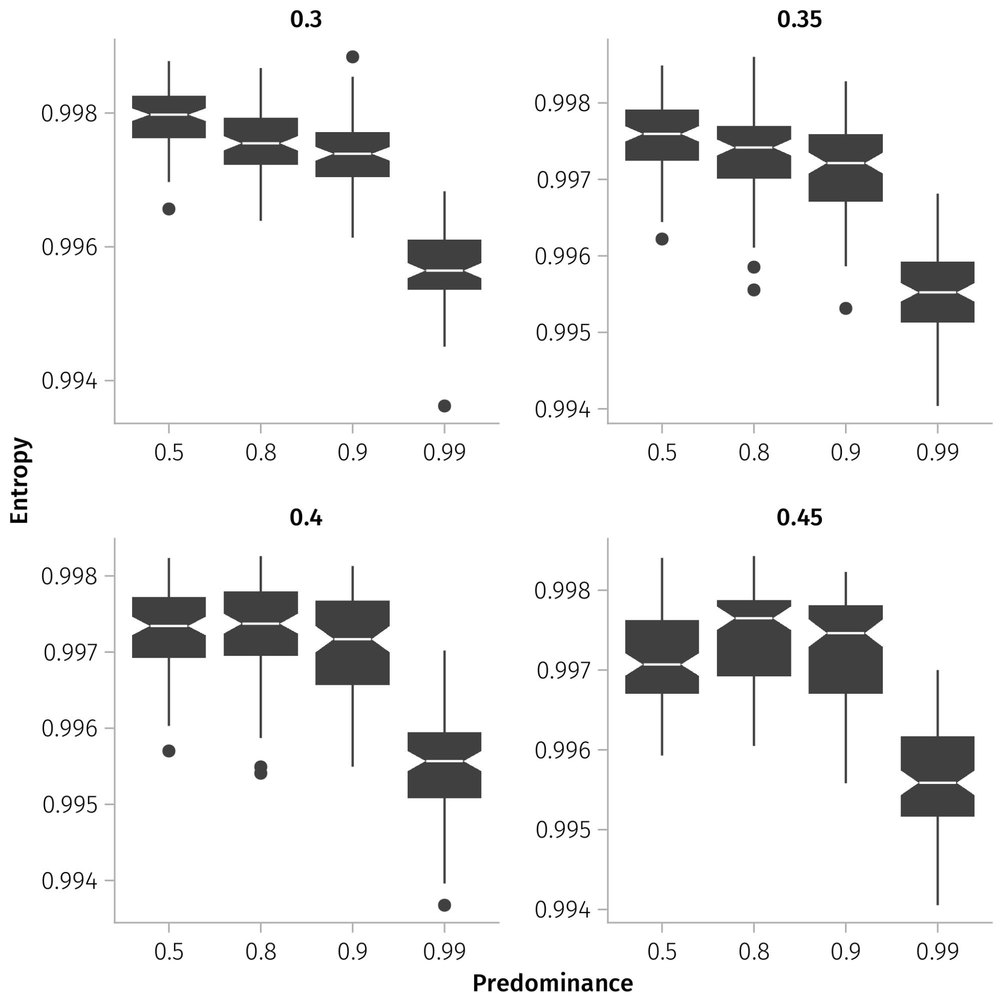


Figure 8: von Neumann entropy of the stochastic blockmodel as the sizes of the blocks are changed.

demonstrate that polarisation does not linearly decrease as one group becomes predominant, as was expected, and that the behaviour of the stochastic block model is more complicated. This is also a useful result to be aware of because groups that are the target of political polarisation are often only a small percentage of the population, such as ethnic minority or refugee populations; as such, societies where this is the case may be structurally vulnerable to polarisation.

5. Conclusions

We have demonstrated a novel method for measuring polarisation through the embedded dimensionality of random dot-product graphs. This is a reliable and straightforward implementation of the correlation-based approach to polarisation suggested by Baldassarri and Gelman. Our method captured the presence of polarisation in all the scenarios where it was expected and had been found by other researchers, in both simulated data and real social media networks. The RDPG approach also allows us to easily see that the process of polarisation is occurring in a network, through its embedded dimensionality reducing, rather than relying on a binary test of whether the network was polarised or non-polarised.

Another advantage of the RDPG approach is that it is computationally light; the main bottleneck is the computation of the first singular values of a large matrix, but this is well known in computer science literature and has already been strongly optimised. We found that the SVD was feasible even when used on networks with millions of nodes. Bimodality-based methods typically use SVD or correspondance analysis to determine the dimension they will test for bimodality, so our approach is at least as efficient.

Our approach is also highly interpretable, without forcing the latent ideological distributions into an artificially unidimensional space. Rather than creating a unidimensional space and then interpreting its political meaning (such as pro- and anti-climate, or left- and right-wing), the dimensionality method instead focuses on the number of dimensions rather than what those dimensions are. In high-dimensional spaces, we do not need to know exactly what ideologies the dimensions correspond to; the important part is that they signal that there are ideological connections being made between nodes that would not be possible if the network was polarised.

The von Neumann entropy of the network did not relate to the dimensionality as closely as we expected, though it did reflect major changes in the networks when they occurred. As such, we think it is best to use the embedded dimensionality of the network to measure its polarisation.

A major limitation of this method which could be improved is that it does not capture affective polarisation very well. Our method functionally considers any interaction between two nodes to be “good”; this means that it is not capable of capturing antagonistic interactions between nodes, and as a result it may overestimate the dimensionality of the network by mistaking brief antagonistic reactions for positive social bonds. There is a great deal of scope for integrating the concept of affective polarisation into our model, through methods such as using signed matrices and classification systems such as sentiment analysis to determine whether interactions in a social network are positive, negative, or neutral before determining the dimensionality.

Another possible extension of this method would be to implement a nonparametric two-sample hypothesis test[30], since this would allow a hypothesis test of whether the two networks are significantly different as additional evidence of polarisation having occurred. We believe that being able to observe the dimensionality of the graph alone is useful; however, we understand that sometimes a hypothesis test is demanded, and we believe this would help demonstrate that changes in the dimensionality of the network are significant.

Since we found some unexpected results when testing common axioms about polarisation, it would also be worthwhile to experiment more with the basics of the field using our method. For example, this paper has only explored the two-block case; many instances of online ”echo chambers” have a large number of groups who all hate people different to them, and it would be useful to see what happens to the dimensionality in such cases. Similarly, our testing on group prevalence showed a decrease in dimensionality when one group was 100 times the size of the other, but we did not test how the dimensionality changed as the size of the smaller group approaches zero.

5.0.1. Acknowledgements

References

[1] J. A. Tucker, A. Guess, P. Barberá, C. Vaccari, A. Siegel, S. Sanovich, D. Stukal, B. Nyhan, Social media, political polarization, and political disinformation: A review of the scientific literature, Political polarization, and political disinformation: a review of the scientific literature (March 19, 2018) (2018).

[2] J. Hebenstreit, Voter polarisation in germany: Unpolarised western but polarised eastern germany?, German Politics (2022).

[3] T. Rodon, Affective and territorial polarisation: The impact on vote choice in spain, South European Society and Politics (2022).

[4] A. R. Kozłowski, G. Krzykowski, G. Fallon, Clustering of polish citizens on the bases of their support for leaving and remaining the european union, Polish Political Science Yearbook (2023).

[5] C. Teney, L. K. Rupieper, A new social conflict on globalisation-related issues in germany? a longitudinal perspective, KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie (2023).

[6] M. J. Hetherington, Review Article: Putting Polarization in Perspective, British Journal of Political Science (2009).

[7] J. A. Hartigan, P. M. Hartigan, The dip test of unimodality, The annals of Statistics (1985).

[8] E. Kopacheva, V. Yantseva, Users' polarisation in dynamic discussion networks: The case of refugee crisis in sweden, PLOS ONE (2022).

[9] L. Röllicke, Polarisation, identity and affect - conceptualising affective polarisation in multi-party systems, Electoral Studies (2023).

[10] C. M. Valensise, M. Cinelli, W. Quattrociocchi, The dynamics of online polarization, arXiv (2022).

[11] S. Iyengar, G. Sood, Y. Lelkes, Affect, Not Ideology, Public Opinion Quarterly (2012).

[12] A. Tanesini, Affective polarisation and emotional distortions on social media, Royal Institute of Philosophy Supplements (2022).

[13] M. Del Vicario, G. Vivaldo, A. Bessi, F. Zollo, A. Scala, G. Caldarelli, W. Quattrociocchi, Echo Chambers: Emotional Contagion and Group Polarization on Facebook, Scientific Reports (2016).

[14] Y. Gao, F. Liu, L. Gao, Echo chamber effects on short video platforms, Scientific Reports (2023).

[15] J.-M. Esteban, D. Ray, On the Measurement of Polarization, Econometrica (1994).

[16] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, E. Lefebvre, Fast unfolding of communities in large networks, Journal of Statistical Mechanics: Theory and Experiment (2008).

[17] M. Falkenberg, A. Galeazzi, M. Torricelli, N. Di Marco, F. Larosa, M. Sas, A. Mekacher, W. Pearce, F. Zollo, W. Quattrociocchi, A. Baronchelli, Growing polarisation around climate change on social media, arXiv (2022).

[18] E. Schubert, Stop using the elbow criterion for k-means, arXiv (2022).

[19] D. Baldassarri, A. Gelman, Partisans without constraint: Political polarization and trends in american public opinion, American Journal of Sociology (2008).

[20] A. C. Kozlowski, J. P. Murphy, Issue alignment and partisanship in the american public: Revisiting the 'partisans without constraint' thesis, Social Science Research (2021).

[21] D. DellaPosta, Pluralistic collapse: The "oil spill" model of mass opinion polarization, American Sociological Review (2020).

[22] E. Laclau, C. Mouffe, Hegemony and socialist strategy, Verso, London, 1985.

[23] A. Athreya, D. E. Fishkind, M. Tang, C. E. Priebe, Y. Park, J. T. Vogelstein, K. Levin, V. Lyzinski, Y. Quin, Statistical inference on random dot product graphs: a survey., The Journal of Machine Learning Research (2017).

[24] M. Zhu, A. Ghodsi, Automatic dimensionality selection from the scree plot via the use of profile likelihood, Computational Statistics & Data Analysis 51 (2) (2006) 918–930.

[25] R. Gu, Y. Shao, How long the singular value decomposed entropy predicts the stock market? — evidence from the dow jones industrial average index, Physica A: Statistical Mechanics and its Applications (2016).

[26] E. Pielou, Ecological diversity, Wiley, New York, 1975.

[27] J. Bezanson, A. Edelman, S. Karpinski, V. B. Shah, Julia: A fresh approach to numerical computing, SIAM Review 59 (1) (2017) 65–98. doi:10.1137/141000671.
URL https://epubs.siam.org/doi/10.1137/141000671

[28] J. Fairbanks, M. Besançon, S. Simon, J. Hoffiman, N. Eubank, S. Karpinski, Juliagraphs/graphs.jl: an optimized graphs package for the julia programming language (2021).
URL https://github.com/JuliaGraphs/Graphs.jl/

[29] Dominique, Alexis, A. Noack, JSOBot, J. Chen, MonssafToukal, tmigot, A. S. Siqueira, J. TagBot, Juliasmoothoptimizers/propack.jl: v0.5.0 (Oct. 2022). doi:10.5281/zenodo.7150606.
URL https://doi.org/10.5281/zenodo.7150606

[30] M. Tang, A. Athreya, D. L. Sussman, V. Lyzinski, C. Priebe, A nonparametric two-sample hypothesis testing problem for random graphs, American Journal of Sociology (2017).