Inevitability of Polarization in Geometric Opinion Exchange

Abdou Majeed Alidou[1], Júlia Baligács[2], Max Hahn-Klimroth[3], Jan Hązła[1], Lukas Hintze[4], Olga Scheftelowitsch[5]

[1]AIMS Rwanda, [2]TU Darmstadt, [3]Goethe University Frankfurt, [4]Universität Hamburg, [5]TU Dortmund.

abdou@aims.edu.gh

Motivation

Our society functions as a complex network of individuals with diverse beliefs and motivations, constantly reshaping their opinions through interactions. Understanding the evolution of opinions is of great scientific interest, e.g., in computer science and economics.

Mathematical models have been proposed to describe the evolution of individuals' opinions in society and this remains an active area of research. Opinions are generally represented by numbers or vectors, and update rules are defined to describe the evolution of the opinions under external influence (advertisement, state laws) or through interaction between individuals/agents.

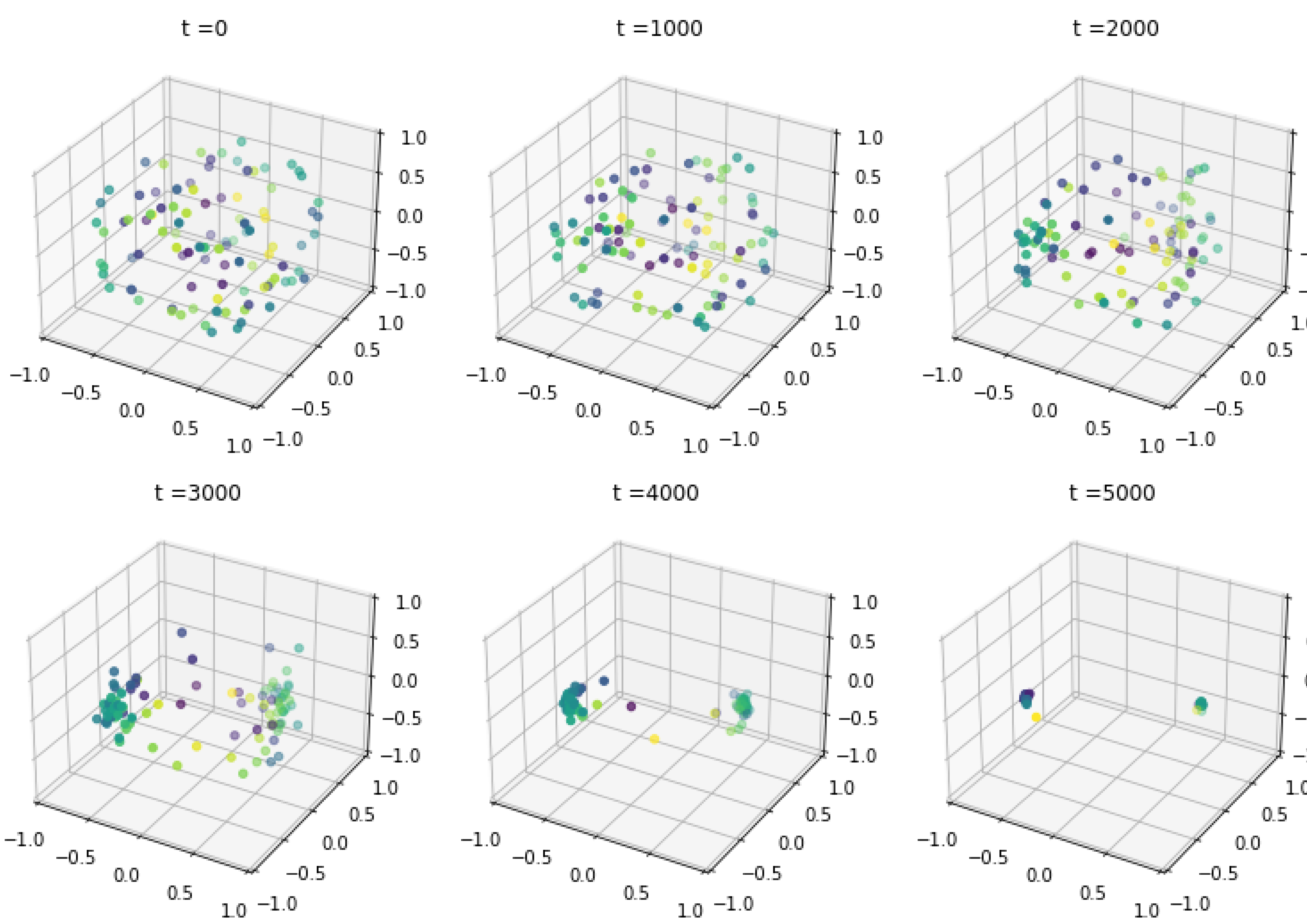
In an opinion exchange setting, agents interact with each other and update their beliefs based on the opinions of the agents they interact with. Often, we are interested in the final distribution of opinions of the agents: whether they reach a consensus, whether the society successfully gets rid of incorrect beliefs, and whether some agents can manipulate beliefs to their advantage (e.g. media, politicians, advertisers), as discussed in [MT17] and [AO11].

One phenomenon present in some societies is the tendency of individuals to align themselves into groups with strongly opposing views on specific topics, this is called **issue radicalization**. Additionally, we sometimes observe unexpected correlations between opinions on seemingly unrelated subjects such as the Palestine/Israel conflict, climate change, and immigration policy (**issue alignment**). More broadly, these phenomena are known as **polarization of opinions**.

Most existing mathematical models of opinion evolution struggle to comprehensively explain these occurrences. In [ABHH+24], we expand upon a geometric model from [HJMR23]. Our main modeling assumption is the tendency to interpret information to fit one's beliefs (**biased assimilation**).

Simulation of the polarization process with 100 agents and 4 topics

Simulation of the polarization process with 100 agents and 4 topics. The initial configuration was uniform on the unit sphere in 4 dimensions. The color represents the fourth dimension. The update rule used is the asymmetric $f(A) = \begin{cases} 0.5 \cdot A & \text{if } A \geq 0 \\ 0.1 \cdot A & \text{if } A < 0 \end{cases}$.



The Model

Let $d, n \geq 2$ denote the number of dimensions/topics and the number of agents, respectively. The *opinion* $\vec{u}_i$ of agent $1 \leq i \leq n$ is a d-dimensional vector on the unit sphere, in other words satisfying $\|\vec{u}_i\| = 1$.

A configuration $\mathcal{U}$ is a collection of n opinions. We say that a configuration is **polarized** when there exists a vector $\vec{v}$ such that for every agent i, either $\vec{u}_i = \vec{v}$ or $\vec{u}_i = -\vec{v}$.

Let f be a function from $[-1, 1]$ to $\mathbb{R}$. We consider discrete time-steps $t \in \{1, 2, \dots\}$, and an infinite sequence of i.i.d. random variables $I^{(1)}, I^{(2)}, \dots$ drawn from a distribution $\mathcal{D}$ on the set of all pairs $(i, j) \in [n] \times [n]$.

For a given initial configuration of opinions $\mathcal{U}^{(0)}$, we define the configurations at time $t = 1, 2, \dots$ as follows: if $I^{(t)} = (i, j)$, then agent j influences agent i and the opinion $\vec{u}_i$ is updated as follows:

$$\vec{u}_i^{(t+1)} = \frac{\vec{w}}{\|\vec{w}\|} \quad \text{with } \vec{w} = \vec{u}_i^{(t)} + f\left(A_{ij}^{(t)}\right) \cdot \vec{u}_j^{(t)},$$

where $\vec{u}_i^{(t)}, \vec{u}_j^{(t)}$ are the opinion of agents i and j at time t and $A_{ij}^{(t)}$ is the correlation (dot product) between $\vec{u}_i^{(t)}$ and $\vec{u}_j^{(t)}$.

Results

We aim to show that polarization almost surely occurs for a broad class of update functions.

Stable update function: A function f from $[-1, 1]$ to $\mathbb{R}$ is stable if it is continuous and if

$$\operatorname{sign}(f(A)) = \operatorname{sign}(A),$$

for all A. For example, the scaled identity

$$f(A) = \eta A, (\eta > 0)$$

is a stable update function.

Theorem in 2D: Let $d = 2$, $n \geq 2$, the interaction distribution $\mathcal{D}$ have full support, f be a stable update function, and $\mathcal{U}^{(0)}$ be any initial configuration which is not separable (i.e. not formed by two orthogonal sub-vector spaces).

Then, almost surely, the random process $\mathcal{U}^{(t)}$ polarizes.

Acknowledgment

This research project was initiated as a part of DAAD activity in Dortmund organized by Amin-Coja Oghlan and Nicola Kistler and their research groups.

References

[ABHH+24] Abdou Majeed Alidou, Júlia Baligács, Max Hahn-Klimroth, Jan Hązła, Lukas Hintze, and Olga Scheftelowitsch. "Inevitability of Polarization in Geometric Opinion Exchange". In: *arXiv preprint arXiv:2402.08446* (2024).

[AO11] Daron Acemoglu and Asuman Ozdaglar. "Opinion Dynamics and Learning in Social Networks". In: *Dynamic Games and Applications* 1.1 (Mar. 2011), pp. 3–49.

[HJMR23] Jan Hązła, Yan Jin, Elchanan Mossel, and Govind Ramnarayan. "A Geometric Model of Opinion Polarization". In: *Mathematics of Operations Research* (2023).

[MT17] Elchanan Mossel and Omer Tamuz. "Opinion exchange dynamics". In: *Probability Surveys* 14 (2017), pp. 155–204.