

# Simulating Opinion Dynamics over Social Networks

Hoang-Son Nguyen

December 7, 2023

## 1 Introduction

### 1.1 Background and Motivation

Nowadays, online social network platforms such as Twitter (now X) and Facebook are prevalent in daily life, allowing users to connect with friends, exchange ideas, discuss recent news, etc. The online social networks are believed to play a large role in forming and changing users' opinions, for example in recent presidential elections in the United States [1]. Consequently, there is a huge interest from economists, political scientists, policy makers,... to study how opinions evolve over time in a social network [2].

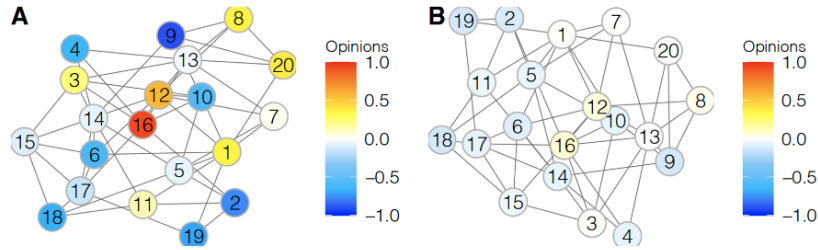


Figure 1: Network configurations and opinions arrangement for Friedkin-Johnsen opinion dynamic model, (A) at the beginning and (B) at the end (figure taken from [3])

It is observed recently that the online social networks such as Facebook, Twitter, etc. facilitates the polarization and disagreement of opinions, instead of consensus [4], [5], [6]. The polarization of opinions potentially cause harm to society, as individuals are tilted towards one end of the extremes, therefore creating unbridgeable gaps that might lead to devastating, unnecessary social conflicts. Motivated by this intriguing phenomenon of the social systems, in this project we investigate how the social network affects opinions of individuals, in particular the polarization and disagreement of opinions.

### 1.2 Goals

We aim study how continuous opinions evolve over time on a social network structure. Our central research question is **how structure of the social network affects the opinion dynamics**.

1. In the first part, we review the stochastic block model, which can describe real-world social networks that are known for their cluster structures. We then report on two most common dynamics models for continuous opinions, namely the linear Friedkin-Johnsen model [7] and the non-linear Hegselmann-Krause model [8], together with some of their notable mathematical properties.
2. In the second part, we perform simulations to study opinion dynamics over networks, using both generated data and real-world data. Using the synthetic experiments, we generate multiple networks for simulation, and compare steady state, convergence rate, disagreements and polarization, etc. in different network topologies.

### 1.3 Simulation Methodology

Our simulation consists of two stages. In the first stage, we use `NetworkX` library to randomly generate a network according to stochastic block models [9], along with some statistics of the networks. In the second stage, we use `NDlib` library to simulate opinion dynamics over the generated graph. We then gather the simulation data to visualize the process, as well as to compute various statistics regarding the opinions, such as convergence time, polarization, disagreement, etc.

## 2 Model Description

### 2.1 Modeling Social Networks

A social network with  $n$  individuals can be described as a  $n$ -node undirected graph  $\mathcal{G} = (\mathcal{V}, \mathcal{E})$  where  $\mathcal{V} = \{1, 2, \dots, n\}$  and  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ . The graph  $\mathcal{G}$  can be represented by the corresponding (weighted) adjacency matrix  $\mathbf{A} \in \mathbb{R}^{n \times n}$ , in which  $\mathbf{A}_{ij} \neq 0$  if  $(i, j) \in \mathcal{E}$  and  $\mathbf{A}_{ij} = 0$  if  $(i, j) \notin \mathcal{E}$ . The Laplacian matrix of a graph is defined as  $\mathbf{L} = \mathbf{D} - \mathbf{A}$ , where  $\mathbf{D} = \text{diag}(\mathbf{A}\mathbf{1})$  is the diagonal degree matrix of  $\mathcal{G}$ .

A popular statistical model that can describes many prevalent features of social networks is the latent variable model, in which the latent variables (or “social” variables) determine the relationship of one individual to others [10]. A prominent latent variable model for social networks is the stochastic block model (SBM) [9]. The key idea behind SBM is that “birds of a feather flock together”, which is rooted in observations made by sociologists that individuals having similar characteristics are more likely to connect. Likewise, in an SBM model, the nodes that belong to the same community are more likely to share an edge, whereas two nodes from two different clusters are less likely to be connected.

Mathematically, an SBM of  $n$  nodes and  $k$  communities can be described by a membership matrix  $\mathbf{Z} \in \{0, 1\}^{n \times k}$  denoting which nodes belong to which communities, and a connectivity matrix  $\mathbf{B} \in [0, 1]^{k \times k}$  specifying probabilities of an edge between two nodes from two communities [11]. Due to the homophily effect, it is often assumed that  $\mathbf{B}_{ii} \gg \mathbf{B}_{ij}$  for  $1 \leq i, j \leq k$  and  $i \neq j$ , i.e. it is more probable that two nodes in the same community are linked. Then, the expectation of the adjacency matrix  $\mathbf{A}$  is  $\mathbb{E}[\mathbf{A}] = \mathbf{Z}\mathbf{B}\mathbf{Z}^\top$ . In this project, we focus on the planted partition model (PPM), which is a simple SBM that assumes  $\mathbf{B} = (a - b)\mathbf{I} + b\mathbf{1}\mathbf{1}^\top$  for  $0 \leq b \leq a \leq 1$ <sup>1</sup>. Using `NetworkX`, we can generate an instance of PPM with 60 nodes in 3 clusters; see Fig. 2.

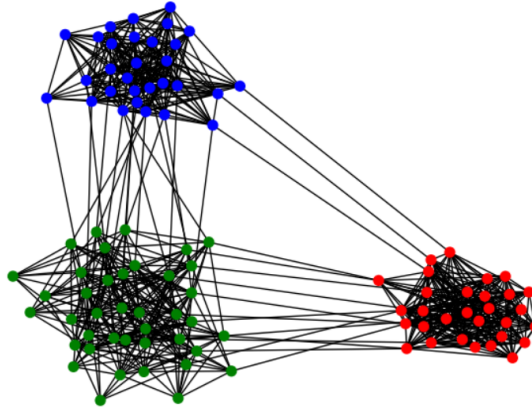


Figure 2: An instance of PPM with 60 nodes and 3 blocks (represented by 3 different colors).

### 2.2 Modeling Dynamics of Continuous Opinions

In one of the pioneering work in studying how opinions evolve in a social network, DeGroot in 1974 proposed a very simple and elegant mathematical model for linear dynamics of continuous opinions: at

<sup>1</sup>That means, two nodes in the same cluster connect with probability  $a$ , and two nodes from different clusters connect with probability  $b < a$ .

each iteration, the opinion of an individual is updated to be the (weighted) average of their neighbors [12]. The model is shown to be statistically significantly better at describing social learning phenomenon than Bayesian learning model, in which agents update their beliefs according to Bayes' rule [13].

Denote  $x_i(t) \in \mathbb{R}$  to be the opinion of  $i^{th}$  agent at time  $t$ . Then, the vector of opinions of a social network at time  $t$  can be represented by  $\mathbf{x}(t) = (x_1(t), \dots, x_n(t)) \in \mathbb{R}^n$ . At time  $t + 1$ , the opinion  $x_i(t + 1)$  is

$$x_i(t + 1) = \sum_{j \in \mathcal{N}_i} a_{ij} x_j(t) \quad (1)$$

where  $\mathcal{N}_i$  denotes the set of neighbors of node  $i$ , and  $a_{ij} \neq 0$  is the weight of the undirected edge  $(i, j) \in \mathcal{E}$ . One can compactly write the above update as

$$\mathbf{x}(t + 1) = \mathbf{A}\mathbf{x}(t) \quad (2)$$

where  $\mathbf{A} = [a_{ij}]_{1 \leq i, j \leq n}$  is the (weighted) adjacency matrix. With agents' opinions initiated at  $\mathbf{x}(0)$ , the update equation (2) implies

$$\mathbf{x}(t + 1) = \mathbf{A}^{t+1}\mathbf{x}(0) \quad (3)$$

and the steady state of opinions  $\mathbf{x}$  satisfies

$$\mathbf{x} = \mathbf{A}\mathbf{x} \quad (4)$$

i.e.  $\mathbf{x}$  is the eigenvector of  $\mathbf{A}$  that corresponds to eigenvalue 1. We note that DeGroot model resembles the model of a discrete-time Markov chain, hence the theory of Markov chain can be applied to DeGroot models to study their stationary state. Specifically, by [14, Proposition 1] or [15, Theorem 1], when the matrix  $\mathbf{A}$  is a stochastic matrix (i.e. the edge weights are non-negative and normalized to be sum to 1), the opinions converge to a unique steady state if and only if  $\mathbf{A}$  induces a connected, aperiodic graph. Therefore,

It is worth to mention that under various DeGroot models, there might emerge *stubborn agents* or *opinion leaders*, which are individuals that can greatly influence many other agents while their own opinions are not affected by others. In addition, one can also study *external influence* on the opinion dynamics of DeGroot model, in which there are sources outside of the social network influencing some agents, thereby affecting the social learning process of the whole network. We refer the readers to two surveys [15] and [16] for a more thorough picture on opinion dynamics and DeGroot models in particular.

The works by DeGroot laid the first foundation for non-Bayesian opinion dynamics models, from which many extensions are proposed. In this project, we mainly focus on simulating and analyzing two famous extensions of DeGroot model, namely Friedkin-Johnsen model [7] and Hegselmann-Krause model [8]. These models are widely used today for studying opinion dynamics in modern social networks, together with many fascinating social phenomena such as polarization and disagreement [4], [5], [6].

**Friedkin-Johnsen Model** As discussed previously, under weak conditions of the network structure, opinions in the original DeGroot model converge to consensus. However, it is empirically observed that opinions are often heterogeneous and polarized, which DeGroot model fails to capture [17]. Friedkin-Johnsen model further extends the traditional linear DeGroot model by incorporating innate opinions (that do not change) and exposed opinions (that can change) into modeling [7]. Mathematically speaking, for all  $1 \leq i \leq n$ , the linear update equation at time  $t + 1$  is

$$x_i(t + 1) = \frac{s_i + \sum_{j \in \mathcal{N}_i} a_{ij} x_j(t)}{1 + \sum_{j \in \mathcal{N}_i} a_{ij}} \quad (5)$$

where  $s_i$  is the innate opinion of agent  $i$ . One can observe that the final state of opinions  $\mathbf{x}$  can be compactly written as [7]

$$\mathbf{x} = (\mathbf{I} + \mathbf{L})^{-1}\mathbf{s} \quad (6)$$

**Hegselmann-Krause Model** A key assumption of many linear DeGroot models such as Friedkin-Johnsen model is that two agents readily exchange their opinions, whereas in practice two individuals with vastly different opinions would not talk to each other. A popular *non-linear* extension emerge in the early 2000s [8], [18], attempting to explain the aforementioned observation by using *bounded confidence*, where the idea is that two agents should only influence each other if their opinions are sufficiently close.

In this project, we choose to focus on [8]. The Hegselmann-Krause model has the same update rule as the original DeGroot model, but the update only happens when  $|x_i(t) - x_j(t)| \leq \epsilon$  for some pre-defined  $\epsilon > 0$ . One can observe that as  $\epsilon \rightarrow +\infty$ , the Hegselmann-Krause model recovers the DeGroot model. Mathematically speaking, at time  $t + 1$ , for all  $1 \leq i \leq n$ , the update equation now is

$$x_i(t+1) = \frac{\sum_{j \in \Gamma_\epsilon(i)} a_{ij} x_j(t)}{|\Gamma_\epsilon(i)|} \quad (7)$$

where  $\Gamma_\epsilon(i) = \{j \in \mathcal{N}_i : |x_i(t) - x_j(t)| \leq \epsilon\}$ , and  $|\Gamma_\epsilon(i)|$  denote the cardinality of the set  $\Gamma_\epsilon(i)$ .

Using `NDlib` library, we performed simulations on a generated graph from SBM with 100 nodes in 3 communities, using Hegselmann-Krause model. At each iteration, we randomly choose a node, then perform the update according to the model. We run the simulations with 10000 iterations. An interesting observation is that for a small enough  $\epsilon$ , the opinions diverge instead of converge. Therefore,  $\epsilon$  must be sufficiently large to ensure convergence to consensus of opinions; otherwise, agents with different opinions would not communicate with each other.

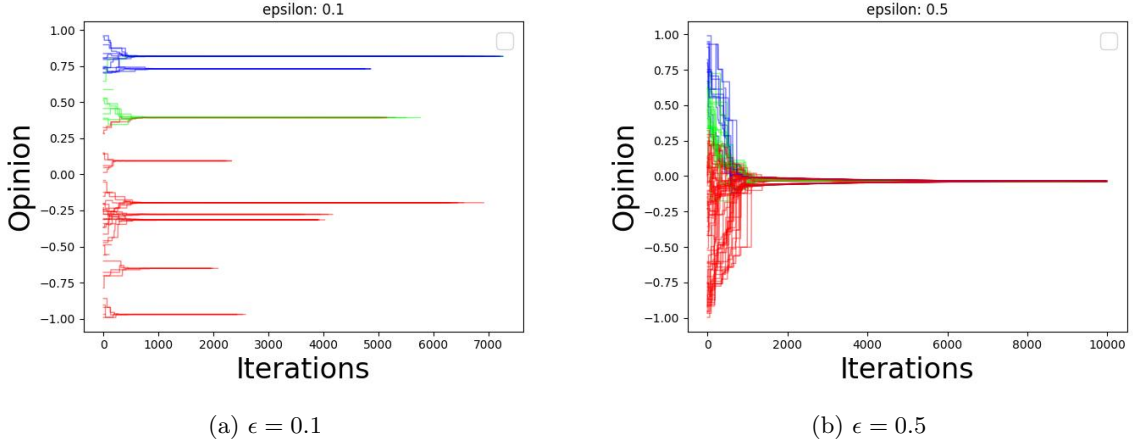


Figure 3: Simulation of two Hegselmann-Krause Models with different  $\epsilon$  on an SBM graph

### 3 Simulations

In the synthetic simulations, we utilize `NDlib` to simulate both Friedkin-Johnsen model and Hegselmann-Krause model, under different PPM models. For simplicity, we use  $\text{SBM}(3, 60, a, b)$  graphs, with undirected and unweighted edges. Specifically, we investigate the properties of opinion dynamics through (1) visualization of opinion evolutions, (2) convergence time of opinions, (3) polarization and disagreement scores of the steady opinions. For the convergence time, we report the number of iterations needed for all opinion changes to fall below a certain pre-defined threshold. Regarding the various measures related to stationary opinions, we chose the polarization and disagreement scores, as these statistics can reflect how far away from consensus the final opinions in a social network are, unlike merely means or variances of the opinions [4], [5], [6]. In the followings, we recall definitions of the two measures.

**Definition 1** (Polarization [6]). *Given opinions  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$  and the mean of its entries  $\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$ , the polarization of  $\mathbf{x}$  is*

$$P(\mathbf{x}) = \sum_{i=1}^n (x_i - \bar{x})^2 = \|\tilde{\mathbf{x}}\|_2^2$$

where  $\tilde{\mathbf{x}} = \mathbf{x} - \bar{x}\mathbf{1}$  are the mean-centered opinions.

**Definition 2** (Disagreement [6]). Given a vector of opinions  $\mathbf{x} = (x_1, \dots, x_n) \in \mathbb{R}^n$ , the (total) disagreement of the social network is

$$D(\mathbf{x}) = \sum_{(i,j) \in \mathcal{E}} (x_i - x_j)^2$$

In the simulations, we test *how the community structure of the social network can influence the convergence time, polarization score and disagreement score of the opinion dynamics*. Specifically, for the three-community SBM graph with  $N = 60$ , while keeping the parameter  $a = 10 \ln(N)/N$ , we vary the parameter  $b = 0.1 \ln(N)/N, 0.2 \ln(N)/N, \dots, 10 \ln(N)/N$ ; the smaller  $b$  corresponds to a more modular graph with the communities being less connected, and the bigger  $b$  corresponds to a graph with well-connected communities. The initial opinions are sampled from the  $\text{Normal}(0, 1)$ , and each simulation runs for 100 iterations. All updates consider the effect of self-loop.

### 3.1 Friedkin-Johnsen Model

For Friedkin-Johnsen model, we set the innate opinions of nodes in the first community to follow  $\text{Normal}(-1.0, 1)$ , those of nodes in the second community to follow  $\text{Normal}(0, 1)$ , and those of nodes in the third community to follow  $\text{Normal}(1.0, 1)$ . We provide the visualization of opinion evolution on two graphs,  $\text{SBM}(3, 60, 10 \ln(60)/60, 0.1 \ln(60)/60)$  and  $\text{SBM}(3, 60, 10 \ln(60)/60, 1.0 \ln(60)/60)$ , in Fig. 6 and Fig. 7.

From the simulation result in Fig. 4, as the nodes between different clusters are more well-connected, the convergence time slightly increases. In addition, we observe an interesting phenomenon in Fig. 5, as polarization and disagreement scores drop when the nodes between two different communities are better connected. Therefore, it suggests that *in a social network where people from a community readily communicate with members of another community, their opinions are less polarized even though their innate opinions differ*. We note that the simulation result aligns with a theoretical result provided in [6, Proposition 3], which states that

$$P(\mathbf{x}) \leq \frac{P(\mathbf{s})}{(1 + \frac{1}{2}d_{\min}h_{\mathcal{G}}^2)^2} \quad (8)$$

where  $\mathbf{x} = (\mathbf{I} + \mathbf{L})^{-1}$  is the steady state of opinions,  $\mathbf{s}$  is the innate opinions,  $d_{\min}$  is the minimum degree of the network  $\mathcal{G}$ , and  $h_{\mathcal{G}}$  is the Cheeger constant of a graph  $\mathcal{G}$ , measuring connectivity between large subsets of vertices in a graph [19]. As the communities in our simulation are well-connected, we can expect that  $h_{\mathcal{G}}$  becomes larger and hence the polarization index  $P(\mathbf{x})$  of the final opinions is smaller.

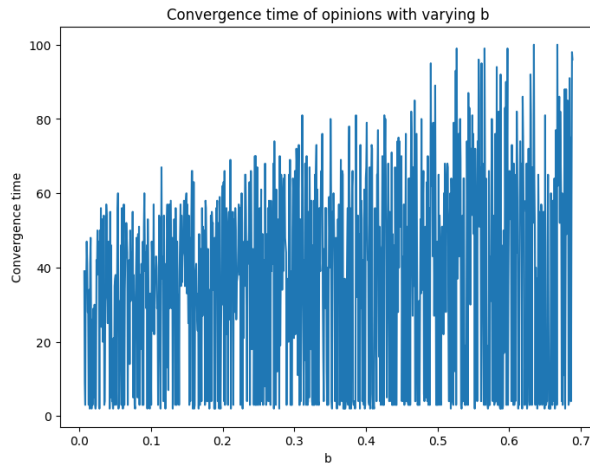
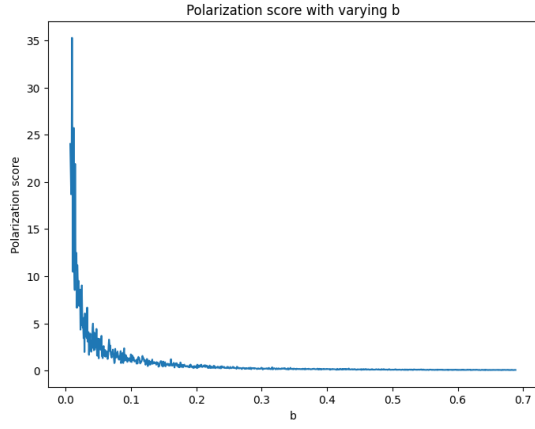
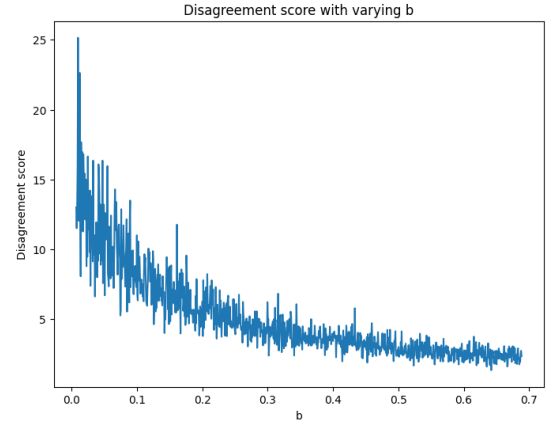


Figure 4: Convergence time of Friedkin-Johnsen model, with varying inter-community connectivity

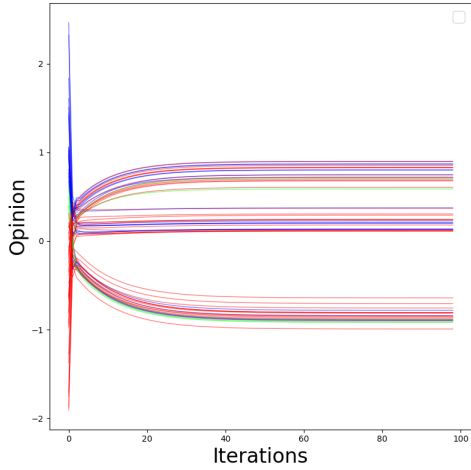


(a) Polarization score

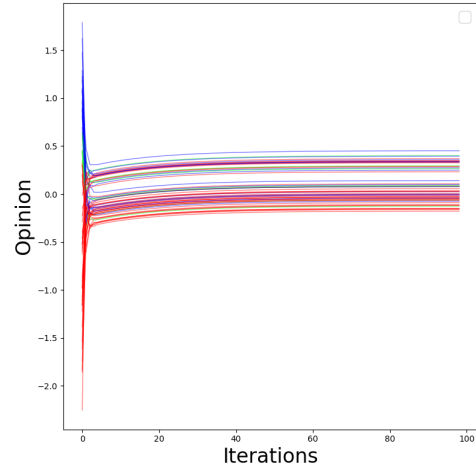


(b) Disagreement score

Figure 5: Polarization and disagreement scores of steady opinions following Friedkin-Johnsen model, with varying inter-community connectivity



(a) SBM(3, 60,  $10 \ln(60)/60$ ,  $0.1 \ln(60)/60$ )



(b) SBM(3, 60,  $10 \ln(60)/60$ ,  $1.0 \ln(60)/60$ )

Figure 6: Opinion evolutions of Friedkin-Johnsen model in two graphs with different inter-cluster connectivity; it is observed that as the different clusters are better connected, the agents achieve consensus

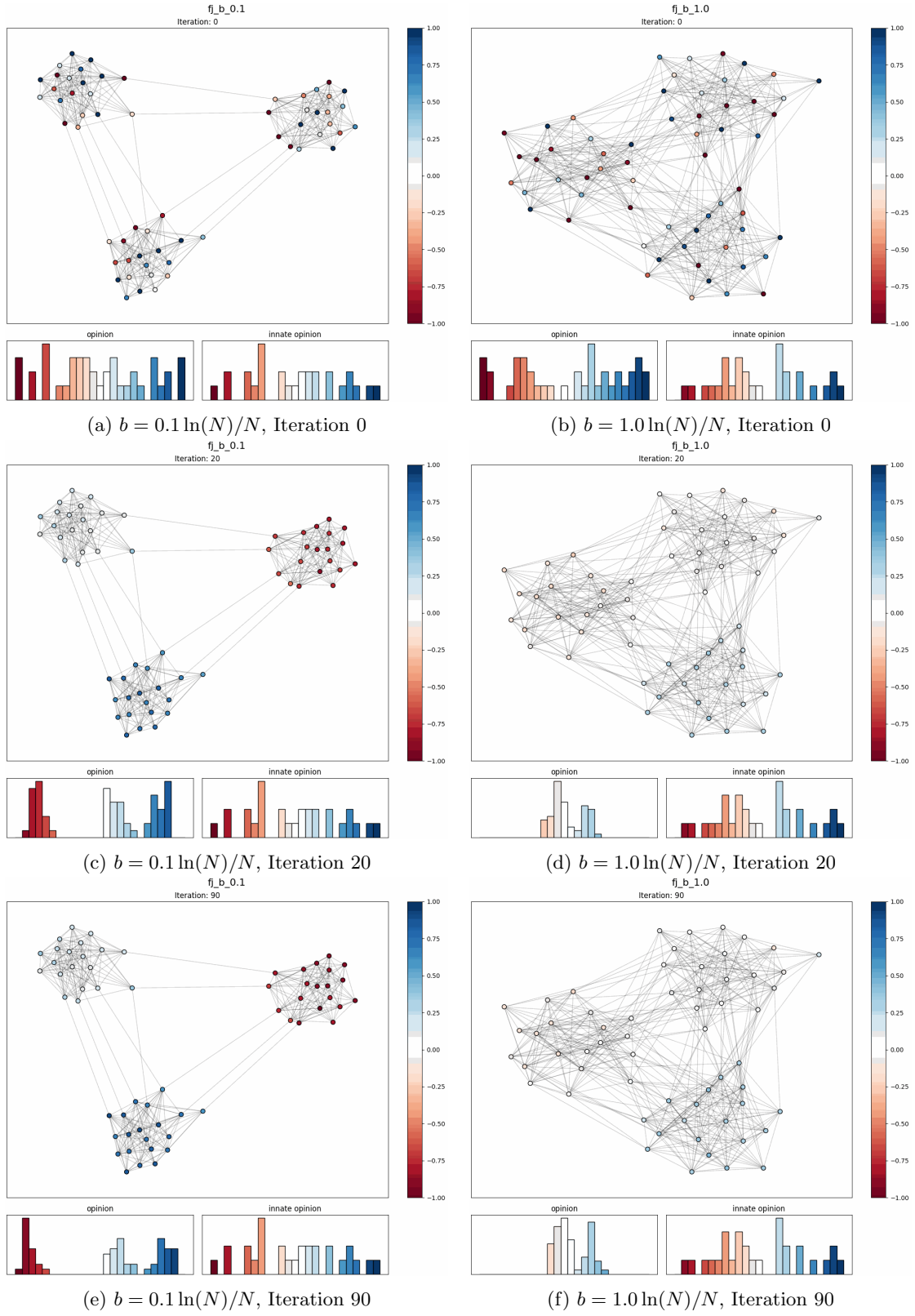


Figure 7: Comparing Friedkin-Johnsen opinion dynamics on two graphs with different inter-cluster connectivity  $b = 0.1$  and  $b = 1.0$

### 3.2 Hegselmann-Krause Model

For the Hegselmann-Krause model, we set the threshold for exchanging opinions to be  $\epsilon = 1.0$ , and perform the same simulation procedures as in the previous section. We observe that the convergence time of Hegselmann-Krause model to the same threshold is smaller than that of Friedkin-Johnsen model, by comparing Fig. 4 and Fig. 8. However, the polarization score and disagreement score reported in Fig. 9 are highly fluctuating and have no clear relationship with inter-community connectivity as in the case of Friedkin-Johnsen model. We speculate that this phenomenon is due to the non-linear nature of the Hegselmann-Krause model. Another reason for fluctuating polarization and disagreement indices is the *emergence of stubborn agents or opinion leaders in the social network*. As we can see from Fig. 10 and Fig. 11, for the case of SBM(3, 60,  $10 \ln(60)/60$ ,  $5.9 \ln(60)/60$ ) with a relatively high connectivity between clusters, three stubborn agents still emerge, which leads all opinions to converge to three opinions. A theoretical analysis for when and why stubborn agents emerge in Hegselmann-Krause model in SBM is left for future works.

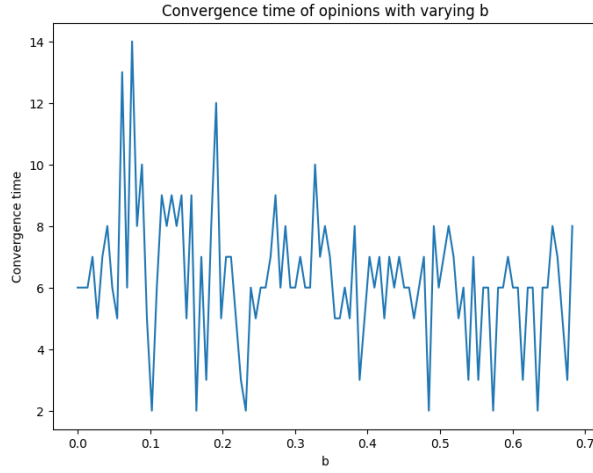


Figure 8: Convergence time of Hegselmann-Krause model, with varying inter-community connectivity

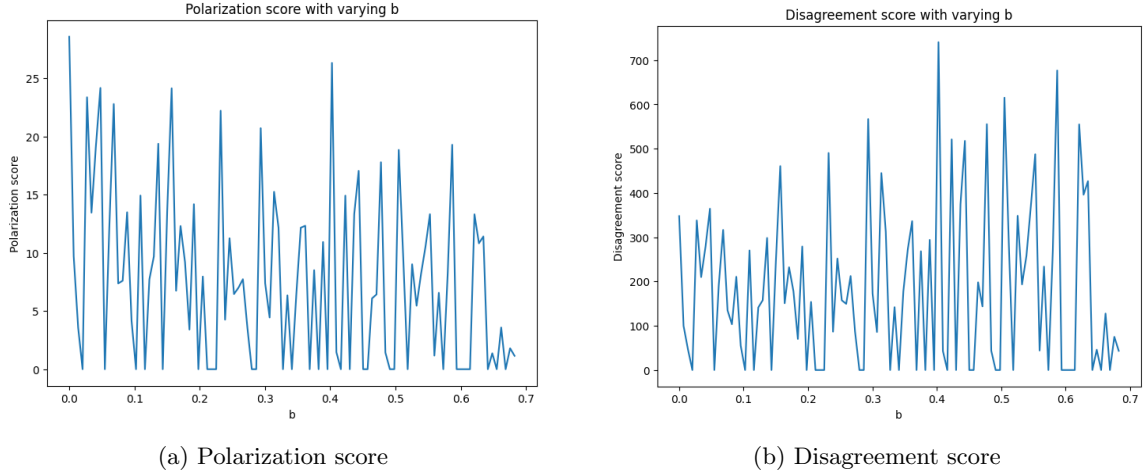
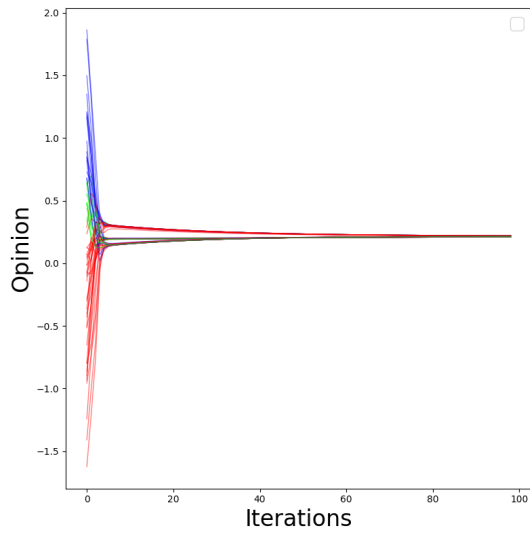
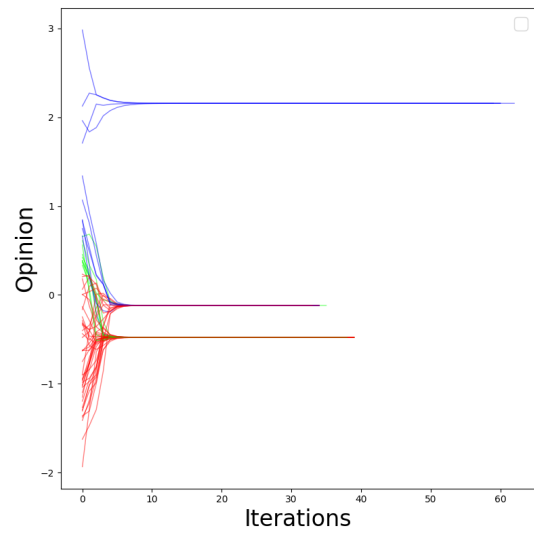


Figure 9: Polarization and disagreement scores of steady opinions following Hegselmann-Krause model, with varying inter-community connectivity





(a)  $\text{SBM}(3, 60, 10 \ln(60)/60, 0.1 \ln(60)/60)$



(b)  $\text{SBM}(3, 60, 10 \ln(60)/60, 5.9 \ln(60)/60)$

Figure 10: Opinion evolutions of Hegselmann-Krause model in two graphs with different inter-cluster connectivity; it is observed that as the different clusters are better connected, the agents achieve consensus

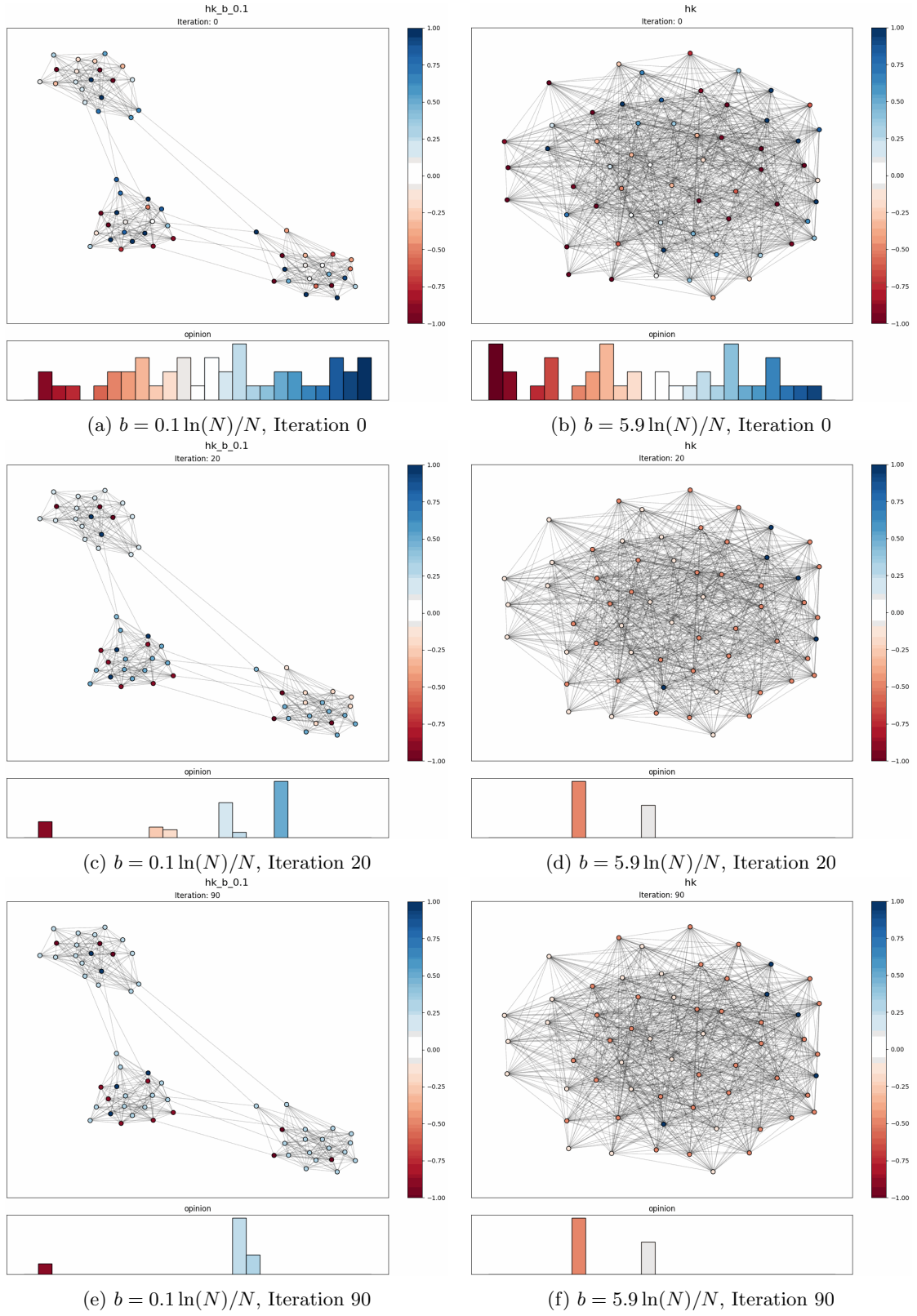


Figure 11: Comparing Hegselmann-Krause opinion dynamics on two graphs with different inter-cluster connectivity:  $b = 0.1 \ln(N)/N$  vs.  $b = 5.9 \ln(N)/N$

## 4 Conclusion

In this project, we investigate the opinion dynamics on social networks, in particular via two models Friedkin-Johnsen and Hegselmann-Krause. We first survey a random graph for social network modeling, then briefly discuss the two opinion dynamics models together with some of their theoretical properties. Then, using `NetworkX` and `NDlib`, we run synthetic simulations with the two mentioned models.

In the synthetic experiments, using generated networks from stochastic block models, we investigate the relationship between the opinion dynamics and the network structure. For the Friedkin-Johnsen models, we observe that as the different communities in a network are better connected, the polarization and disagreement of steady-state opinions decrease, and vice versa. However, for Hegselmann-Krause model, we did not observe a clear pattern in polarization or disagreement scores in terms of community structure. We speculate that the reason behind this observation is due to emergence of stubborn agents, of when and why we leave for future works.

## References

- [1] Thomas Fujiwara, Karsten Müller, and Carlo Schwarz. “The Effect of Social Media on Elections: Evidence from the United States”. In: *Journal of the European Economic Association* (Oct. 2023). ISSN: 1542-4766. DOI: 10.1093/jeea/jvad058.
- [2] Daron Acemoglu and Asuman Ozdaglar. “Opinion Dynamics and Learning in Social Networks”. In: *Dynamic Games and Applications* 1.1 (Oct. 2010). DOI: 10.1007/s13235-010-0004-1.
- [3] Elisabetta Biondi. *Analysis of opinion dynamics over a realistic dynamic social network – report of a successful TNA experience*. Sept. 2023. URL: <http://sobigdata.eu/blog/analysis-opinion-dynamics-over-realistic-dynamic-social-network-report-successful-tna> (visited on 11/01/2023).
- [4] Cameron Musco, Christopher Musco, and Charalampos E. Tsourakakis. “Minimizing Polarization and Disagreement in Social Networks”. In: *Proceedings of the 2018 World Wide Web Conference*. WWW ’18. 2018, pp. 369–378. ISBN: 9781450356398. DOI: 10.1145/3178876.3186103.
- [5] Liwang Zhu, Qi Bao, and Zhongzhi Zhang. “Minimizing Polarization and Disagreement in Social Networks via Link Recommendation”. In: *Advances in Neural Information Processing Systems*. Vol. 34. 2021, pp. 2072–2084. URL: [https://proceedings.neurips.cc/paper\\_files/paper/2021/file/101951fe7ebe7bd8c77d14f75746b4bc-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2021/file/101951fe7ebe7bd8c77d14f75746b4bc-Paper.pdf).
- [6] Miklos Z. Rácz and Daniel E. Rigobon. “Towards Consensus: Reducing Polarization by Perturbing Social Networks”. In: *IEEE Transactions on Network Science and Engineering* 10.6 (2023), pp. 3450–3464. DOI: 10.1109/TNSE.2023.3262970.
- [7] Noah E. Friedkin and Eugene C. Johnsen. “Social influence and opinions”. In: *The Journal of Mathematical Sociology* 15.3-4 (1990), pp. 193–206. DOI: 10.1080/0022250X.1990.9990069.
- [8] Rainer Hegselmann and Ulrich Krause. “Opinion Dynamics and Bounded Confidence: Models, Analysis and Simulation”. In: *Journal of Artificial Societies and Social Simulation* 5.3 (2002), p. 2. URL: <https://jasss.soc.surrey.ac.uk/5/3/2.html>.
- [9] Paul W. Holland, Kathryn Blackmond Laskey, and Samuel Leinhardt. “Stochastic blockmodels: first steps”. In: *Social Networks* 5.2 (1983), pp. 109–137. ISSN: 0378-8733. DOI: 10.1016/0378-8733(83)90021-7.
- [10] Adrian E Raftery Peter D Hoff and Mark S Handcock. “Latent Space Approaches to Social Network Analysis”. In: *Journal of the American Statistical Association* 97.460 (2002), pp. 1090–1098. DOI: 10.1198/016214502388618906.
- [11] Karl Rohe, Sourav Chatterjee, and Bin Yu. “Spectral clustering and the high-dimensional stochastic blockmodel”. In: *The Annals of Statistics* 39.4 (Aug. 2011). ISSN: 0090-5364. DOI: 10.1214/11-aos887. URL: <http://dx.doi.org/10.1214/11-AOS887>.
- [12] Morris H. Degroot. “Reaching a Consensus”. In: *Journal of the American Statistical Association* 69.345 (Mar. 1974), pp. 118–121. DOI: 10.1080/01621459.1974.10480137.
- [13] Arun G. Chandrasekhar, Horacio Larreguy, and Juan Pablo Xandri. “Testing Models of Social Learning on Networks: Evidence From Two Experiments”. In: *Econometrica* 88.1 (2020), pp. 1–32. ISSN: 0012-9682. DOI: 10.3982/ecta14407.

- [14] Benjamin Golub and Matthew O. Jackson. “Naïve Learning in Social Networks and the Wisdom of Crowds”. In: *American Economic Journal: Microeconomics* 2.1 (Feb. 2010), pp. 112–49. DOI: 10.1257/mic.2.1.112.
- [15] Elchanan Mossel and Omer Tamuz. “Opinion exchange dynamics”. In: *Probability Surveys* 14.none (2017), pp. 155–204. DOI: 10.1214/14-PS230.
- [16] Antonio F. Peralta, János Kertész, and Gerardo Iñiguez. *Opinion dynamics in social networks: From models to data*. 2022. arXiv: 2201.01322 [physics.soc-ph].
- [17] Pranav Dandekar, Ashish Goel, and David T. Lee. “Biased assimilation, homophily, and the dynamics of polarization”. In: *Proceedings of the National Academy of Sciences* 110.15 (Mar. 2013), pp. 5791–5796. ISSN: 1091-6490. DOI: 10.1073/pnas.1217220110.
- [18] Guillaume Deffuant et al. “Mixing beliefs among interacting agents”. In: *Advances in Complex Systems* 03.01n04 (Jan. 2000), pp. 87–98. ISSN: 1793-6802. DOI: 10.1142/s0219525900000078.
- [19] Fan R. K. Chung. *Spectral Graph Theory*. American Mathematical Soc., 1997.