

# Evolution of Opinions with Estimation and Interference

Yifa Liu<sup>1,2</sup>, Long Cheng<sup>1,2\*</sup>

1. State Key Laboratory for Management and Control of Complex Systems, Institute of Automation, Chinese Academy of Sciences,  
Beijing 100190, China  
E-mail: long.cheng@ia.ac.cn

2. School of Artificial Intelligence, University of Chinese Academy of Sciences, Beijing 100049, China

**Abstract:** The evolution of opinion is largely influenced by individual subjectivity, however, this is neglected in many studies about opinion dynamics. This paper analyzes the evolution of self-cognition and its impact on the spread of opinion from the perspective of individual judgment. Furthermore, this paper considers the effect of external interference on the cluster by altering social interaction, and discusses the impact of demagoguery on individual self-cognition bias and polarization of the cluster.

**Key Words:** Opinion dynamics, Self-cognition, Estimation

## 1 Introduction

There is no more complex system in the world than human beings. And the most complex system of human is the human mind [1]. Accordingly, human cognition is so extremely complex that it has become more incomprehensible than the cognition of human cognition. Due to the complexity of human cognition, a great number of scientists attempt to understand how people think and how people's opinions transfer from various perspectives [2, 3]. And opinion dynamics have also attracted more and more attention [4].

However, as a complex interdisciplinary subject, opinion dynamics is difficult to analyze and verify through intuitive physical experiments. Not to mention the emergence phenomenon when the opinions of different individuals continuously evolve according to individual behavior and interaction with other individuals [5]. And ultimately, the attitude of the whole society which is constituted by the opinions of all individuals often does not match rational schemes and predictions.

### 1.1 Related Works

Although there are many studies have investigated opinion dynamics in different ways, most studies worked backward from the results to the evolutionary patterns of opinions.

The DeGroot model proposed in 1974 is the earliest and most basic model in the field of opinion dynamics [6]. The DeGroot model assumes an individual's opinion at the next moment becomes the weighted average of its own opinion and the opinions of neighboring individuals in the network, where the weights are given by this individual. Under this model, regardless of the initial population's opinions, there is always a consensus in the end. However, this situation of complete consensus is consistent with actual observations, and thus, there is a lot of studies to explain the reason. In [7], each individual's paranoia about its initial beliefs is taken into account, namely only part of each individual's beliefs can be changed. This model is known as the Friedkin-Johnsen model. Personal paranoia about own opinion may not be always related to the original opinion. In [8], the self-persistence degree is introduced to improve the DeG-

root model to describe stubborn individuals. The higher the self-persistence degree, the more difficult it is for the individual to change its ideas. The bias phenomenon is studied in [9], where the entrenchment parameter is introduced to model the bias and backfire effect, and therefore this model allows the polarization. In many studies, the evolution of opinions is deterministic, resulting in no change when reaches the steady state. Some studies consider the effects of uncertainty and noise [10–13]. Most studies on the opinion dynamics focus on a specific opinion, however, two corresponding opinions on similar or related things may influence each other [15], and thus the individual's opinions on different things are considered. In previous studies, there is only one idea of an individual and it is expressed in social networks. The expression of opinions is studied in [16], where the phenotype reflected in people's herd mentality is not certainly those their inner approval of. The Hegselmann-Krause model considers the situation of whether individuals interact by introducing confidence bound [14], where two adjacent individuals do not exchange opinions if their beliefs are very different. Therefore, the Hegselmann-Krause model is nonlinear. In addition to studying how opinions evolve, there are also studies on taking advantage of the laws of opinion dynamics. It is pointed out that those who are stubborn and unwilling to change their minds are more likely to become leaders in the network [17]. And how to facilitate the crowd to reach a specific consensus by adding the fewest edges in the human-machine network is investigated in [18].

### 1.2 Contributions

Many existing studies have analyzed the evolution of the opinions of individuals from different perspectives, however, the opinion is taken for granted as an entity in these models. And many studies about opinion dynamics does not involve individuals' feeling, which may be too heartless. Opinions are not objective but subjective, and the response and action of the individual play an important role in the evolution of its idea. This issue is ignored in most studies.

Therefore, this paper analyzes the possible attitude of an individual to the interaction of ideas in social networks from the perspective of individual self-cognition and behavior. And the investigation of self-cognition can also explain the inconsistency between phenotype and actual inner idea.

This paper first considers individual judgment of curren-

This work was supported in part by the National Natural Science Foundation of China (Grants 61633016 61873268 62025307 U1913209) and the Beijing Natural Science Foundation (Grant JQ19020).

t cognition through integrating the historical cognition and the interactions with other individuals, and it is proved that this fusion approach is effective. Then this paper analyzes that the individuals can reach an agreement with their neighbors and the entire cluster eventually reaches a consensus in the case of sincere exchanges of opinions between individuals. It is further shown that in the presence of external interference, the continuous reinforcement of cognitive shift through the interaction of the affected individuals may eventually lead to the deviation of the individuals' self-cognition and the polarization of the cluster. By introducing the self-cognition and biased decision, this paper presents a possible explanation for the phenomenon of polarization in opinion dynamics.

**Notation:**  $\mathbf{1}$  and  $\mathbf{0}$  denote the vector with all ones and all zeros, respectively.  $I$  and  $\mathbf{O}$  denote the identity matrix and the matrix with all zeros, respectively.  $\mathcal{D}(M)$  denotes the degree matrix of matrix  $M$ , i.e.  $\mathcal{D}([m_{ij}]_{\zeta \times \zeta}) = \text{diag}(\sum_{j=1}^{\zeta} m_{1j}, \dots, \sum_{j=1}^{\zeta} m_{\zeta j})$ .  $\vec{e}_i$  denotes the  $i$ -th column of the identity matrix.

## 2 Opinion Dynamics Model in Social Network

This section first introduces the model of cluster and its social networks, then presents the evolution model of the individual opinion with subjective cognition and social behavior. This paper mainly focuses on the discrete-time opinion evolution problem.

### 2.1 Cluster Social Network Structure and Opinion Drift Model

Considering a social network composed of  $n$  closely connected individuals, the evolution of individual opinion involves the influence of its neighbors and their subjective changes.

First, the following assumptions are adopted.

**Assumption 1.** Individual interactions are bidirectional, i.e., edges in the social network are undirected.

**Rationality to Assumption 1.** This paper considers regular individuals and their behaviors. For parts with a large range of influence, such as celebrities, media, presidents, etc., their interferences are modeled separately later. In addition, social interaction is usually two-way in real society, and it is difficult for one individual not affected by another but only outputs influence in the interaction.

**Assumption 2.** This social network is connected, namely there is a path between any two individuals.

**Rationality to Assumption 2.** On the one hand, according to Six Degrees of Separation [19], the vast majority of individuals in the world can establish connections through a path within 7 steps. On the other hand, even ignoring the weak interaction between individuals in reality, this paper discusses a relatively isolated cluster, which is interconnected within this cluster, such as a community. If the clusters are not connected, the cluster is split into multiple connected clusters. If an individual in the cluster has an interactive relationship with an individual outside, the individual outside is included in the cluster until the cluster has no out-of-cluster nodes with which it significantly interacts.

**Assumption 3.** The evolution of an individual's opinion is potentially influenced by the neighbors' opinion.

**Rationality to Assumption 3.** It is believed in political economy that the nature of man is the sum of all social relations. The development and evolution of individuals are affected by the interaction of social networks everywhere.

This social network is modeled as a weighted graph  $(\mathcal{V}, \mathcal{E}, \mathcal{A})$ , where  $\mathcal{V} = (1, 2, \dots, N)$  denotes the set of all individuals which describes this cluster,  $v_i$  represents the  $i$ -th individual;  $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$  denotes the set of interactive relationship,  $e_{ij} = (i, j) \in \mathcal{E}$  if and only if the  $j$ -th individual has a considerable influence on the  $i$ -th individual;  $\mathcal{A} = [a_{ij}] \in \mathbb{R}^{N \times N}$  denotes the weighted adjacency matrix,  $a_{ii} > 0$ ,  $a_{ij} \geq 0$  and  $a_{ij} = 0$  if  $e_{ij} \notin \mathcal{E}$ .  $\mathcal{N}_i = \{j \in \mathcal{V}, e_{ij} \in \mathcal{E}\}$  denotes the set of neighborhood of  $i$ -th node.

Under these assumptions, a discrete-time evolution model of individual opinions is presented as

$$x_i[k+1] = a_{ii}x_i[k] + \sum_{j \in \mathcal{N}_i} a_{ij}x_j[k] + b_i u_i[k] + w_i[k], \quad (1)$$

where  $x_i \in \mathbb{R}^n$  and  $u_i \in \mathbb{R}^p$  are the opinion and action of the  $i$ -th individual at time  $k$ ,  $a_{ii} + \sum_{j \in \mathcal{N}_i} a_{ij} = I$ . Since  $x_i$  may be a multidimensional vector,  $a_{ij}$  can be a matrix.  $w_i$  is i.i.d Gaussian noise vector,  $w_i[k] \sim N(0, Q_i)$ ,  $Q_i, i \in \mathcal{V}$  is a small positive definite matrix. And the opinion evolution of the entire cluster can be expressed by

$$X[k+1] = \mathcal{A}X[k] + \mathcal{B}U[k] + W[k], \quad (2)$$

where  $X = [x_1^T \ x_2^T \ \dots \ x_n^T]^T$ ,  $\mathcal{B} = [b_1^T \ b_2^T \ \dots \ b_n^T]^T$ ,  $W = [w_1 \ w_2 \ \dots \ w_n]$ .  $(\mathcal{A}, \mathcal{B})$  is controllable.

In most related studies,  $n = 1$ ,  $\sum_{j=1}^n a_{ij} = 1$ ,  $u_i = 0$ . For the sake of simplification, this paper also sets  $n = 1$ , correspondingly,  $p = 1$  and  $b_i > 0$ .

**Remark 1.** Some studies assume that only individuals with similar opinions communicate with each other, and individuals with very different opinions do not interact. In this model, this can be achieved by adjusting the adjacency matrix to reduce these interact effects. In addition, that two closely interacting individuals do not affect each other at all may be too ruthless in reality.

**Assumption 4.** The world is changing slowly, i.e.,  $\|\mathcal{A} - I\| < \epsilon_1$ ,  $\|\mathcal{B}\| < \epsilon_2$ , where  $\epsilon_1, \epsilon_2$  are sufficiently small positive value. Matrices  $\mathcal{A}$  and  $\mathcal{B}$  are time-invariant.

**Rationality to Assumption 4.** This assumption may be contrary to the normal worldview. Currently, people may always complain that the world is changing so drastically. However, the time scales adopted by these two statements are not uniform. Even if there is a lot of variation from year to year, there may not be many variations from day to day. If this cannot be satisfied, it can be divided into smaller units such as hours. In short, man adjusts faster than the world.

### 2.2 Social Interaction between Individuals

Many studies about opinion dynamics are result-oriented, focusing on the outcome of interactions between individuals rather than the process of interactions, which leads to the neglect of the interaction sections. However, the polarization

of ideas, including false ideas such as rumors, is actually achieved through repeated reinforcement in the social interaction process.

Interactions can include social behaviors, evaluations, and so on. Providing that the main social behavior occurs between two individuals, this paper ignores the scene where more than two people interact or split it into multiple 1-to-1 interactions.

**Assumption 5.** *Each individual authentically presents itself, i.e.,  $\mathbb{E}(y_i - x_i) = 0$ .*

**Rationality to Assumption 5.** *Although the expressed opinions are proposed and studied in [16], long-term successful camouflage in continuous social interaction requires superb skills and decisive courage. This paper only discusses the sincere scenario, and the possibility of camouflage may be investigated in future work. The later part gives a possible explanation for the difference between phenotype and true value through a cognitive shift.*

The output of the interaction of the  $i$ -th individual with the  $j$ -th individual is modeled as

$$z_{ij}[k] = h_{ij}^i x_i[k] + h_{ij}^j x_j[k] + v_{ij}[k], \quad e_{ij} \in \mathcal{E}, \quad (3)$$

where  $z_{ij} \in \mathbb{R}^m$ ,  $m \geq n$  is the interaction result between individuals  $i$  and  $j$ ,  $h_{ij}^i \in \mathbb{R}^{m \times n}$  and  $h_{ij}^j \in \mathbb{R}^{m \times n}$  are full column rank. Since  $n$  is set to 1,  $h_{ij}^{iT} h_{ij}^i > 0$ .  $v_{ij}$  is the *i.i.d.* Gaussian noise vector,  $v_{ij}[k] \sim N(0, R_{ij})$ ,  $R_{ij}, e_{ij} \in \mathcal{E}$  is a small positive definite matrix.

Besides, each individual can portray his own image as

$$y_i[k] = x_i[k] + \mu_i[k], \quad (4)$$

where  $y_i \in \mathbb{R}^n$  is presentation of the  $i$ -th individual, namely the image it maintains in the public,  $\mu_i$  is the *i.i.d.* Gaussian noise,  $\mu_i[k] \sim N(0, S_i)$ ,  $S_i$  is a small positive definite matrix.

### 3 Opinion Transformation under the Impacts of Cognition and Action

Section 2 discusses the basic patterns of individual opinion evolution and interaction, however, many questions remain undetermined. This section aims at analyzing the judgment and decision-making from the perspective of an individual.

Many studies model the interaction between individuals directly into the change of opinions, which leads to the roughness of the interaction description, where the opinions are input as an object to the evolution process. Although it is difficult to evaluate and quantify human subjectivity, individual cognition and behavioral decision-making can greatly affect opinion evolution. If an individual's cognition of itself is biased, which is actually reasonable in present observations of the post-truth era, the trajectory of its opinion cannot be as ideal as expected.

**Assumption 6.** *The hypothesis of rational man is adopted.*

**Rationality to Assumption 6.** *In order to analyze the problem within a controllable framework, this paper adopts the hypothesis of rational man, which is an assumption often adopted in economics. It is believed that most human actions are still reasonable.*

### 3.1 Self-Cognition

Before adopting the adjustment strategy to deal with the changing world, a rational individual should first recognize the world and itself. This subsection discusses how a rational individual achieves effective self-cognition.

Denote  $\hat{x}_i[k]$  as the  $i$ -th individual's cognition of itself at time  $k$  and  $\hat{X} = [\hat{x}_1^T \hat{x}_2^T \cdots \hat{x}_n^T]^T$ .

Consider an extremely confident individual who disregards any social events and believes in its judgment and unaffectedness. It is noted that  $a_{ij}, i \neq j$  is not accessible to individuals. Then the law of its self-cognition evolution can be derived by

$$\hat{x}_{ii}[k+1] = \left( \sum_{j=1}^n a_{ij} \right) \hat{x}_i[k] + b_i u_i[k], \quad (5)$$

where  $\hat{x}_{ii}[k]$  is the current cognition inference at time  $k$  made by the  $i$ -th individual based on its historical cognition. It is noted that this individual does not think that others affect it, though psychological research has proved that even a stubborn individual is still influenced by others.

Conversely, consider another extremely unconfident individual, which is similar to the one discussed in [6]. This individual is easier to accept others' comments and interaction results, and tends to acquire self-recognition from the interaction process. The law of its self-cognition evolution can be derived by

$$\hat{x}_{ij}[k] = (h_{ij}^{iT} h_{ij}^i)^{-1} h_{ij}^{iT} (z_{ij}[k] - h_{ij}^j y_j[k]), \quad j \in \mathcal{N}_i, \quad (6)$$

where  $\hat{x}_{ij}[k]$  is the self appraisal of the  $i$ -th individual after socializing with the  $j$ -th individual.

It seems that both cases are unusual in practice. For the former, it is difficult to integrate into society and easy to get cognitive bias. For the latter, nice guys finish last. Therefore, under the rationality-first hypothesis, the following fusion estimation method is adopted for individuals to realize the cognitions

$$\hat{x}_i[k] = \omega_{ii} \hat{x}_{ii}[k] + \sum_{j \in \mathcal{N}_i} \omega_{ij} \hat{x}_{ij}[k], \quad (7)$$

where  $\omega_{ij}$  is weight of the  $i$ -th individual to the  $j$ -th individual,  $\sum_{j=1}^N \omega_{ij} = 1$ ,  $0 < \omega_{ij} < 1$  if  $j \in \mathcal{N}_i \cup \{i\}$ , otherwise,  $\omega_{ij} = 0$ .

The cognitive error can be measured by the difference between the estimated value of its beliefs and the true value  $e_i = x_i - \hat{x}_i$ ,  $E = [e_1^T e_2^T \cdots e_n^T]^T$ . The following theorem points out that under the assumptions of this paper, each individual can control its cognitive error within an acceptable range through long-time observation.

**Theorem 1.** *When no adjustment action is taken, each individual can finally get an unbiased cognition of his own opinion through continuous social interaction, namely math expectation the cognitive error of rational people is zero.*

*Proof.* See the proof of Theorem 1 in Appendix A.  $\square$



### 3.2 Opinion Drift

Subsection 3.1 analyzes the cognitive behavior of rational individuals. After cognition, individuals need to take countermeasures according to the actual situation. Then how should an individual respond rationally is discussed in this subsection.

Since the cluster taken in this paper is a closely related group, the following control law to identify with other individuals is adopted,

$$u_i[k] = K \sum_{j \in \mathcal{N}_i} \omega_{ij} (y_j[k] - \hat{x}_i[k]), \quad (8)$$

where  $K < 1 / \min_{i=1, \dots, n} \{b_i\}$ .

**Theorem 2.** *By adopting (8), the cluster with rational individuals can reach a dynamic agreement after long-term evolution, i.e.,  $E(x_i - x_j) \rightarrow 0, t \rightarrow \infty$ .*

*Proof.* See the proof of Theorem 2 in Appendix B.  $\square$

### 3.3 Opinion and Cognitive Evolution in Transitive Deception

Subsection 3.2 discusses the evolution logic of opinion within a cluster in the absence of external interference. However, the individuals in this cluster are all ordinary, and the model of the individual is not suitable for things with wide influence such as the media and the former president.

Therefore, this subsection considers the existence of an element with widespread influence outside of this cluster. Its interference on the cluster is more reflected in the daily interactions between clusters. For example, people may often talk about something, such as rumors, to reinforce cognition. If both are the target audiences, they trust the rumor more.

Many studies model this element as a leader, which causes all individuals to eventually rush towards this leader under the condition of connectivity. However, in fact, each individual may not necessarily reach this leader. For example, most people's access to false information is often second-hand and above, and even then, the wide spread and negative impact of fake news are prevented.

Interactions between affected individuals are described as

$$z_{ij}^a[k] = h_{ij}^i x_i[k] + h_{ij}^j x_j[k] + v_{ij}[k] + \alpha_{ij}[k], e_{ij} \in \mathcal{E}, \quad (9)$$

where  $\alpha_{ij}[k]$  is the interference exerted in the interaction between the  $i$ -th individual and the  $j$ -th individual. This paper adopts only one kind of effective external interference, it is assumed that  $h_{ij}^{iT} \alpha_{ij} \geq 0$  and there is at least one(two)  $h_{ij}^{iT} \alpha_{ij} > 0$ .

Under these circumstances, the persistent demagogic effect can deviate individuals' self-cognitions and then trigger biased decisions, which eventually leads to the polarization of the cluster.

**Theorem 3.** *When external interference can always act on the cluster, the general opinion of this cluster can eventually be polarized, i.e.,  $\mathbf{1}^T x[t] |_{\{\sum_t \sum_{i,j} h_{ij}^{iT} \alpha_{ij}[t] \rightarrow \infty\}} \rightarrow \infty$ .*

*Proof.* See the proof of Theorem 3 in Appendix C.  $\square$

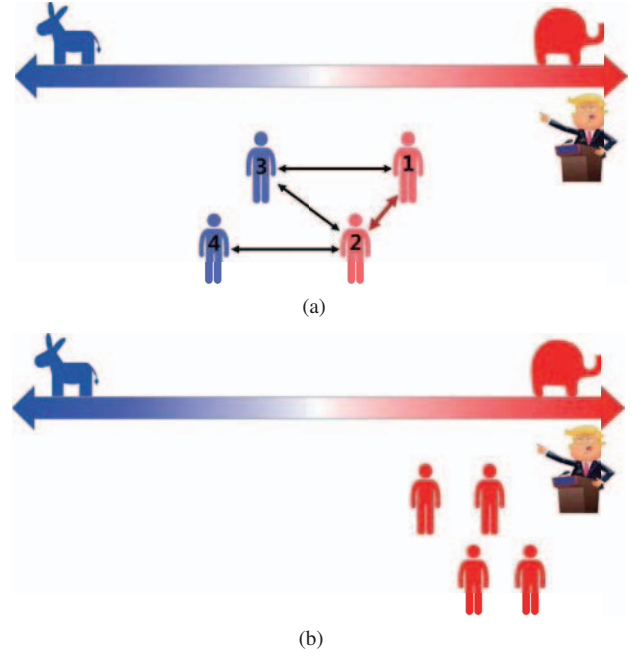


Fig. 1: Interactions between individuals of this a cluster: (a) social network topology; (b) outcome prediction.

## 4 Numerical Simulation

This section is to verify the correctness of the theorems, namely whether individuals can get the right self-cognition and reach a consensus and whether this cluster polarized by external interference.

Consider four closely connected individuals, of which individuals 1 and 2 are conservatives and individuals 3 and 4 are liberals but not staunch. Their interaction relationship is shown in Fig. 1.

The weighted matrix is shown as follows

$$\Omega = \begin{bmatrix} 0.5 & 0.25 & 0.25 & 0 \\ 0.2 & 0.5 & 0.1 & 0.2 \\ 0.25 & 0.25 & 0.5 & 0 \\ 0 & 0.5 & 0 & 0.5 \end{bmatrix}.$$

The initial opinion values are  $x_1[0] = 2, x_2[0] = 1, x_3[0] = -1, x_4[0] = -2$ , where higher positive values represent higher conservatism leanings and negative values with greater absolute values represent higher liberalism leanings. Positive infinity represents the far right. The relevant parameters are set to  $\mathcal{A} = I, b_i = 1, K = 0.1$ , and all noises are set to  $\sim N(0, 0.1^2)$ .

Although former President Trump constantly held rallies, his rallies are far from enough to radiate all the rednecks. Most rednecks interact with other rednecks in daily life, and they encourage and inspire each other.

In the first scenario without additional interference, estimation strategy (7) and action strategy (8) are adopted. Figure 2 shows the evolution processes of the opinion and cognition of each individual. It can be seen that all individuals achieve moderate points of view and correct cognitions.

In the second scenario, suppose there is a conservative pioneer who gives the conservative supporter the creed. Then the interaction between the conservative individuals 1 and 2 is influenced

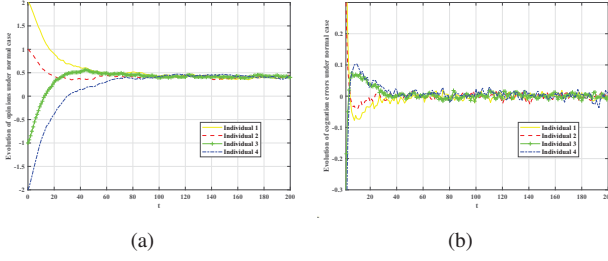


Fig. 2: Evolutions of opinions and cognition errors under normal case: (a) opinions reach a consensus; (b) cognition errors converge to 0.

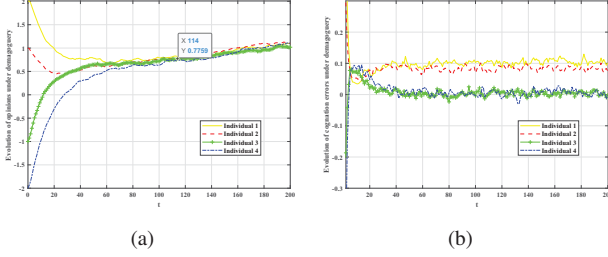


Fig. 3: Evolutions of opinions and cognition errors under demagoguery: (a) opinions are being polarized; (b) there exist cognition errors of affected individuals.

$$z_{12}^a = z_{12} + 0.1, \quad z_{21}^a = z_{21} + 0.1,$$

that is, the affected conservatives are more willing to encourage each other and call for action to support MAGA. Other conditions remain the same. Figure 3 shows the evolution processes of the opinion and cognition in this case. It can be seen that the opinion of clusters is constantly being polarized, and it is noted that affected individuals 1 and 2 have biased self-cognition while other individuals do not.

## 5 Conclusion

Human cognition of himself or herself is largely hindered by their own subjectivity, resulting in the cognition being easily biased. However, if everyone sincerely accepts the views of others, the entire society is developing in the direction of harmony. Though in dynamic equilibrium, the society is stable, until the demagoguery. A rumor does not need to be repeatedly released, it only needs to enter into social communication. And then in social circulation, this rumor is repeatedly strengthened and eventually causes considerable prejudice. If these kinds of lies are constantly made up, society will be polarized. Correspondingly, tailored lies to different clusters can lead to the tearing of social.

## Appendix A: The Proof of Theorem 1

*Proof.* Substitute two estimation strategies (5) and (6) into weighted fusion estimation method (7), the final evolution

equation of individual cognition can be obtained as

$$\begin{aligned} \hat{x}_i[k+1] &= \omega_{ii} \left( \sum_{j=1}^n a_{ij} \hat{x}_j[k] + b_i u_i[k] \right) \\ &+ \sum_{j \in \mathcal{N}_i} \omega_{ij} (h_{ij}^{iT} h_{ij}^i)^{-1} h_{ij}^{iT} (z_{ij}[k+1] - h_{ij}^j y_j[k+1]) \\ &= \omega_{ii} \sum_{j=1}^n a_{ij} \hat{x}_j[k] + \omega_{ii} b_i u_i[k] + \sum_{j \in \mathcal{N}_i} \omega_{ij} x_j[k+1] \\ &+ \sum_{j \in \mathcal{N}_i} \omega_{ij} (h_{ij}^{iT} h_{ij}^i)^{-1} h_{ij}^{iT} (v_{ij}[k+1] - h_{ij}^j \mu_j[k+1]) \\ &= \omega_{ii} \hat{x}_i[k] + (\omega_{ii} + \sum_{j \in \mathcal{N}_i} \omega_{ij}) b_i u_i[k] + \sum_{j \in \mathcal{N}_i} \omega_{ij} a_{ii} x_i[k] \\ &+ (\sum_{j \in \mathcal{N}_i} \omega_{ij}) (\sum_{j \in \mathcal{N}_i} a_{ij} x_j[k]) + \sum_{j \in \mathcal{N}_i} \omega_{ij} w_i[k] \\ &+ \sum_{j \in \mathcal{N}_i} \omega_{ij} [(h_{ij}^{iT} h_{ij}^i)^{-1} h_{ij}^{iT} (v_{ij}[k+1] - h_{ij}^j \mu_j[k+1])]. \end{aligned}$$

Then together with (1), the cognitive error can be derived by

$$\begin{aligned} e_i[k+1] &= x_i[k+1] - \hat{x}_i[k+1] \\ &= a_{ii} x_i[k] + \sum_{j \in \mathcal{N}_i} a_{ij} x_j[k] + b_i u_i[k] + w_i[k] \\ &- \omega_{ii} \hat{x}_i[k] - b_i u_i[k] - \sum_{j \in \mathcal{N}_i} \omega_{ij} a_{ii} x_i[k] \\ &- (\sum_{j \in \mathcal{N}_i} \omega_{ij}) (\sum_{j \in \mathcal{N}_i} a_{ij} x_j[k]) - \sum_{j \in \mathcal{N}_i} \omega_{ij} w_i[k] \\ &- \sum_{j \in \mathcal{N}_i} \omega_{ij} [(h_{ij}^{iT} h_{ij}^i)^{-1} h_{ij}^{iT} (v_{ij}[k+1] - h_{ij}^j \mu_j[k+1])] \\ &= \omega_{ii} [a_{ii} (x_i[k] - \hat{x}_i[k]) + \sum_{j \in \mathcal{N}_i} a_{ij} (x_j[k] - \hat{x}_i[k]) + w_i[k]] \\ &+ \sum_{j \in \mathcal{N}_i} \omega_{ij} [(h_{ij}^{iT} h_{ij}^i)^{-1} h_{ij}^{iT} (h_{ij}^j \mu_j[k+1] - v_{ij}[k+1])] \\ &= \omega_{ii} e_i[k] + \omega_{ii} \sum_{j \in \mathcal{N}_i} a_{ij} (x_j[k] - x_i[k]) + \omega_{ii} w_i[k] \\ &+ \sum_{j \in \mathcal{N}_i} \omega_{ij} [(h_{ij}^{iT} h_{ij}^i)^{-1} h_{ij}^{iT} (h_{ij}^j \mu_j[k+1] - v_{ij}[k+1])]. \end{aligned}$$

First ignoring the effects of all noises, the evolution of opinion and cognitive error for the entire cluster can be obtained,

$$\begin{aligned} \begin{bmatrix} X[k+1] \\ E[k+1] \end{bmatrix} &= \begin{bmatrix} \mathcal{A}X[k] \\ \Xi(\mathcal{A} - \mathcal{D}(\mathcal{A}))X[k] + \Xi E[k] \end{bmatrix} \\ &= \begin{bmatrix} \mathcal{A} & \mathbf{O} \\ \Xi(\mathcal{A} - \mathcal{D}(\mathcal{A})) & \Xi \end{bmatrix} \begin{bmatrix} X[k] \\ E[k] \end{bmatrix} \\ &\triangleq \Phi \begin{bmatrix} X[k] \\ E[k] \end{bmatrix} \end{aligned}$$

where  $\Xi = \text{diag}(\omega_{11}, \omega_{22}, \dots, \omega_{nn})$ .

Denote  $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$  as the eigenvalues of  $(\mathcal{A} - \mathcal{D}(\mathcal{A}))$ . From Gershgorin disc theorem and the connectivity of this cluster, it can be obtained that  $\lambda_1 > -1$ ,  $\lambda_n = 0$  and

all eigenvalues of  $\mathcal{A}$  belongs to  $(0, 1]$ . It is noted that the powers of  $\Phi$  are still the triangular block matrix,

$$\Phi^k = \begin{bmatrix} \mathcal{A}^k & \mathbf{O} \\ \sum_{i=0}^k \Xi^i [\Xi(\mathcal{A} - \mathcal{D}(\mathcal{A}))] \mathcal{A}^{k-i} & \Xi^k \end{bmatrix}$$

Then the evaluation of estimation error can be written as follows

$$\begin{aligned} & \begin{bmatrix} X[k+1] \\ E[k+1] \end{bmatrix} \\ = & \Phi \begin{bmatrix} X[k] \\ E[k] \end{bmatrix} + \begin{bmatrix} W[k] \\ \Xi W[k] \end{bmatrix} \\ & + \begin{bmatrix} \mathbf{0} \\ \sum_{i=1}^n [\tilde{e}_i^T \Omega (Dg_i^T Dg_i)^{-1} Dg_i^T (V_i + H_i^T \mathcal{M}_i)] \tilde{e}_i \end{bmatrix} \\ \triangleq & \Phi \begin{bmatrix} X[k] \\ E[k] \end{bmatrix} + \begin{bmatrix} N_1[k] \\ N_2[k] \end{bmatrix}, \end{aligned}$$

where  $Dg_i = \text{diag}(v_{i1}, v_{i2}, \dots, v_{in})$ ,  $V_i = [v_{i1}^T \dots v_{in}^T]^T$ ,  $H_i^T = [h_{i1}^1 \dots h_{in}^n]$ ,  $\mathcal{M}_i = [\mu_1 \dots \mu_n]$ .

$$\begin{bmatrix} X[k] \\ E[k] \end{bmatrix} = \Phi^k \begin{bmatrix} X[0] \\ E[0] \end{bmatrix} + \sum_{i=1}^{k-1} \Phi^{i-1} \begin{bmatrix} N_1[k-i] \\ N_2[k-i] \end{bmatrix}.$$

The final cognitive error can be obtained that

$$\begin{aligned} E[\infty] &= \lim_{k \rightarrow \infty} E[k] \\ &= \lim_{k \rightarrow \infty} \sum_{i=0}^k \Xi^{i+1} (\mathcal{A} - \mathcal{D}(\mathcal{A})) \mathcal{A}^{k-i} X[0] \\ &\quad + \lim_{k \rightarrow \infty} \Xi^k E[0] + \lim_{k \rightarrow \infty} \sum_{i=1}^{k-1} \Xi^{i-1} N_2[k-i] \\ &\quad + \lim_{k \rightarrow \infty} \sum_{i=1}^{k-1} \sum_{j=0}^k \Xi^{j+1} (\mathcal{A} - \mathcal{D}(\mathcal{A})) \mathcal{A}^{k-j} N_1[k-i] \end{aligned}$$

Since the absolute values of eigenvalues of  $(\mathcal{A} - \mathcal{D}(\mathcal{A}))$  and  $\Xi$  are less than 1 and eigenvalues of  $(\mathcal{A})$  are no more than 1, the mathematical expectation and variance of  $E[\infty]$  are bounded.  $\square$

## Appendix B: The Proof of Theorem 2

*Proof.* Applying action of individual (8) into (1) and (7) leads to the evolution equation of the individual

$$\begin{aligned} & \hat{x}_i[k+1] \\ = & \omega_{ii} \hat{x}_i[k] + b_i K \sum_{j \in \mathcal{N}_i} \omega_{ij} (y_j[k] - \hat{x}_i[k]) \\ & + (1 - \omega_{ii}) (a_{ii} x_i[k] + \sum_{j \in \mathcal{N}_i} a_{ij} x_j[k]) + (1 - \omega_{ii}) w_i[k] \end{aligned}$$

$$\begin{aligned} & + \sum_{j \in \mathcal{N}_i} \omega_{ij} [(h_{ij}^{iT} h_{ij}^i)^{-1} h_{ij}^{iT} (v_{ij}[k+1] - h_{ij}^j \mu_j[k+1])] \\ = & \omega_{ii} \hat{x}_i[k] + b_i K \sum_{j \in \mathcal{N}_i} \omega_{ij} (x_j[k] - \hat{x}_i[k]) \\ & + (1 - \omega_{ii}) (a_{ii} x_i[k] + \sum_{j \in \mathcal{N}_i} a_{ij} x_j[k]) + (1 - \omega_{ii}) w_i[k] \\ & + \sum_{j \in \mathcal{N}_i} \omega_{ij} [(h_{ij}^{iT} h_{ij}^i)^{-1} h_{ij}^{iT} [v_{ij}[k+1] + h_{ij}^i b_i K \mu_j[k] \\ & - h_{ij}^j \mu_j[k+1]]]. \end{aligned}$$

Still ignore noises first, and then opinion and cognition of this cluster can be derived by

$$\begin{aligned} & \begin{bmatrix} X[k+1] \\ \hat{X}[k+1] \end{bmatrix} \\ = & \begin{bmatrix} [\mathcal{A} + BK(\Omega - \Xi)] X[k] - BK(\mathcal{D}(\Omega) - \Xi) \hat{X}[k] \\ [(I - \Xi)\mathcal{A} + BK(\Omega - \Xi)] X[k] + [\Xi - BK(\mathcal{D}(\Omega) - \Xi)] \hat{X}[k] \end{bmatrix} \\ = & \begin{bmatrix} \mathcal{A} + BK(\Omega - \Xi) & -BK\mathcal{D}(\Omega - \Xi) \\ (I - \Xi)\mathcal{A} + BK(\Omega - \Xi) & \Xi - BK\mathcal{D}(\Omega - \Xi) \end{bmatrix} \begin{bmatrix} X[k] \\ \hat{X}[k] \end{bmatrix} \\ \triangleq & \Psi \begin{bmatrix} X[k] \\ \hat{X}[k] \end{bmatrix}. \end{aligned}$$

Denote

$$NN = \frac{1}{2n} \begin{bmatrix} \mathbf{1}\mathbf{1}^T & \mathbf{1}\mathbf{1}^T \\ \mathbf{1}\mathbf{1}^T & \mathbf{1}\mathbf{1}^T \end{bmatrix}.$$

It is noted that the following equations hold,

$$\begin{aligned} \Psi &= \begin{bmatrix} I & I \\ \mathbf{O} & I \end{bmatrix} \\ & \begin{bmatrix} \mathbf{O} & -\Xi \\ (I - \Xi)\mathcal{A} + BK(\Omega - \Xi - \mathcal{D}(\Omega - \Xi)\mathcal{A}) & \Xi - BK\mathcal{D}(\Omega - \Xi) \end{bmatrix} \\ & \begin{bmatrix} I & \mathbf{O} \\ -\mathcal{A} & I \end{bmatrix}, \\ \Psi \times NN &= NN, \\ NN \times NN &= NN, \\ NN &= \begin{bmatrix} I & I \\ \mathbf{O} & I \end{bmatrix} \begin{bmatrix} \mathbf{O} & \mathbf{O} \\ \frac{1}{n}\mathbf{1}\mathbf{1}^T & \frac{1}{2n}\mathbf{1}\mathbf{1}^T \end{bmatrix} \begin{bmatrix} I & \mathbf{O} \\ -\mathcal{A} & I \end{bmatrix}. \end{aligned}$$

Then the follows prove  $\Psi$  converges to  $NN$ .

$$\begin{aligned} & \lim_{k \rightarrow \infty} (\Psi - NN)^k \\ = & \begin{bmatrix} I & I \\ \mathbf{O} & I \end{bmatrix} \\ & \begin{bmatrix} \mathbf{O} & -\Xi \\ (I - \Xi)\mathcal{A} - \mathbf{1}\mathbf{1}^T/n & \Xi - BK\mathcal{D}(\Omega - \Xi) - \mathbf{1}\mathbf{1}^T/2n \\ +BK(\Omega - \Xi - \mathcal{D}(\Omega - \Xi)\mathcal{A}) & \end{bmatrix} \\ & \begin{bmatrix} I & \mathbf{O} \\ -\mathcal{A} & I \end{bmatrix}. \end{aligned}$$

If  $K < 1/\min\{b_i\}$ , the absolute values of eigenvalues of  $(I - \Xi)\mathcal{A} - \mathbf{1}\mathbf{1}^T/n + BK(\Omega - \Xi - \mathcal{D}(\Omega - \Xi)\mathcal{A})$  are less than 1, together with all elements of  $\Xi$  less than 1,  $\lim_{k \rightarrow \infty} (\Psi - NN)^k = \mathbf{O}$  and then  $\lim_{k \rightarrow \infty} \Psi^k = NN$ .

And by using the similar method in Appendix A, it can be proven that noises do not make the opinion diverge.

Finally, the opinions of the cluster will reach quasi-consensus, that is, fluctuate around the mean value.  $\square$

### Appendix C: The Proof of Theorem 3

*Proof.* Still ignore noises first, and then opinion and cognition of this cluster under demagoguery can be derived by

$$\begin{bmatrix} X[k+1] \\ \hat{X}[k+1] \end{bmatrix} = \Psi \begin{bmatrix} X[k] \\ \hat{X}[k] \end{bmatrix} + \begin{bmatrix} \mathbf{O} \\ \mathbf{H} \end{bmatrix} \text{vec}([\alpha_{ij}[k+1]]),$$

where  $\mathbf{H} = \text{diag}(H_1, H_2, \dots, H_n)$ ,  $H_i = [\omega_{i1} (h_{i1}^{iT} h_{i1}^i)^{-1} h_{i1}^{iT} \omega_{i2} (h_{i2}^{iT} h_{i2}^i)^{-1} h_{i2}^{iT} \dots \omega_{in} (h_{in}^{iT} h_{in}^i)^{-1} h_{in}^{iT}]$ .

$$\begin{bmatrix} X[k] \\ \hat{X}[k] \end{bmatrix} = \Psi^k \begin{bmatrix} X[0] \\ \hat{X}[0] \end{bmatrix} + \sum_{t=1}^k \Psi^{k-t} \begin{bmatrix} \mathbf{O} \\ \mathbf{H} \end{bmatrix} \text{vec}([\alpha_{ij}[t]]),$$

$$\begin{aligned} & \lim_{k \rightarrow \infty} \left( \begin{bmatrix} X[k+1] \\ \hat{X}[k+1] \end{bmatrix} - \begin{bmatrix} X[k] \\ \hat{X}[k] \end{bmatrix} \right) \\ &= \lim_{k \rightarrow \infty} \left\{ \sum_{t=1}^k (\Psi^{k+1-t} - \Psi^{k-t}) \begin{bmatrix} \mathbf{O} \\ \mathbf{H} \end{bmatrix} \text{vec}([\alpha_{ij}[t]]) \right. \\ & \quad \left. + \begin{bmatrix} \mathbf{O} \\ \mathbf{H} \end{bmatrix} \text{vec}([\alpha_{ij}[k+1]]) \right\}. \end{aligned}$$

It is noted that  $\mathbf{H} \text{vec}([\alpha_{ij}])$  is semi-positive definite matrix, and therefore

$$\begin{aligned} & \lim_{k \rightarrow \infty} [\mathbf{1}^T \mathbf{1}^T] \left( \begin{bmatrix} X[k+1] \\ \hat{X}[k+1] \end{bmatrix} - \begin{bmatrix} X[k] \\ \hat{X}[k] \end{bmatrix} \right) \\ & \geq [\mathbf{1}^T \mathbf{1}^T] \begin{bmatrix} \mathbf{O} \\ \mathbf{H} \text{vec}([\alpha_{ij}[k+1]]) \end{bmatrix} > 0. \end{aligned}$$

Then when  $k \rightarrow \infty$ , together with the conclusion in Appendix B, it can be derived by

$$\begin{aligned} & \mathbf{1}^T X[2k+j] |_{\{\alpha_{ij}[t] \equiv 0, k+j < t \leq 2k+j\}} \\ & \geq \mathbf{1}^T [I \mathbf{O}] N N \begin{bmatrix} X[k+1+j] \\ \hat{X}[k+1+j] \end{bmatrix} > \mathbf{1}^T X[k+j], \quad j \ll k. \end{aligned}$$

Since  $\mathbf{H}$  is constant, by repeating this process and from the arbitrariness of  $j$ , it can be obtained that

$$\mathbf{1}^T X[2k+j] > \mathbf{1}^T X[k+j],$$

and finally

$$\lim_{n \rightarrow \infty, k \rightarrow \infty} \mathbf{1}^T X[nk+j] = \infty, \quad \lim_{t \rightarrow \infty} \mathbf{1}^T X[t] = \infty.$$

Then take the impact of noises into account. From Appendix B, the covariance matrix of noises is converge, then, the sum of all noise effects is bounded at any moment. Therefore, the general opinion of this cluster is eventually polarized.  $\square$

### References

- [1] G. H. Lewes, Problems of Life and Mind. London: Trübner & Company, 1877.
- [2] H. Noorazar, K. R. Vixie, A. Talebanpour, and Y. Hu, From classical to modern opinion dynamics, *International Journal of Modern Physics*, 31(7): 2050101, 2020.
- [3] B. D. O. Anderson and M. Ye, Recent advances in the modelling and analysis of opinion dynamics on influence networks, *International Journal of Automation and Computing*, 16(2): 129–149, 2019.

- [4] L. Wang, Y. Tian, and J. Du, Opinion dynamics in social networks, *Scientia Sinica Informationis*, 48: 3–23, 2018.
- [5] M. M. Waldrop, Complexity: the Emerging Science at the Edge of Order and Chaos. New York: Simon and Schuster, 1993 Holland.
- [6] M. H. DeGroot, Reaching a consensus, *Journal of the American Statistical Association*, 69(345): 118–121, 1974.
- [7] N. Friedkin and E. Johnsen, Social influence networks and opinion change models of opinion formation, *Advances in Group Processes*, 16: 1–29, 1999.
- [8] Q. Zhou, Z. Wu, A. H. Altalhi, and F. Herrera, A two-step communication opinion dynamics model with self-persistence and influence index for social networks based on the DeGroot model, *Information Sciences*, 519: 363–381, 2020.
- [9] X. Chen, P. Tsaparas, J. Lijffijt, and T. Bie, Opinion dynamics with backfire effect and biased assimilation, *PLoS ONE*, 16(9): e0256922, 2021.
- [10] E. Kurmyshev, H. A. Jurez, R. A. Gonzalez-Silva, Dynamics of bounded confidence opinion in heterogeneous social networks: Concord against partial antagonism, *Physica A: Statistical Mechanics and its Applications*, 390(16): 2945–2955, 2011.
- [11] A. Carro, R. Toral, and M. S. Miguel, The role of noise and initial conditions in the asymptotic solution of a bounded confidence, continuous opinion model, *Journal of Statistical Physics*, 151(12): 131–149, 2013.
- [12] W. Su, G. Chen, and Y. Hong, Noise leads to quasi-consensus of hegselmannKrause opinion dynamics, *Automatica*, 85: 448–454, 2017.
- [13] G. Chen, W. Su, S. Ding, and Y. Hong, Heterogeneous Hegselmann-Krause dynamics with environment and communication noise, *IEEE Transactions on Automatic Control*, 65(8): 3409–3424, 2020.
- [14] R. Hegselmann, U. Krause, Opinion dynamics and bounded confidence models, analysis, and simulation, *Journal of Artificial Societies and Social Simulation*, 5: 1–24, 2002.
- [15] F. Xiong, Y. Liu, L. Wang, and X. Wang, Analysis and application of opinion model with multiple topic interactions, *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 27(8): 083113, 2017.
- [16] M. Ye, Y. Qin, A. Govaert, B. D. O. Anderson, and M. Cao, An influence network model to study discrepancies in expressed and private opinions, *Automatica*, 107: 371–381, 2019.
- [17] W. Jager, F. Amblard, Uniformity, bipolarization and pluriformity captured as generic stylized behavior with an agent-based simulation model of attitude change, *Computational & Mathematical Organization Theory*, 10(4): 295–303, 2005.
- [18] Y. Dong, Z. Ding, L. Martínez, and F. Herrera, Managing consensus based on leadership in opinion dynamics, *Information Sciences*, 397: 187–205, 2017.
- [19] J. Travers and S. Milgram, An experimental study of the small world problem, *Sociometry*, 32(4): 425–443, 1969.