Reproduction of results: Dynamics of opinion polarization

Salom Iochanas @ TUC
Thomas Lagkalis @ TUC

June 11, 2025

Abstract

This code report follows the paper *"Dynamics of Opinion Polarization", E.Biondi et al.*[1] and documents the implementation and computational experiments used to reproduce its theoretical findings. The core of the analysis revolves around the generalized Friedkin–Johnsen (gFJ) model of opinion dynamics. The code implements various configurations of influence and susceptibility matrices and computes polarization metrics (P2, P3, P4) under different network topologies, including random graphs, the Karate Club network, and Barabasi–Albert graphs. Our implementation is available at github.com/ThomasLagkalis/Opinion-Polarization-Dynamics

1 Introduction

With the rise of online platforms, the mechanisms driving opinion polarization, where groups drift toward extreme positions, have gained remarkable attention. Social networks, though, do not merely provide a transparent technological substrate that facilitates interactions in the online dimension. The Dynamics of Opinion Polarization paper explores this phenomenon through the lens of the generalized Friedkin–Johnsen (gFJ) model, a well-established framework in opinion dynamics that integrates both personal prejudice and susceptibility to social influence.

This code report presents the methodology and implementation used to reproduce the theoretical results from the paper. Specifically, dives in the formulation used to implements the system's parameters (section 2). Then, provides the formulas used to calculate polarizing vectors based on the Theorems analyzed in the paper (sections 3 and 4). Finally, it showcases the reproduction results and compares them with those of the paper (section 5).

2 System formulation

As a first step we construct the generalized Friedkin-Johnsen (gFJ) model. The system requires the following main inputs:

1. s: the vector of the initial opinions (prejudice)
2. Λ: The diagonal matrix of susceptibilities
3. W: The influence matrix of the graph

From theorem 1 of the paper, we can compute the steady-state solution for the gFJ, for systems where the convergence condition for ΛW is satisfied. This closed form matrix H_g is easily computable and plays a central role in analyzing the polarization in the model.

$$z = (I - \Lambda W)^{-1}(I - \Lambda)s = H_g s$$

We consider three different settings in this implementation:

Random:
In this setting, both Λ and W are initialized with uniformly random values for n nodes. The matrix W is then normalized to ensure it is row-stochastic.

Karate Club:
This dataset provides the basic graph $G = (V, E)$, from that we can derive the weight and susceptibility matrices as follows:

- The influence weights are set proportionally to node degrees.
- the Λ matrix is set using one of three ways:
  1. Constant $\lambda_i = 0.8$
  2. Proportional λ to PageRank Centrality
  3. Inverse proportional λ to PageRank Centrality

Barabasi-Albert:
To study a more influencer heavy network, we generate a Barabasi–Albert graph. The matrices W and Λ are processed in the same way as the Karate Club. The only addition is that the number of connections for each node is set to $m = 2$.

3 Polarization vectors for P2 and P3

Following the paper as a guide the first objective we tackle is the generating polarization vectors for the global metrics $P2$ and $P3$:

$$P2(x) = \frac{1}{|V|}\sum_{i\in V} x_i^2 = \frac{1}{|V|}\|x\|_2^2$$

$$P3(x) = \sum_{i\in V} x_i^2 = \|x\|_2^2$$

According to Theorem 5 in the paper, we can derive initial opinion vector $s_{B1(1)}$ with following steps:

- Find the eigenvalues and eigenvectors of $H_g^T H_g$.
- Set $s_{B1(1)}$ the top eigenvector, which corresponds to the local maximum for the $P2, P3$ polarization on the L_2 ball.

A variation of this vector $s_{B1(t)}$ can be obtained by scaling for $t = 1/s_{B2}^{(k)}(1)$, where $s_{B2}^{(k)}(1)$ is the largest $s_{B1(1)}$ entry.

To evaluate how close a given initial opinion vector is to the most polarizing configuration (in terms of metrics $P2, P3$), we solve the following non-convex minimization problem using scipy (approximately):

$$f(s) = -\Delta P3(s) = -(\|z\|_2^2 - \|s\|_2^2) = -\|H_g s\|_2^2 + \|s\|_2^2$$

$$\text{method: SLSQP, bounds: } s \in [0,1]^n$$

4 Polarization vectors for P4

The second task is to find an initial opinion vector that maximizes polarization under the $P4$ metric:

$$P4(x) = \sum_{i\in V} |x_i| = \|x\|_1$$

Theorem 6 suggests the following method for finding one:

- Find the column with the highest column sum of H_g, $j = \text{argmax}\sum_i h_{ij}$
- Set $s_{B1(1)}$ to $e_j = (000...010...000)$

Similarly we want to compare it to the most polarizing configuration for $P4$. We define the following minimization problem:

$$f(s) = -\Delta P4(s) = -(\|z\|_1 - \|s\|_1) = -\|H_g s\|_1 + \|s\|_1$$

$$\text{method: SLSQP, bounds: } s \in [0,1]^n$$

5 Comparing the results

With the aforementioned components in place, we can now evaluate every initial opinion vector across the different environments. We will use a similar format as the paper to present the results.

First, we present the tables of the metrics for every environment. As expected, the polarization shifts are in tune with the theory:

$$\Delta P(s_{B_2(1)}) \leq \Delta P(s_{B_2(t)}) \leq \Delta P(s_{\max}^{P_2,P_3})$$

$$\Delta P(s_{B_1(1)}) \leq \Delta P(s_{\max}^{P_4}).$$

	Δ_{P_1}	Δ_{P_2}	Δ_{P_3}	Δ_{P_4}	Δ_{NDI}	Δ_{GDI}
s_{unif}	-1.48	-0.21	-2.11	-2.51	-3.38	-14.978
$s_{B_1(1)}$	-0.49	0.08	0.82	1.32	-0.99	-4.87
$s_{B_2(1)}$	-0.31	-0.02	-0.09	0.74	-0.64	-1.54
$s_{B_2(t)}$	-1.08	0.16	1.65	2.15	-2.09	-10.85
$s_{max}^{P_2,P_3}$	-0.26	0.04	0.43	0.77	-0.53	-2.63
$s_{max}^{P_4}$	-0.19	0.02	0.20	1.22	-0.33	-1.93

Table 1: Uniformed random for constant λ

	Δ_{GDI}	Δ_{NDI}	Δ_{P_1}	Δ_{P_2}	Δ_{P_3}	Δ_{P_4}
s_{unif}	-321.39	-14.71	-9.45	-0.26	-8.69	-8.96
$s_{B_2(1)}$	-53.42	-8.88	-1.57	0.07	2.52	5.46
$s_{B_2(t)}$	-53.42	-8.88	-1.57	0.07	2.52	5.46
$s_{max}^{P_2,P_3}$	-85.65	-7.08	-2.52	0.16	5.56	5.56
$s_{B_1(1)}$	-27.65	-6.48	-0.81	-0.02	-0.54	2.23
$s_{max}^{P_4}$	-252.99	-23.11	-7.44	0.04	1.42	9.10

Table 2: Karate Club constant λ

	Δ_{GDI}	Δ_{NDI}	Δ_{P_1}	Δ_{P_2}	Δ_{P_3}	Δ_{P_4}
s_{unif}	-2730.45	-46.01	-27.30	-0.27	-27.32	-32.37
$s_{B_2(1)}$	-258.03	-14.92	-2.58	0.04	4.03	11.60
$s_{B_2(t)}$	-258.03	-14.92	-2.58	0.04	4.03	11.60
$s_{max}^{P_2,P_3}$	-426.52	-15.74	-4.27	0.12	11.98	11.98
$s_{B_1(1)}$	-83.35	-7.06	-0.83	-0.01	-0.67	3.18
$s_{max}^{P_4}$	-2097.50	-68.86	-20.97	-0.01	-1.46	22.54

Table 3: Barabasi-Albert constant λ

	Δ_{GDI}	Δ_{NDI}	Δ_{P_1}	Δ_{P_2}	Δ_{P_3}	Δ_{P_4}
s_{unif}	-96.98	-10.77	-2.85	-0.08	-2.84	-2.19
$s_{B_2(1)}$	-1.58	-0.21	-0.05	0.00	0.05	0.05
$s_{B_2(t)}$	-1.58	-0.21	-0.05	0.00	0.05	0.05
$s_{max}^{P_2,P_3}$	-1.49	-0.22	-0.04	0.00	0.05	0.05
$s_{B_1(1)}$	-9.35	-0.52	-0.28	-0.01	-0.27	0.04
$s_{max}^{P_4}$	-75.81	-9.10	-2.23	-0.05	-1.87	0.29

Table 4: Karate Club $\lambda \propto C$

	Δ_{GDI}	Δ_{NDI}	Δ_{P_1}	Δ_{P_2}	Δ_{P_3}	Δ_{P_4}
s_{unif}	-2097.50	-68.86	-20.97	-0.01	-1.46	22.54
$s_{B_2(1)}$	-1.86	-0.11	-0.02	0.00	0.02	0.02
$s_{B_2(t)}$	-1.86	-0.11	-0.02	0.00	0.02	0.02
$s_{max}^{P_2,P_3}$	-1.79	-0.11	-0.02	0.00	0.02	0.02
$s_{B_1(1)}$	-20.76	-0.55	-0.21	-0.00	-0.21	0.02
$s_{max}^{P_4}$	-468.57	-21.01	-4.69	-0.04	-4.41	0.24

Table 5: Barabasi-Albert $\lambda \propto C$

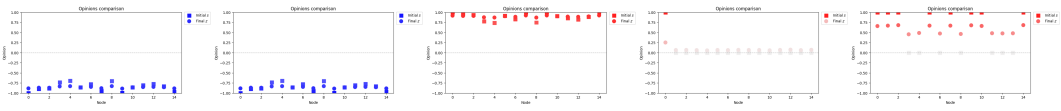
	Δ_{GDI}	Δ_{NDI}	Δ_{P_1}	Δ_{P_2}	Δ_{P_3}	Δ_{P_4}
s_{unif}	-125.52	-6.87	-3.69	-0.11	-3.68	-2.48
$s_{B_2(1)}$	0.84	-1.70	0.02	0.02	0.59	1.54
$s_{B_2(t)}$	0.84	-1.70	0.02	0.02	0.59	1.54
$s_{max}^{P_2,P_3}$	-26.93	-2.40	-0.79	0.06	2.01	2.01
$s_{B_1(1)}$	-0.81	-1.37	-0.02	0.00	0.07	1.07
$s_{max}^{P_4}$	-85.82	-9.11	-2.52	0.01	0.31	3.74

Table 6: Karate Club $\lambda \propto 1/C$

	Δ_{GDI}	Δ_{NDI}	Δ_{P_1}		Δ_{P_3}	Δ_{P_4}
s_{unif}	-545.31	-12.65	-5.45	-0.05	-5.41	-4.21
$s_{B_2(1)}$	0.51	-0.87	0.01	0.00	0.30	1.28
$s_{B_2(t)}$	0.51	-0.87	0.01	0.00	0.30	1.28
$s_{max}^{P_2,P_3}$	-96.59	-3.35	-0.97	0.03	2.62	2.62
$s_{B_1(1)}$	-1.58	-0.82	-0.02	0.00	0.00	0.66
$s_{max}^{P_4}$	-510.59	-16.82	-5.11	-0.01	-0.99	5.46

Table 7: Barabasi-Albert $\lambda \propto 1/C$

To visualize individual opinion shifts, we plot each node's initial opinion (indicated by a square) alongside its corresponding final opinion (indicated by a circle). This visualization is provided for each environment.



(a) B2(1) (b) B2(t) (c) max_P2P3 (d) B1(1) (e) max_p4

Figure 1: Opinion shifts on the Random network

(a) B2(1) (b) B2(t) (c) max_P2P3 (d) B1(1) (e) max_p4

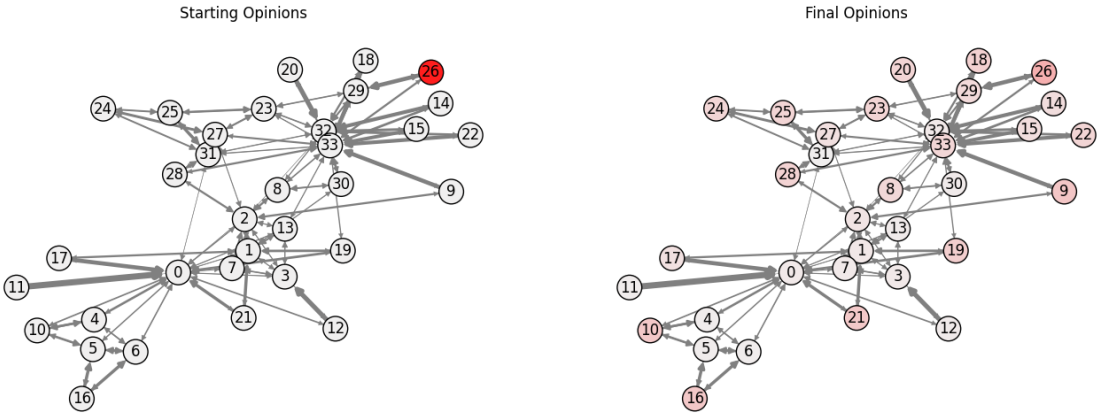
Figure 2: Opinion shifts on the Karate Club

(a) B2(1) (b) B2(t) (c) max_P2P3 (d) B1(1) (e) max_p4

Figure 3: Opinion shifts on Barabasi-Albert

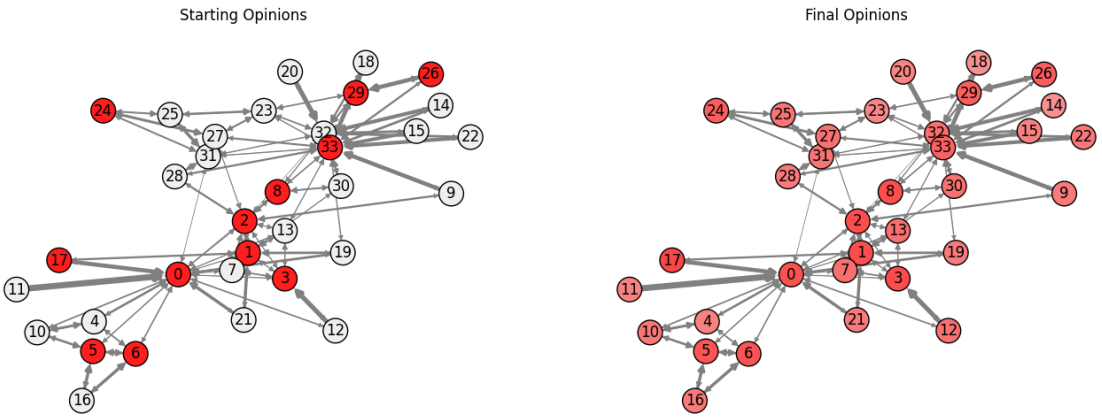
To illustrate the effects of the different polarization vectors, we include representative visualizations of the network. These plots can showcase the spread and influence of certain nodes.

For the $P4$ polarizing vectors, we observe the effect of the most influential node alone selected across the network. This can be compared to the optimal solution, which appears to strategically select a set of influential nodes rather than relying on just one.



(a) initial opinions (b) final opinions

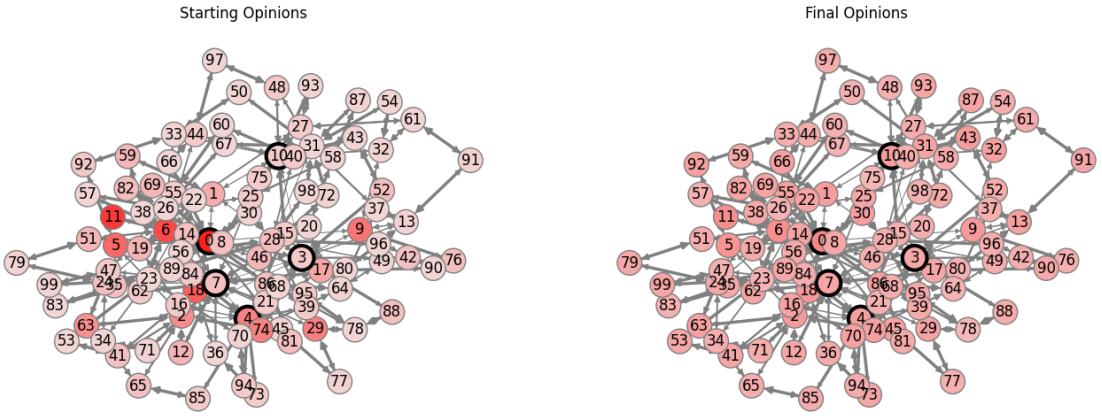
Figure 4: Karate Club B1(1) initial opinion vector



(a) initial opinions (b) final opinions

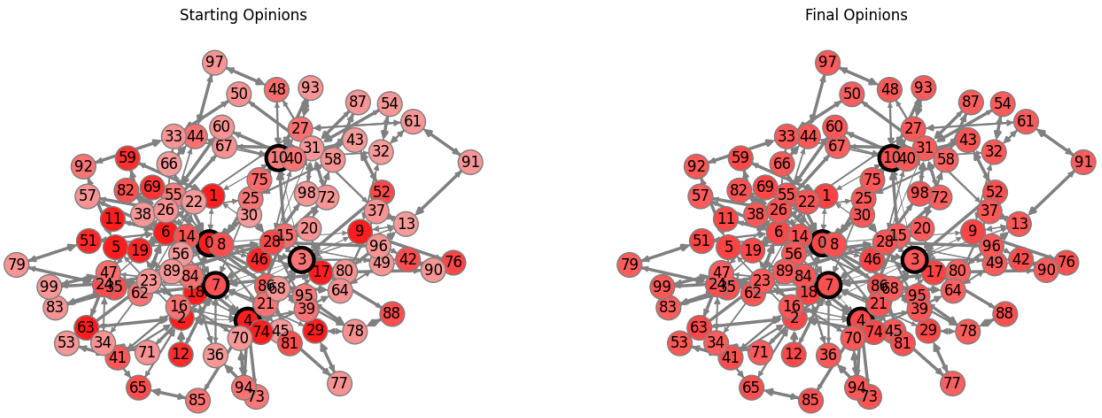
Figure 5: Karate Club max P4 initial opinion vector

For the Barabasi-Albert we have highlighted the top 5 most influential nodes (aka hubs) with a thick black border. Notably, both $s_{B1(t)}$ and the optimal solutions for $P2$ and $P3$ do not assign extreme opinions to all influential nodes, but rather a subset of them.



(a) initial opinions (b) final opinions

Figure 6: Barabasi-Albert B2(t) initial opinion vector



(a) initial opinions

(b) final opinions

Figure 7: Barabasi-Albert max P2, P3 initial opinion vector

References

[1] Elisabetta Biondi et al. "Dynamics of Opinion Polarization". In: *IEEE Transactions on Systems, Man, and Cybernetics: Systems* 53.9 (2023), pp. 5381–5392. DOI: 10.1109/TSMC.2023.3268758.