Quantifying Attraction to Extreme Opinions in Online Debates

Davide Perra[1], Andrea Failla[1,2], and Giulio Rossetti[2]

[1] University of Pisa, Department of Computer Science, Pisa, 56127, Italy
d.perra@studenti.unipi.it, andrea.failla@phd.unipi.it
[2] National Research Council, Institute of Information Science and Technologies "A. Faedo" (ISTI), Pisa, 56127, Italy
giulio.rossetti@isti.cnr.it

Abstract. Opinion polarization and political segregation are key societal concerns, especially on social media. Although these phenomena have been traditionally attributed to homophily — preference for like-minded individuals — recent work in social psychology suggests that acrophily — preference for extreme rather than moderate opinions — might play a role as well. In this work, we introduce a methodology to estimate the degree of preference for connecting with users who hold strong opinions on social media. Our framework is composed of four phases: (i) opinion estimation, (ii) opinion thresholding, (iii) network construction, and (iv) acrophily estimation. We apply it to study the climate change debate on Reddit and find that users show higher-than-expected acrophilic patterns, especially if they are climate skeptics or have extreme opinions. Acrophilic patterns are stable over time, while polarization gradually leaves space for pluralism.

Keywords: acrophily · opinion polarization · climate change · social media · social networks

1 Introduction

Online Social Platforms, such as X/Twitter and Reddit, are digital environments where individuals freely manifest their opinions and typically engage in discussions with peers. When it comes to debating hot-button issues such as climate change [23], minority discrimination [9], and abortion rights [2], users self-organize in opposed opinion-driven subgroups called Echo Chambers [17, 2]. In these settings, users' beliefs are reinforced by exposure to like-minded peers, and opposing views are instead discredited. As a consequence, opportunities for change are scarce, and opinions might radicalize over time [12]. Human biases such as homophily [8] – preference for connecting with similar others – were shown to play a pivotal role in these dynamics. Moreover, recent work in social psychology has unveiled that acrophily – preference for like-minded peers with more extreme opinions (as opposed to more moderate) – also influences tie and opinion formation mechanisms [5], and that the combined effects of homophily and acrophily ultimately harshen polarization effects [6]. Therefore, measuring these effects becomes of utmost importance to understand (and consequently, reduce) the effects of these warning phenomena. Nonetheless, acrophily is a rather recent concept, and literature lacks a methodology to measure it in the wild. To bridge this gap, we introduce a framework to model, measure, and interpret attraction to extreme opinions in online debates. Our framework is flexible, and can be applied to discussions on any social media platform, as it leverages features that are common to virtually all platforms. In this work, we apply it to Reddit discussions on climate change and study the evolution of polarization and homophilic/acrophilic effects across time. Notably, this is the first research effort providing a standardized methodology to quantify acrophily in online social networks, and also to study its temporal evolution.

The remainder of this work is organized as follows. In section 2, we provide an overview of related literature; section 3 introduces the methodology and details its steps; subsequently, section 4 describes the data collection process, and section 5 introduces and discusses the experiments; finally, section 6 concludes the work and suggests future research directions.

2 Related Work

This section discusses relevant literature related to the present work. Specifically, it focuses on acrophily and homophily, human biases that were claimed to drive social interactions online.

Acrophily. Acrophily is a social tie formation mechanism first introduced in [5], where it is defined as the tendency to affiliate with more extreme (as opposed to more moderate) like-minded peers. In this study, participants were asked to rate their emotions toward political policies in a controlled setting. After viewing peers' responses to the same policies, they selected which peers they wanted to affiliate with. The findings indicate that participants' choices were influenced by both homophily and acrophily. Moreover, the inclination toward acrophily was linked to the perception that more extreme expressions are seen as more representative of one's political group. More recent work has studied the presence of acrophily on Twitter (now X) [6]. The authors associated users with political leaning estimated from media consumption and then compared the probability of retweeting extreme users with controlled simulations. The findings suggest that liberals' social ties may be driven by both homophily and acrophily, while conservatives show extreme acrophilic preferences. Authors hint that research should move toward exploring acrophily on other platforms as well, a suggestion we address in the present work.

Homophily. Homophily is the principle that *similarity breeds connection* [8], and describes the tendency to associate with peers sharing similar tastes, beliefs and/or demographics. This phenomenon was observed in a variety of online and offline settings [10, 1, 18]. In the online domain, it was linked to the rise of (especially political) polarization [2, 17, 20]. Several methods exist to quantify homophilic tendencies in social networks. Newman's assortativity coefficient is the most popular measure, and quantifies attribute correlation between adjacent node pairs [10]. More recent work [13, 15] has moved to study local patterns of homophily, to address the fact that, in large networks, homophilic and heterophilic wiring patterns may coexist. Attribute-driven wiring patterns were studied among higher-order interactions as well, by defining measures over simplicial complexes and hypergraphs [16, 22].

3 Methodology

In this section, we introduce our analytical pipeline for estimating the preference for extreme opinions in online debates. Of the two works on acrophily to date, only one studies this phenomenon on social media data [6]. While relying on a sound methodology, and yielding interesting results, acrophily is studied on retweet networks, a choice that makes it impossible to generalize their strategy to other social platforms like Reddit. Instead, we tackle this issue by leveraging information on social interactions and semantics extracted from user-generated content – elements shared by any social platform. Our framework is composed of the following four phases: (i) Opinion estimation, (ii) Opinion thresholding, (iii) Network construction, and (iv) Acrophily estimation. These steps are described in detail below.

Opinion Estimation. This step deals with the task of identifying user ideology with respect to a controversial issue. Specifically, the aim is to understand (i) which of the two stances the user agrees with, and (ii) how strong their opinion is. Estimating user opinions is a challenging task because most users do not necessarily associate themselves with an explicit label, e.g., political ideology. Moreover, estimating the *intensity* of such opinion is paramount difficult. We model the problem of jointly predicting stance and opinion intensity as a text classification task. Informally, a classifier learns to discern between two stances (e.g., democrats vs. conservatives) based on annotated examples of each ideology (e.g., posts by prominent democrats/conservatives). The assumption here is that an individual's opinion is conveyed by the posts he/she produces within the debate, and that like-minded individuals share similar language to an extent.
As a proxy of opinion intensity, we use the model's class probability, a value in [0,1] identifying the model's confidence level on a particular prediction. Here, values closer to 1 reflect a stronger belief in one stance, and values closer to 0 indicate a stronger belief in the other. The intuition behind this choice is that strong opinions emerge more clearly from text, and thus are assigned scores closer to 0 and 1 than moderate opinions. After obtaining scores for each post, we compute the average score over a predefined time period (e.g., one month, semester, year, etc.) and assign the results to each user. Lastly, we rescale and

center these values on zero by applying the following transformation:

$$y = 2x - 1, \tag{1}$$

where x is the average model score, and y is the resulting opinion value. After applying this function, the opinion sign is indicative of the stance (pro/against), while the (absolute) value is indicative of the opinion strength. For instance, in an abortion debate, stances may hold values in [-1, 1] where 1 represents a strong "pro-abortion rights" ideology, -1 represents a strong "pro-life/anti-abortion" ideology, and 0 implies neutrality.

Opinion Thresholding. After estimating user ideology, a key step of this pipeline involves distinguishing between extreme vs. moderate opinion holders. This typically involves setting a threshold, so that opinions falling beyond it are considered strong, and otherwise are considered moderate. The choice of an appropriate threshold strictly depends on the data, and specifically on the platform and discussion topic. We suggest choosing thresholds depending on the opinion distribution and the topic of interest. Namely, on a controversial issue, one would expect a considerable fraction of debaters to have extreme opinions. However, other analysts might choose other approaches, such as setting symmetric thresholds (e.g., 0.7 and -0.7), or relying on external data such as survey data.

Network Construction. In order to study connections to users with extreme opinions, we must first build a model representing their social ties. Social topology on Online Social Networks typically comes in multiple layers [3]. Indeed, on most platforms, users can follow/befriend each other, and comment, share and like other people's content. Since we are interested in the dynamics of online debates, we will focus on networks emerging from conversations. We model an online debate as a directed graph $\mathcal{G} = (V, E)$, where V is the set of nodes representing social media users, and E is the set of edges that encode conversations between such users (e.g. comment/reply interactions). Each node is further enriched with information on its stance toward a specific topic on a binary spectrum (the result of the first step), and on whether this stance is extreme or not (the result of the second step).

Acrophily Estimation. This phase involves computing measures on the structure and semantics obtained from previous steps. To measure attraction to extreme opinions, we rely on the concept of acrophily [6, 5] and formalize it as follows.

Definition 1 (Acrophily Index). *Let $G = (V, E)$ be a node-attributed graph where V is the set of nodes, and E is the set of edges. Let each node $v \in V$ be enriched by a tuple (p_v, q_v) where p_v is a binary value describing user ideology among two opposing factions, and q_v is a binary value indicating whether user ideology is extreme (1) or moderate (0).*

Let S_v identify the set of successors of the node v, and $S_v^{same} \subset S_v$ its subset of nodes adhering to the same faction, namely, all successors s for which $p_v == p_s$ is true. Let M_v^{same} be the subset of S_v^{same} containing nodes with *extreme ideologies, i.e., all successors* $s \in S_v^{same}$ *for which* $q_s == 1$ *is true. The Acrophily Index for node* v *is computed as:*

$$Acrophily(v) = \frac{|M_v^{same}|}{|S_v^{same}|}. \tag{2}$$

This function returns a value in [0, 1], where 1 indicates strong acrophilic behavior (preference for like-minded peers with more extreme ideology), and 0 indicates the absence of such behavior. The measure is undefined for $|S_v^{same}| = 0$, and thus is computed only for nodes having at least one like-minded successor. The Acrophily Index above can be generalized to the whole graph by averaging over the number of non-null values.

4 Data Collection

We apply our framework to Reddit debates on climate change. In this section, we detail the data collection and preparation processes.
In this phase, our aim is to find text content that accurately represents each of the two opposing stances in the climate change debate, namely climate skeptics vs. those who believe climate change is genuine. To do so, we identify a collection of subreddits where discussions on this issue take place, and retrieve all their submissions[3] and comments missed from 2019 to 2022 via the Pushshift API[4]. Since we want to capture the controversial nature of the debate, we make sure to include subreddits that clearly relate to each of the ideological positions (e.g., r/climateskeptics, r/ClimateActionPlan). We rely on subreddit names and descriptions to identify relevant subreddits. Moreover, we sample 50 random submissions from each of the considered subreddits and manually verify adherence to the topic. Since a previous study suggests that many debates on climate change occur in general-purpose settings/subreddits [19], we also include subreddits with a broader scope (e.g., r/TrueAskReddit). In these cases, we filter out submissions that do not contain the word "climate" to ensure adherence to the topic of interest. This process resulted in 40,872 submissions and 661,024 comments. The final dataset contains both textual information — which is used to estimate the user ideology — and relational information — which is used to build the interaction network.

5 Attraction to Extreme Opinions in the Climate Change Debate

In the following, we apply our methodology to Reddit data on climate change debates. Our aim is twofold. First, we want to discover whether users show signs

[3] Reddit distinguishes between a ”submission”, i.e., a post that initiates a discussion, and a ”comment”, i.e., a post that replies to a submission or to another comment
[4] https://api.pushshift.io/

of attraction to extreme opinions, and whether this phenomenon occurs at a higher rate than expected. Then, we investigate whether opinions in climate change debates are polarized and how polarization evolves over time. All data produced during this study is anonymized and released in a dedicated Zenodo repository along with the code to reproduce the experiments (see [14]).

5.1 Opinion Estimation

In order to estimate user opinions on the climate change issue, we rely on BERT, a neural language model based on a bidirectional transformer architecture [21]. We finetune the BERT base uncased model on annotated examples extracted from our dataset. Specifically, we use submissions on stance-specific subreddits (e.g., r/climateskeptics), where users argue on their own view of the topic. We use submissions because they are usually longer pieces of text where opinions emerge more clearly. The filtered dataset contains 19K posts by climate skeptics — to which we assign label `'1'`, and 28K by supporters, to which we assign label `'0'`. Text is then processed via a standard pipeline, including lowercasing, removing non-printable characters, XSLT tags, URLs, numbers, punctuation, extra spaces, and English stopwords. Then, we remove posts with less than 15 characters to ensure a minimum level of informativity. This leaves 5K skeptics submissions, and 9K supporters submissions. This data is split into train (70%), validation (10%), and test (20%) sets, and the former two are fed to the BERT model for the fine-tuning phase.

We experiment with several configurations of the batch size, number of epochs, and input length. The batch size refers to the number of training examples that are used by the model during a single weights update; the number of epochs is the amount of complete iterations across the entire training dataset throughout the learning process; finally, the input length is the maximum number of tokens allowed, after which the input is truncated. In the best configuration, we set these parameters at 22, 2, and 155, respectively. Performance metrics for the

Table 1: BERT Performance Metrics

Metric	Training	Validation	Test
Accuracy	0.954	0.888	0.818
Precision	0.957	0.899	0.848
Recall	0.963	0.906	0.883
F1 Score	0.960	0.903	0.865

best model performances are shown in Table 1. The model is able to learn the patterns in the training data well, as all metrics are beyond 0.95 in this phase. Validation statistics are slightly lower. This indicates that the model generalizes well but is slightly less accurate on unseen data compared to the training data.

The difference between training and validation metrics suggests a small degree of overfitting. The model's performance further decreases when applied to completely unseen data, which is expected. Still, the metrics indicate that the model is suitable to be applied to new contexts as it reaches a respectable F1 score of over 0.85. Fig. 1 displays a confusion matrix with the counts of true/false positives/negatives obtained on the test set. The confusion matrix does not suggest significant bias, given the good performance on both classes, and similar magnitudes of misclassification class-wise.

As a last step, we apply the model to unlabeled data from discussions on climate change in general subreddits. Since we want to investigate whether preference for extreme opinions changes over time, we partition the 2022 data by quarters. Moreover, we average the model's class probabilities for each user across each trimester. Finally, each value is rescaled and centered at 0 via Eq. 1 to obtain the final opinion scores. In summary, the output of this step is, for each user discussing climate change in a non-specialized scenario (e.g. general discussion of news or politics), a score that quantifies (i) their stance on climate change in a given quarter of 2022, and (ii) the intensity of this belief.

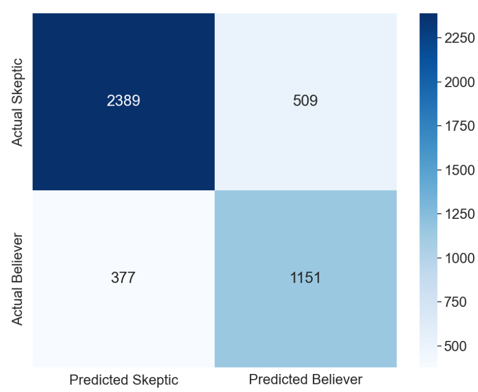


Fig. 1: Confusion Matrix of BERT performance on the test set

5.2 Opinion Thresholding

Fig. 2 displays the Kernel Density Estimations of opinion distributions across the four timesteps. Opinions show bimodal distributions with peaks nearing extreme values, which is a typical sign of polarization [7, 9]. However, as highlighted in the previous section, selecting appropriate thresholds that separate extreme from moderate users is a challenging task. Ideally, in a polarized scenario, we would expect to find several users with extreme opinions. We study the impact that cutting at different percentiles on both ends has on the volume of extreme users. We study the following percentiles: 5th, 10th, 15th, 20th, 25th for the left threshold; 65th, 70th, 75th, 80th, 85th, 90th, 95th percentiles for the right threshold. We choose to cut (i) at the percentile immediately after the peak on the left and (ii) at the percentile immediately before the peak on the right. The left and right thresholds for each snapshot are represented in Fig. 2 as red dashed lines, and their values are reported in grey.

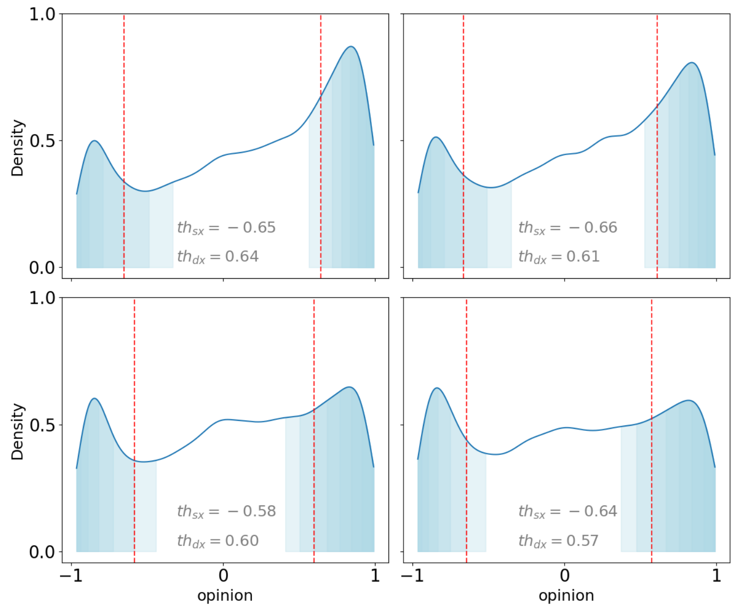


Fig. 2: Opinion distribution at each time step. Blue shades outline different percentiles, with darker ones indicating more restrictive thresholds. Dashed lines outline the thresholds used in this work.

5.3 Network Construction

To account for the temporal evolution of user ideology and social topology, we build four networks from data, each encompassing three months of online interactions. Nodes are Reddit users, and directed edges refer to replies/comments to submissions or comments in the studied subreddits. Information on who replies to who is available in the Reddit post metadata. In accordance with the previous section, we enrich the nodes with information on (i) their estimated faction at that time step and (ii) whether their opinion is extreme or not. Summary statistics are reported in Table 2.

Table 2: Network statistics for each snapshot. $|V|$ and $|E|$ refer to the number of nodes and edges respectively. k_{avg} refers to the average degree. C_{max} refers to the number of nodes in the largest weakly connected component.

Quarter	$\|V\|$	$\|E\|$	k_{avg}	$\|C_{max}\|$	#supporters	#skeptics
Q1	23,253	57,128	4.91	21,069	8753	14,500
Q2	28,426	78,913	5.55	27,373	11,049	17,377
Q3	30,048	87,749	5.84	29,443	13,203	16,845
Q4	19,854	51,964	5.23	19,358	9,358	10,496

5.4 Acrophily Estimation

In the following, we compute Acrophily for all networks via Eq. 2. KDE plots are shown in Fig. 3. Distributions of acrophily values are almost identical across all observation periods. Most users show an absence of acrophilic behaviors, as evident from the peak on the left. Still, some peaks can be observed at around 0.5 and 1, highlighting a considerable presence of acrophilic users as well. Overall, 25 to 30% of users in each timestamp show strong acrophilic tendencies ($>$ 0.5).

As shown in Figure 4 (a), acrophily shows some variation across time, although minor. In general, however, we find that (i) skeptical users show higher acrophily than supporters (with the exception of Q3, where values are comparable), and (ii) extreme users are more acrophilic than moderates. To understand whether this debate exhibits more acrophily than expected at random, we compare average acrophily values with null models. Specifically, we compare each value with those obtained from a directed configuration model [11] (DCM), which shuffles connections while preserving in and out-degree distributions. Each node in the DCM is assigned an opinion value drawn from the empirical opinion distribution of the corresponding time period. This process ensures that connections in the null model are random, while taking into account both the structural properties of the system and the opinion distribution. Figure 4 (b) shows the ratios of real over expected acrophily values for different categories. In all timestamps, the climate change debaters show higher-than-expected acrophilic tendencies. This is true both on a global scale, where our data is up to 1.4 times more acrophilic than the DCM, and class-wise. Again, extreme users and skeptical users show the highest ratios with respect to the DCM (up to 1.6 and up to 1.7, respectively), suggesting that climate change debates on Reddit are particularly heated and at risk of polarization.

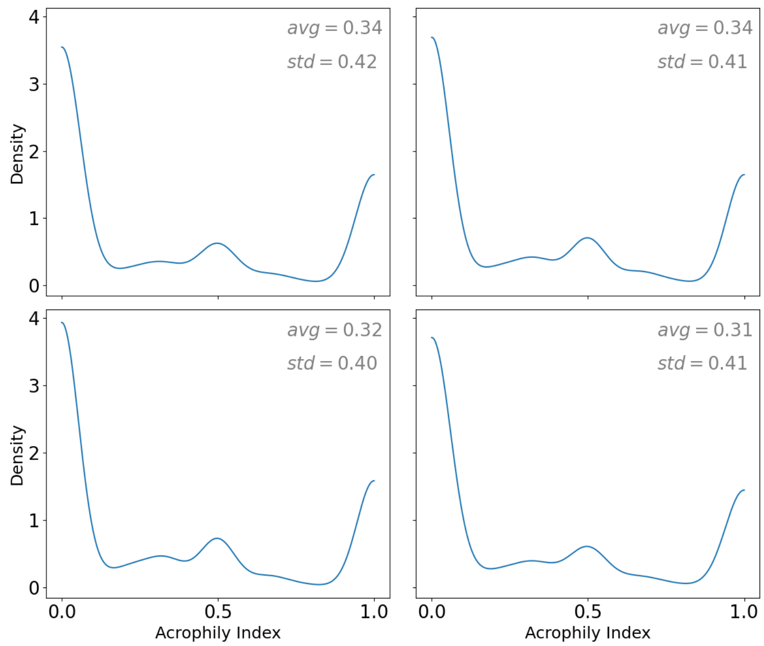
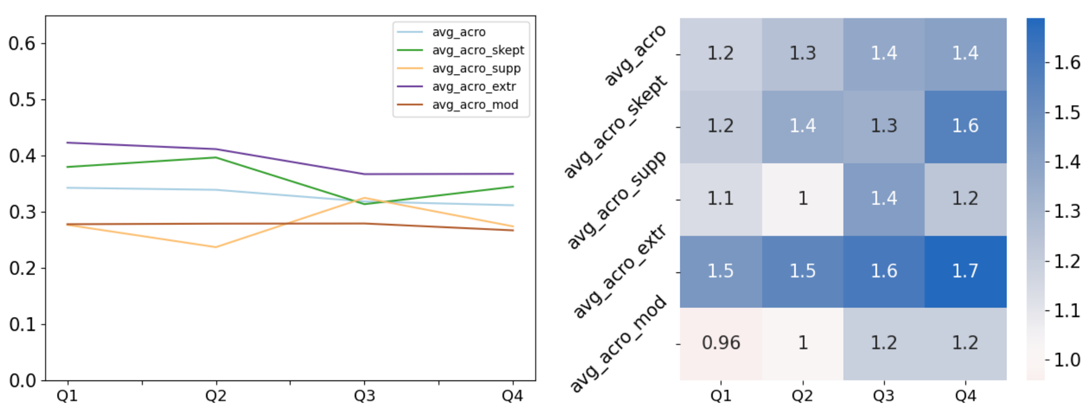


Fig. 3: KDE of acrophily values at each timestep. Mean and standard deviation values are reported in the top-right corner.



(a) (b)

Fig. 4: Temporal trends of acrophily (a) and ratios over the directed configuration model (b)

To assess whether users are embedded in local echo chambers, we compare their opinions with those of the users they actively interact with [20]. Fig. 5 displays this comparison across the four snapshots. We find that users are typically surrounded by peers with similar opinions (both in stance and magnitude), and especially so for extreme climate skeptics. Howvever, pluralism emerges as time goes by (as denoted by the brighter areas now aligning horizontally, see [20]), especially in the second half of the year, and opinions become more moderate. Indeed, the Pearson coefficient for user opinion and average successors' opinion lies around 0.4 in the first two quarters, but goes down to around 0.2 afterwards. Note that all coefficients are statistically significant ($pvalue < 0.01$). Interestingly, the highest pluralism is observed in the last quarter, concurrently with the United Nations Climate Change Conference (more commonly, COP27). We hypothesize that relevant events in the second part of 2022 — such as COP27 — might have induced Reddit users to engage with users with opposing views, with the aim of convincing and/or criticizing.

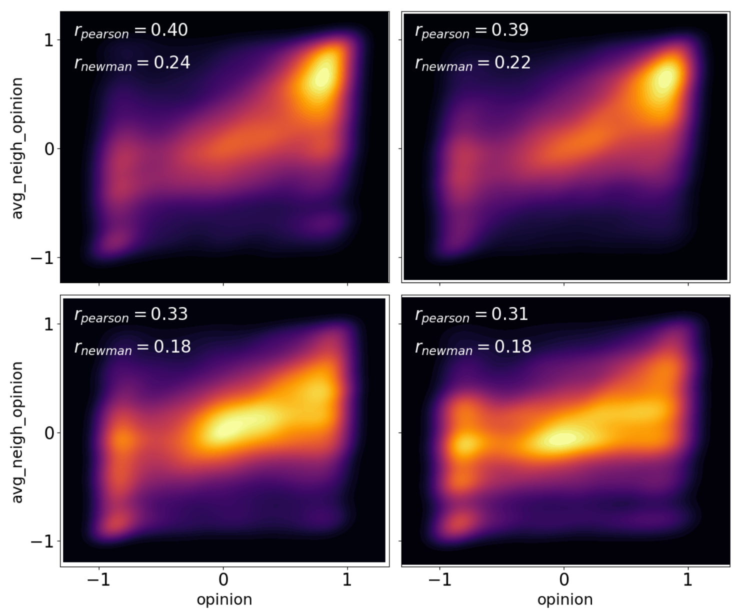


Fig. 5: Contour plot of user opinion (x-axis) against average opinions of users they interact with (y-axis). $r_{pearson}$ refers to the Pearson correlation coefficient between the two variables. r_{newman} refers to the Newman categorical assortativity coefficient for the stance label.

6 Conclusion

With the aim of better understanding polluting online dynamics, we have introduced a descriptive framework to measure attraction to extreme opinions within debates on social platforms. Contrary to previous work on acrophily – which relied on surveys and/or retweet-based Twitter simulations – our framework is scalable and platform-independent. This is due to the fact that it relies strictly on user-generated content and conversation structure, which are present in most – if not all – social networks. We apply our framework to Reddit debates on climate change, and study the evolution of acrophily during 2022. In our analysis, we discover varying degrees of acrophily. Moreover, we find that extreme opinion holders and climate skeptics are the categories that show the highest acrophilic tendencies, and do so at higher-than-expected rates. Users also show a degree of opinion polarization, especially in the first half of the year. Subsequently, polarization leaves space for pluralism, possibly tied to relevant real-world events.

Limitations. Our work is not free of limitations. First of all, due to the lack of ground truth data, there is no way to evaluate results in a sound manner. This is a common issue of data-driven works that estimate user opinions [9], which emphasizes the need for open, topic-specific datasets with this kind of information. Moreover, we use class probabilities of a deep learning model to estimate opinions, and leverage extreme values (i.e., close to 1 or 0) to distinguish between extreme vs. moderate stances. We do this assuming that strong opinions emerge more clearly from the text. Although reasonable, this might not always be true. Still, we believe it is the most systematic way of doing so at scale. Relatedly, to gather enough data for a reasonable user opinion classification we use large window sizes (three months per snapshot), which might overlook within-quarter opinion changes. Moreover, a small percentage of users were removed from the dataset becausetheir posts were either deleted or not sufficiently long/informative. This, of course, has an impact on the observed network topology, although minor. Finally, the validation process is based on a comparison of average acrophily values, which does not take into account the variability in acrophily across different network realisations generated by the null model. We also stress that our findings are strictly relative to the data and should not be generalized to larger populations without further statistical analyses.

Future research. Future works might attempt to tackle the issues discussed above and/or apply this framework to other controversial debates, such as those related to vaccination, abortion rights, and more. Other platforms should be investigated as well, and especially recent/understudied ones such as the Twitter-like Bluesky [4] and Mastodon [24]. This would allow to understand whether different algorithmic choices and platform affordances may impact observed acrophilic tendencies. Finally, we plan to devise new measures to estimate acrophily that account for chance and class size imbalance, and to further study the combined effects of acrophily with other biases.

Acknowledgments. This work is supported by (i) the European Union – Horizon 2020 Program under the scheme "INFRAIA-01-2018-2019 – Integrating Activities for Advanced Communities", Grant Agreement n.871042, "SoBigData++: European Integrated Infrastructure for Social Mining and Big Data Analytics" (http://www.sobigdata.eu); (ii) SoBigData.it which receives funding from the European Union – NextGenerationEU – National Recovery and Resilience Plan (Piano Nazionale di Ripresa e Resilienza, PNRR) – Project: "SoBigData.it – Strengthening the Italian RI for Social Mining and Big Data Analytics" – Prot. IR0000013 – Avviso n. 3264 del 28/12/2021; (iii) EU NextGenerationEU programme under the funding schemes PNRR-PE-AI FAIR (Future Artificial Intelligence Research).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Barberá, P., Jost, J.T., Nagler, J., Tucker, J.A., Bonneau, R.: Tweeting from left to right: Is online political communication more than an echo chamber? Psychological science **26**(10), 1531–1542 (2015)
2. Cinelli, M., Morales, G.D.F., Galeazzi, A., Quattrociocchi, W., Starnini, M.: Echo chambers on social media: A comparative analysis. arXiv preprint arXiv:2004.09603 (2020)
3. Dickison, M.E., Magnani, M., Rossi, L.: Multilayer social networks. Cambridge University Press (2016)
4. Failla, A., Rossetti, G.: " i'm in the bluesky tonight": Insights from a year worth of social data. arXiv preprint arXiv:2404.18984 (2024)
5. Goldenberg, A., Abruzzo, J.M., Huang, Z., Schöne, J., Bailey, D., Willer, R., Halperin, E., Gross, J.J.: Homophily and acrophily as drivers of political segregation. Nature Human Behaviour **7**(2), 219–230 (2023)
6. Goldenberg, A., Bailey, D., Muric, G., Ferrara, E., Schöne, J., Willer, R., Halperin, E., Gross, J.: Attraction to politically extreme users on social media (2023)
7. Liu, J., Huang, S., Aden, N.M., Johnson, N.F., Song, C.: Emergence of polarization in coevolving networks. Physical Review Letters **130**(3), 037401 (2023)
8. McPherson, M., Smith-Lovin, L., Cook, J.M.: Birds of a feather: Homophily in social networks. Annual review of sociology **27**(1), 415–444 (2001)
9. Morini, V., Pollacci, L., Rossetti, G.: Toward a standard approach for echo chamber detection: Reddit case study. Applied Sciences **11**(12), 5390 (2021)
10. Newman, M.E.: Mixing patterns in networks. Physical review E **67**(2), 026126 (2003)
11. Newman, M.E., Strogatz, S.H., Watts, D.J.: Random graphs with arbitrary degree distributions and their applications. Physical review E **64**(2), 026118 (2001)
12. Nguyen, C.T.: Echo chambers and epistemic bubbles. Episteme **17**(2), 141–161 (2020)
13. Peel, L., Delvenne, J.C., Lambiotte, R.: Multiscale mixing patterns in networks. Proceedings of the National Academy of Sciences **115**(16), 4057–4062 (2018)
14. Perra, D., Failla, A., Rossetti, G.: Reddit climate change debate dataset (2024). https://doi.org/https://doi.org/10.5281/zenodo.13603527, https://doi.org/10.5281/zenodo.13603527

15. Rossetti, G., Citraro, S., Milli, L.: Conformity: a path-aware homophily measure for node-attributed networks. IEEE Intelligent Systems **36**(1), 25–34 (2021)
16. Sarker, A., Northrup, N., Jadbabaie, A.: Generalizing homophily to simplicial complexes. In: International Conference on Complex Networks and Their Applications. pp. 311–323. Springer (2022)
17. Sunstein, C.R.: The law of group polarization. University of Chicago Law School, John M. Olin Law & Economics Working Paper (91) (1999)
18. Traud, A.L., Mucha, P.J., Porter, M.A.: Social structure of facebook networks. Physica A: Statistical Mechanics and its Applications **391**(16), 4165–4180 (2012)
19. Treen, K., Williams, H., O'Neill, S., Coan, T.G.: Discussion of climate change on reddit: Polarized discourse or deliberative debate? Environmental Communication **16**(5), 680–698 (2022)
20. Valensise, C.M., Cinelli, M., Quattrociocchi, W.: The drivers of online polarization: Fitting models to data. Information Sciences **642**, 119152 (2023)
21. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, Ł., Polosukhin, I.: Attention is all you need. Advances in neural information processing systems **30** (2017)
22. Veldt, N., Benson, A.R., Kleinberg, J.: Combinatorial characterizations and impossibilities for higher-order homophily. Science Advances **9**(1), eabq3200 (2023)
23. Williams, H.T., McMurray, J.R., Kurz, T., Lambert, F.H.: Network analysis reveals open forums and echo chambers in social media discussions of climate change. Global environmental change **32**, 126–138 (2015)
24. Zignani, M., Gaito, S., Rossi, G.P.: Follow the "mastodon": Structure and evolution of a decentralized online social network. In: Proceedings of the International AAAI Conference on Web and Social Media. vol. 12, pp. 541–550 (2018)