Attraction by ingroup coherence drives the emergence of ideological sorting

Lucía Pedraza[1], Federico Zimmerman[2], Joaquín Navajas[2], and Pablo Balenzuela[1]

[1]Departamento de Física, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires and IFIBA (Instituto de Física de Buenos Aires) - CONICET.
[2]Laboratorio de Neurociencia, Universidad Torcuato Di Tella, Buenos Aires, Argentina.

Abstract

In this work, we propose an agent-based model to study the mechanism underlying polarization and ideological sorting in a group of agents. Based on experimental results, we assume that the two main mechanisms driving interactions are homophily, in which people who share similar opinions are more likely to interact, and in-group-coherence favoritism, in which they are more attracted to coherent in-groups rather than incoherent or out-group members. Interestingly, by incorporating this latter assumption, we were able to observe ideological alignment as observed in political opinions. Additionally, by solving the model's master equations and fitting the model's parameters with opinions from survey's data and experiments we show that a combination of homophily and in-group coherence is the best explanation for the emergence of ideological sorting.

Keywords— opinion model, ideological sorting, agent based models, master equations

1 Introduction

The increasing political polarization has become a worrying concern in many different countries and a serious threat to society and democracy itself. In response to this scenario, interdisciplinary efforts combining empirical and theoretical approaches, have sought to better understand the involved mechanisms and the conditions under which societies are expected to polarize.

One promising way to address this question is by studying the behavior of agent-based models (ABMs) under different conditions of social influence and interactions. Based on a few simplistic assumptions, most studies have focused on under which conditions individuals may reach consensus or polarize and become more extreme across one single topic, presumably oversimplifying the actual complexity of the current landscape and the underlying social dynamics.

While disagreement on policy issues is increasing, opinions are also becoming more extreme but also more coherent across diverse and seemingly unrelated topics. In several countries people have ideologically sorted in terms of partisanship leading to an issue alignment. For example, in the US, an individual who supports the women's right to voluntarily terminate pregnancy will be more likely to support stricter legislation on gun control, even though these topics are, a priori, unrelated to each other.

A recent paper [1] have shown that people not only hold politically coherent opinions across very different issues but also that this property, i.e., political coherence, increases interpersonal attraction among co-partisans. In other words, individuals who hold coherent opinions are more attractive than those individuals having some degree of ambivalence in their attitudes (e.g., a person who is anti-abortion but supports gun control). However, whether and how this driver of interpersonal attraction relates to macro-level patterns of political polarization and partisan-ideological sorting remains largely unknown.

Interestingly, by incorporating this latter novel experimental finding from political psychology at the microscopic level, issue alignment emerges starting from a random distribution of opinions. We formulated the model, derived the master equation of the system and performed numerical simulations obtaining identical results. Additionally, we compared different final states of the model with data from multiple datasets (which include more than 20,000 opinions on different controversial issues). By fitting the model's parameters with the data, we were able to observe that homophily alone could not explain ideologically sorted states and enhancing the importance of considering ingroup-coherence favoritism in political interactions.

2 The model

We consider a system of N agents. Each one holds a multidimensional opinion, where each dimension reflects the agent's opinion on a particular issue. For the sake of simplicity, we will consider a debate on only two issues. Agent i will hold an opinion $O_i = (O_i^1, O_i^2)$ where each coordinate could be against (-1), neutral (0) or in favour (1). For convenience, when possible, we assume that every political issue expresses a right-wing opinion facilitating us to notate right-wing opinions as positive, left-wing opinions as negatives, and undecided as zero. We also classify all agents by their ideological consistency. This procedure leads to following different communities:

- Coherent agents (C): hold assertive and matching opinions on both issues ($O = (-1, -1)$ or $O = (1, 1)$).
- Incoherent agents (I): hold assertive but opposite opinions on the two different issues ($O = (-1, 1)$ or $O = (1, -1)$).
- Weak agents (W): Agents that hold an assertive opinion in one issue and are undecided regarding the other one ($O = (-1, 0)$, $O = (0, -1)$, $O = (1, 0)$ or $O = (0, 1)$).
- Apathetic agents (A): Agents that are undecided on both issues ($O = (0, 0)$).

Over time, agents interact with each other and these interactions influence agents' opinions. At each time step, two agents are randomly selected and their interaction would lead one of the agents to influence the other with probability P. We incorporate two different mechanisms that impact agents' influence coupled by a parameter k.

- **Homophily** We define similarity as one minus the normalized Manhattan distance between the agents' opinions (L_1 distance, i.e. the absolute value of the sum of the distance in both coordinates): $S_{ij} = 1 - \frac{|O_i - O_j|_1}{4}$
- **Ingroup coherence favoritism** We define agents' ideology as the sum of the opinions on the two issues and we normalize it by dividing it by two: $I_i = \frac{O_i^1 + O_i^2}{2}$. The ideology sign value corresponds to the agent's leaning, a positive value corresponds to right-wing agents and negative one to left-wing agents. The ingroup coherence displayed by agent i to agent j is the absolute value of the ideology of i, if both are in the same group ($\sigma_{ij} = 1$), and zero otherwise: $C_{ij} = |I_i|\sigma_{ij}$.

The probability of influence is defined $P_{ij} = (1 - k)S_{ij} + kC_{ij}$

For high values of k only similarity is taking action and for low values only coherence.

Pairwise interactions lead to opinion changes that could occur only in one of the two issues and only by one unit. When agents are similar, influence will be attractive and agent i will move closer to j by changing one of its own opinions. Conversely, for dissimilar agents interactions are repulsive. In this case, agent i will change one of its opinions moving further from agent j and reducing their similarity.

3 Simulations, analytic results and empirical data

To derive the master equation, we consider that at each interaction, two agents are randomly selected. Therefore, the probability that these agents belong to a community is the fraction of agents in the communities. Then, we compute the flux between populations by considering the likelihood of changing an opinion after an interaction, obtaining the following equations:

$$\begin{aligned}
\frac{dC}{dt} &= WC\left(\frac{k}{16} + \frac{1-k}{4}\right) + W^2\frac{k}{16} \\
\frac{dI}{dt} &= WI\frac{k}{16} + W^2\frac{k}{16} \\
\frac{dA}{dt} &= -AC\frac{k}{2} - AI\frac{k}{2} \\
\frac{dW}{dt} &= -\frac{dC}{dt} - \frac{dI}{dt} - \frac{dA}{dt}
\end{aligned}$$

These equations were numerically solved with randomly distributed initial conditions such that: $C(0) = I(0) = 2/9$, $A(0) = 1/9$ y $W(0) = 4/9$. For numerical simulations of the ABM , we consider a system with N=1000 agents and averaged the results of 100 simulations per condition.

In Figure 1 (A) we show the proportion of coherent and incoherent agents, depending on the parameter k. The parameter k was varied from 0 to 1 with steps of 0.05. We can see the perfcet match between simulations and the theoretical results. In all cases (except $k = 1$), only the coherence and incoherence agents remain whereas the apathetic and the weaks disappear. For higher k, the coherence mechanism is stronger leading the system to a final state where the coherence community dominates the population.

The model's simulations and the analytical formulation display a clear relationship between the influence of in-group coherence favoritism and the final proportion of coherent agents. Next, we analyze to what extent actual opinions on a great variety of controversial issues are sorted. We work with multiple datasets [2, 3, 4] with more than 20,000 responses on different polarizing topics. All responses indicate the participants' agreement,

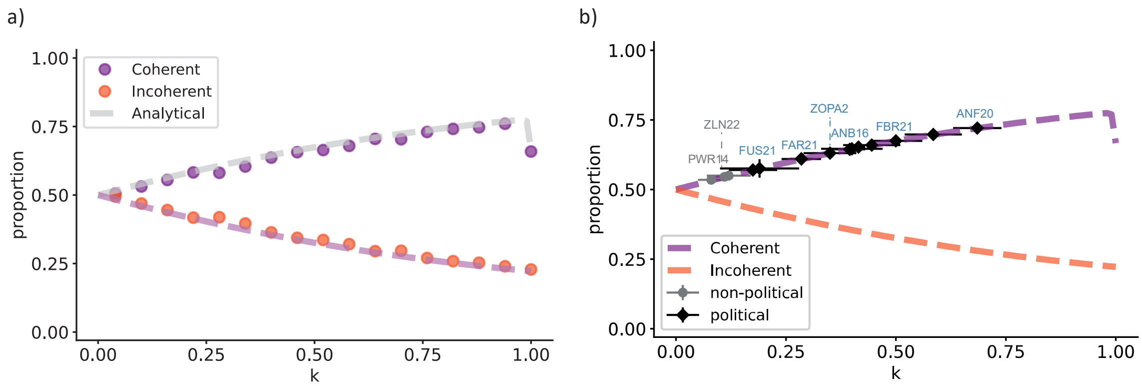

Figure 1: A) The model's final states are shown for different values of k. The figure depicts the mean values of 10 different simulations per scenario. The final proportion of coherent agents is shown in purple and the final proportion of incoherent agents in orange. For $k < 1$, as k increases, so does the final proportion of coherent agentes. B) Every dataset's mean sorting value was mapped to its corresponding model's k value. Non-political datasets are shown in light gray and political ones in black.

disagreement or neutrality to each different issue. In order to contrast the data to the proposed two-dimensional model, we focused on the relation between the proportion of coherent (C) and incoherent (I) agents and did not consider in the analysis the weak and apathetic populations. Thereby, we compute sorting ($S = \frac{C}{C+I}$) for all the possible pairs of opinions within each dataset.

Having the model's analytical solution allows us to map every sorting value into its corresponding model's k parameter (Figure 1(B)). Non-political datasets are shown in light gray dots and black diamonds are used for political datasets. We note that the value of k (the proportion of ingroup coherence) divide non-political from political groups of opinions. Taken altogether, data suggests that homophily alone can not explain the emergence of the observed levels of sorting and political polarization. On the contrary, non-political controversial opinions exhibit the lowest observed levels of sorting which can be explained by homophily and attractive-repulsive interactions.

References

[1] Federico Zimmerman et al. "Political coherence and certainty as drivers of interpersonal liking over and above similarity". In: *Science advances* 8.6 (2022), eabk1909.

[2] *The American National Election Studies (ANES)*. https://electionstudies.org/data-center.

[3] *Pew Research*. https://www.pewresearch.org/.

[4] Lucia Freira et al. "The interplay between partisanship, forecasted COVID-19 deaths, and support for preventive policies". In: *Humanities and Social Sciences Communications* 8.1 (2021), pp. 1–10.