Preprints of the 19th World Congress
The International Federation of Automatic Control
Cape Town, South Africa. August 24-29, 2014

# A Multi-Agent Model of Opinion Formation with Truth Seeking and Endogenous Leaders ⋆

**Qipeng Liu, Jiuhua Zhao, Lin Wang, and Xiaofan Wang**

*Department of Automation, Shanghai Jiao Tong University, and Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, China (e-mail: {sjtu_lqp,jiuhuadandan,wanglin,xfwang}@sjtu.edu.cn)*

**Abstract:** We present an opinion formation model in which individuals are supposed to seek a consensus on the truth. All agents are boundedly confident in the sense that an agent communicates with another one if their opinion difference is not greater than a threshold which is called the bound of confidence. When facing the truth, they are also boundedly confident such that they take the truth into account if the truth is inside their bound of confidence. Any agent influenced by the truth is viewed as a leader, and thus, the role of leader is endogenous with the evolution of opinions. For any pair of bound of confidence and position of truth, we provide a possible range of the strength of the attraction of truth that can guarantee the whole group converging to the truth. We also find that there always exists a suitable strength of the attraction of truth leading the whole group to a consensus on the truth as long as the whole group converge to a consensus in the absence of the attraction of truth.

Keywords: Multi-agent systems; opinion formation; consensus; endogenous leaders.

## 1. INTRODUCTION

Multi-agent models are widely used to study the opinion dynamics in social networks, such as the formation of consensus (DeGroot [1974]), information spreading (Acemoglu et al. [2010]), and wisdom of crowds (Golub and Jackson [2010]). Multi-agent models allow us to investigate the relationship between local interactions and global behavior at a level of higher resolution compared with the models studied in the field of statistical mechanics (Sznajd-Weron and Sznajd [2000], Slanina and Lavicka [2003], Clifford and Sudbury [1973], Holley and Liggett [1975]), most of which can trace back to the Ising model in statistical physics. Generally, multi-agent models are formulated as dynamical systems, and both transient state and steady state can be studied (see e.g., Olfati-Saber et al. [2007], Touri and Nedić [2011], Chen et al. [2013]).

Whether opinions of all agents reach a consensus in the long run is one of the questions frequently mentioned in most opinion formation models. A consensus is very important in the situations such as negotiating, consulting, and making decisions. In some other situations, we might not be satisfied with just reaching a consensus, but also expect the consensus value to be a desired one. For example, when a decision has its payoff, we are supposed to make an optimal decision which has the best payoff; the governments or the big companies would make use of the mass media to lead the public opinions reaching a consensus in favor of their own benefits. To model these desired-value seeking problems, an external or global influence is usually introduced besides the local interactions. In Hegselmann and Krause [2006], a constant in the opinion space, representing the value of truth, is introduced to the original Hegselmann-Krause model (HK model for short). The influence of local interactions and the global influence of truth are merged together in the form of convex combination. A similar schedule is examined in the Deffuant model (Malarz [2006]). Other mathematical models are also proposed in the context of consensus tracking in the field of control theory (Ren [2008, 2007]).

The agents observing or influenced by the external or global signals are regarded as informed agents or leaders, and correspondingly, other agents are uninformed agents or followers. In the literature, the role of leader or follower is usually assigned exogenously in the sense that it is appointed by outside environment or system designers rather than the system itself (Molavi and Jadbabaie [2011], Liu and Wang [2013]). The leaders might be chosen randomly, and the role is fixed during the evolution of opinions. In this paper, we present a truth seeking model with endogenous leaders. All agents are set to be boundedly confident in the sense that each agent only communicates with those whose opinion difference is not greater than a given threshold, call the bound of confidence. In addition, there exists a truth in the opinion space, and all agents are confronted with the attraction of truth in a bounded-confidence way: an agent is influenced by the truth (becomes a leader) if and only if her opinion differs from the truth not greater than her bound of confidence. It is

⋆ This work is supported by the National Natural Science Foundation of China under Grant Nos. 61074125 and 61104137, the Science Fund for Creative Research Groups of the National Natural Science Foundation of China (No. 61221003), and the National Key Basic Research Program (973 Program) of China (No. 2010CB731403).

obvious that rather than assign the role of leaders among the agents exogenously, we let the leaders emerge automatically once certain conditions are satisfied. In respect of the opinion updating rule, the leaders add the influence of truth to the average opinions of neighbors in the form of convex combination, while the followers update their opinions simply by averaging opinions of neighbors. If at the equilibrium all agents reach a consensus on the truth, we call it a correct consensus. Our work is to seek the conditions under which the correct consensus can be achieved. We find that there always exists a suitable strength of the attraction of truth resulting in a correct consensus as long as the whole group reach a consensus in the absence of the influence of truth.

The remainder of this paper is organized as follows: In Section 2, our opinion formation model is presented. In Section 3, theoretical analyses and main results are given. The paper ends by concluding remarks and open questions.

## 2. MODEL DESCRIPTION

### *2.1 Opinions and Bounded Confidence*

Consider a set of agents as $V = \{1, 2, \cdots, n\}$, and the opinion of agent $i$ at time $t$ is denoted by $x_i(t) \in [0, 1]$. The initial opinions $\{x_i(0), i \in V\}$ are set to be uniformly distributed in the interval [0,1].

Each agent is endowed a bound of confidence $\epsilon$ such that any pair of agents trust and communicate with each other only if their opinion difference is not greater than $\epsilon$. All the agents who can communicate with agent $i$ compose her neighbor set $N_i(t)$, i.e., $N_i(t) = \{j \in V : |x_i(t) - x_j(t)| \le \epsilon\}$.

### *2.2 The Truth and Endogenous Leaders*

The truth is represented by a constant in the opinion space, denoted by $T \in [0, 1]$. Any agent $i$ is attracted by the truth only when $T$ is inside her bound of confidence, i.e., $|x_i(t) - T| \le \epsilon$. These agents are called *leaders* in the sense that they have the potential to lead the whole group to the truth. Correspondingly, other agents are called *followers*. It is easy to see that the role of leaders is endogenous during the evolution of opinions.

### *2.3 Updating Rule*

At each time step $t$, each agent $i$ updates her opinion simply by averaging the opinions of her neighbors if the truth is out of her bound of confidence; otherwise the attraction of truth is considered. In summary, the updating rule can be described as follows:

$$x_i(t+1) = \alpha_i(t)T + (1 - \alpha_i(t)) \sum_{j\in N_i(t)} \frac{1}{|N_i(t)|} x_j(t), \quad (1)$$

where $\alpha_i(t) > 0$ if agent $i$ is a leader at time $t$, and $\alpha_i(t) = 0$ otherwise; $|S|$ is the cardinality of the corresponding set $S$. For simplicity, we assume that the truth has a common strength of attraction on all leaders, i.e., $\alpha_i(t) = \alpha \in (0, 1]$ if $\alpha_i(t) > 0$.

*Remark 1:* The models studied in the literature most related to ours are those in Hegselmann and Krause [2006] and Carletti et al. [2006]. In Hegselmann and Krause [2006], a mathematical model similar to (1) is proposed, while the role of leaders are exogenous, i.e., either all agents are leaders or some agents are chosen randomly in advance, and the role is fixed all the time. In Carletti et al. [2006], an opinion formation model with endogenous leaders is presented based on the original Deffuant model (Deffuant et al. [2002]), and the attraction of truth ("propaganda" in their work) periodically appears.

*Remark 2:* It is worth pointing out that the rule (1) is not intentionally followed by the agents. If they realize $T$ is the truth, then they will converge to it immediately, regardless of the opinions of other agents. The interpretation of (1) should be that some signals reflecting the truth to a certain extent can be observed, and accordingly, the agents update their opinions based on the noisy observations (see Hegselmann and Krause [2006] for more detailed discussion). Here we blind out the philosophical discussions and focus on the dynamical system governed by (1).

Even though the updating rule (1) is similar to the formal HK model, and even more similar to the truth-seeking model in Hegselmann and Krause [2006], they have essential difference. In the HK model, the question of interest is whether the opinions converge to a consensus, while, once the truth and opinion leaders are introduced, the question of interest becomes whether the opinions influenced by the leaders can reach the pre-defined truth. Some of our main results provide the range of influence of truth in which a consensus on truth can be achieved. The main difference between the model with endogenous leaders and that with exogenous leaders is that the role of leaders is time-varying based on the evolution of opinions, which makes the performance of the model quite different (e.g., order-preserving property shown in the following).

## 3. THEORETICAL ANALYSES AND MAIN RESULTS

### *3.1 Order-Preserving Property*

The order-preserving property of the original HK model is well known and proved in both Krause [2000] and Blondel et al. [2009]. Once leaders or heterogenous agents are introduced, the order-preserving property might not hold anymore. For instance in Hegselmann and Krause [2006] exogenous leaders are introduced to the HK model, and it is shown that the opinions of agents might change their order as the opinions evolves. In our work, leaders are endogenous, and the following proposition shows that, unlike the truth-seeking model with exogenous leaders, the updating rule (1) preserves the order of opinions.

*Proposition 1.* For any pair of agents updating opinions according to (1), the order of their opinions is preserved, i.e., for all $t \ge 0$, if $x_i(t) \le x_j(t)$, then $x_i(t+1) \le x_j(t+1)$.

**Proof.** Suppose that $x_i(t) \le x_j(t)$. Let $\widetilde{N}_i(t)$ be the set of agents only communicating with $i$ and not $j$, $\widetilde{N}_j(t)$ the set of agents only communicating with $j$ and not $i$, and $\widetilde{N}_{ij}(t)$ the set of agents communicating with both $i$ and $j$. For any $k_1 \in \widetilde{N}_i(t)$, $k_2 \in \widetilde{N}_{ij}(t)$, and $k_3 \in \widetilde{N}_j(t)$, we have $x_{k_1}(t) \le x_{k_2}(t) \le x_{k_3}(t)$. Thus, $\overline{x}_{\widetilde{N}_i(t)} \le \overline{x}_{\widetilde{N}_{ij}(t)} \le \overline{x}_{\widetilde{N}_j(t)}$, where $\overline{x}_{\widetilde{N}_i(t)}$, $\overline{x}_{\widetilde{N}_{ij}(t)}$, and $\overline{x}_{\widetilde{N}_i(t)}$, respectively, is the average of opinions in the corresponding set. Next we discuss the order-preserving property of (1) in different situations classified by different positions of $T$.

(a) *$T$ is outside the bound of confidence of both $i$ and $j$.*

Due to $\alpha_i(t) = \alpha_j(t) = 0$ in (1), we have

$$x_i(t+1) = \frac{|\widetilde{N}_i(t)|\overline{x}_{\widetilde{N}_i(t)} + |\widetilde{N}_{ij}(t)|\overline{x}_{\widetilde{N}_{ij}(t)}}{|\widetilde{N}_i(t)| + |\widetilde{N}_{ij}(t)|} \leq \overline{x}_{\widetilde{N}_{ij}(t)} \quad (2)$$

and

$$x_j(t+1) = \frac{|\widetilde{N}_j(t)|\overline{x}_{\widetilde{N}_j(t)} + |\widetilde{N}_{ij}(t)|\overline{x}_{\widetilde{N}_{ij}(t)}}{|\widetilde{N}_j(t)| + |\widetilde{N}_{ij}(t)|} \geq \overline{x}_{\widetilde{N}_{ij}(t)} \quad (3)$$

Thus, $x_i(t+1) \leq x_j(t+1)$.

(b) *$T$ is inside the bound of confidence of $i$ and not $j$.*

In this situation, $\alpha_i(t) = \alpha > 0$ and $\alpha_j(t) = 0$. For any $k \in \widetilde{N}_{ij}(t)$, we have $T \leq x_k(t)$. Therefore, $T \leq \overline{x}_{\widetilde{N}_{ij}(t)}$. By (1) we have

$$\begin{aligned} x_i(t+1) &= \alpha T + (1-\alpha)\frac{|\widetilde{N}_i(t)|\overline{x}_{\widetilde{N}_i(t)} + |\widetilde{N}_{ij}(t)|\overline{x}_{\widetilde{N}_{ij}(t)}}{|\widetilde{N}_i(t)| + |\widetilde{N}_{ij}(t)|} \\ &\leq \alpha T + (1-\alpha)\overline{x}_{\widetilde{N}_{ij}(t)} \\ &\leq \overline{x}_{\widetilde{N}_{ij}(t)} \end{aligned}$$

Due to $\alpha_j(t) = 0$, we have $x_j(t+1) \geq \overline{x}_{\widetilde{N}_{ij}(t)}$ by (3). Therefore, $x_i(t+1) \leq x_j(t+1)$.

(c) *$T$ is inside the bound of confidence of $j$ and not $i$.*

Here $\alpha_i(t) = 0$ and $\alpha_j(t) = \alpha > 0$. For any $k \in \widetilde{N}_{ij}(t)$, we have $T \geq x_k(t)$. Therefore, $T \geq \overline{x}_{\widetilde{N}_{ij}(t)}$. By (1) we have

$$\begin{aligned} x_j(t+1) &= \alpha T + (1-\alpha)\frac{|\widetilde{N}_i(t)|\overline{x}_{\widetilde{N}_j(t)} + |\widetilde{N}_{ij}(t)|\overline{x}_{\widetilde{N}_{ij}(t)}}{|\widetilde{N}_j(t)| + |\widetilde{N}_{ij}(t)|} \\ &\geq \alpha T + (1-\alpha)\overline{x}_{\widetilde{N}_{ij}(t)} \\ &\geq \overline{x}_{\widetilde{N}_{ij}(t)} \end{aligned}$$

In addition, $x_i(t+1) \leq \overline{x}_{\widetilde{N}_{ij}(t)}$ by (2) since $\alpha_i(t) = 0$. Therefore, $x_i(t+1) \leq x_j(t+1)$.

(d) *$T$ is inside the bound of confidence of both $i$ and $j$.*

In this case, we have

$$\begin{aligned} & x_j(t+1) - x_i(t+1) \\ &= (1-\alpha)\left(\frac{|\widetilde{N}_j|\overline{x}_{\widetilde{N}_j} + |\widetilde{N}_{ij}|\overline{x}_{\widetilde{N}_{ij}}}{|\widetilde{N}_j| + |\widetilde{N}_{ij}|} - \frac{|\widetilde{N}_i|\overline{x}_{\widetilde{N}_i} + |\widetilde{N}_{ij}|\overline{x}_{\widetilde{N}_{ij}}}{|\widetilde{N}_i| + |\widetilde{N}_{ij}|}\right) \\ &\geq 0 \end{aligned}$$

Considering (a), (b), (c), and (d), we obtain that $x_i(t+1) \leq x_j(t+1)$ regardless of the position of $T$. Because of the arbitrarily of $t$, we have that, for all $t \geq 0$, if $x_i(t) \leq x_j(t)$, then $x_i(t+1) \leq x_j(t+1)$. Thus, the order of opinions is preserved.

### *3.2 Influence of the Strength of the Attraction of Truth on Truth Seeking*

Most studies on opinion formation models focus on whether all agents reach a consensus. This question is also very important to our model. However, since we have introduced the influence of truth, the opinions of agents are expected not only reaching a consensus but also converging to the truth. Therefore, we propose a new concept: *reaching a correct consensus*, which means all agents collectively converge to the truth. For a given number of agents and uniformly distributed initial opinions, whether a correct consensus can be achieved depends on three factors:

- the bound of confidence, $\epsilon$;
- the position of truth, $T$;
- the strength of the attraction of truth, $\alpha$.

Our goal is to find the possible range of $\alpha$ under any given $T$ and $\epsilon$ to obtain a correct consensus. To deal with this issue, we set up a $(T, \epsilon)$-plane, where for any given point in the plane, $\alpha$ is the only factor affecting the performance of the system.

To obtain a glance of the $(T, \epsilon)$-plane, we perform a simulation on the model with $n = 100$. In the matter of fact, according to simulations with different number of agents, the layout in the $(T, \epsilon)$-plane does not have apparent difference as the number of agents changes. Figure 1 shows the upper bound of $\alpha$ for a correct consensus in the $(T, \epsilon)$-plane. Due to the symmetry, we only need to show the half part of the $(T, \epsilon)$-plane. Here we choose the part with $T \leq 0.5$. The upper bound of $\alpha$ equal to one means that, at the corresponding point in the $(T, \epsilon)$-plane, a correct consensus can be achieved, and $\alpha$ can be as large as one. The upper bound of $\alpha$ equal to zero means that we cannot find a suitable $\alpha$ to lead the whole group to the truth. Figure 2 shows the lower bound of $\alpha$ for a correct consensus. Here the lower bound of $\alpha$ being very close to zero means that the lower bound of $\alpha$ for a correct consensus at the corresponding point in the $(T, \epsilon)$-plane can be extremely small, and $\alpha = 1$ means that we cannot find a suitable $\alpha$ to lead the whole group to the truth.

By checking the upper bound and lower bound shown in Fig. 1 and Fig. 2, respectively, we obtain a clear division of the $(T, \epsilon)$-plane shown in Fig. 3, which is a full picture of the $(T, \epsilon)$-plane. There exist four zones, and in each of

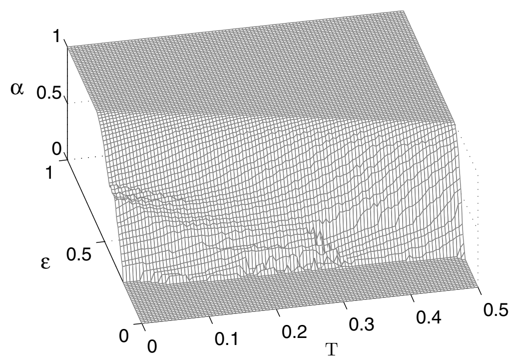


Fig. 1. The upper bound of $\alpha$ for a correct consensus.

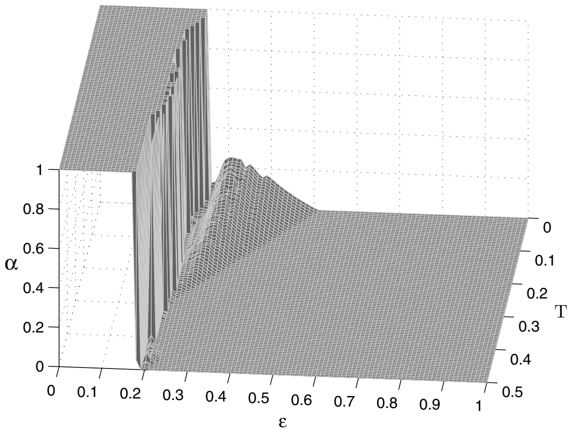


Fig. 2. The lower bound of $\alpha$ for a correct consensus.

them we have a different range of $\alpha$ to obtain a correct consensus. In Zone I, we can set $\alpha$ to be any value in the interval (0,1]. In the rest part of the plane, at least an upper bound constrains the value of $\alpha$. In Zone II, there exists no lower bound for $\alpha$. The strength of the attraction of truth in the zones other than Zone I and Zone II at least needs a lower bound and an upper bound to lead the whole group to the truth. In Zone III, we can still find a suitable $\alpha$ to achieve this goal, while in Zone IV there is no suitable $\alpha$.

Now we introduce a parameter, the critical bound of confidence $\epsilon_c$, which is the smallest bound of confidence keeping the whole group reaching a consensus (not necessarily a correct consensus) without the influence of $T$. An analytical expression of $\epsilon_c$ for any given initial condition has not been established yet. Here we obtain $\epsilon_c$ by simulations. It is shown in Fig. 4 that the approximate value of $\epsilon_c$ is 0.225 when $n = 100$.

Let us turn back to the $(T, \epsilon)$-plane shown in Fig. 3. The dashed line represents the critical bound of confidence, which is near the boundary between Zone III and Zone IV. That is to say, the critical bound of confidence separates Zone IV from other parts. We know that Zone IV is the only region that we cannot find a suitable $\alpha$ to achieve a correct consensus. This implies that *there always exists a suitable strength of the attraction of truth to lead the*

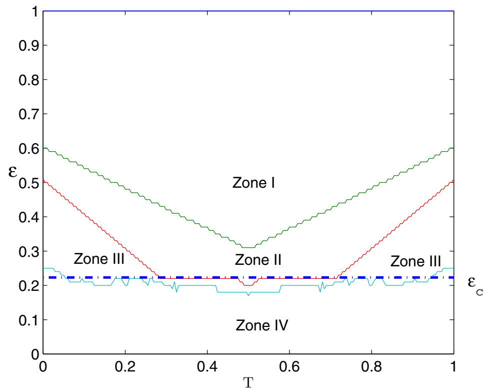


Fig. 3. Four zones in the $(T, \epsilon)$-plane.

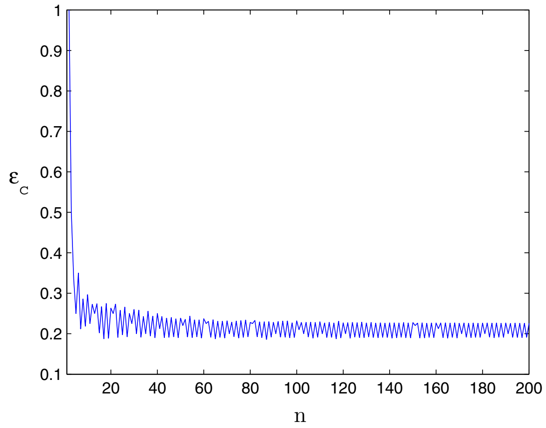


Fig. 4. The critical bound of confidence as a function of the number of agents.

*opinions of all agents to the truth as long as the whole group reach a consensus in the absence of the influence of truth.*

### *3.3 Theoretical Analysis about the $(T, \epsilon)$-Plane*

An opinion formation model with bounded confidence has a strong coupling between opinions and neighboring relationships. The dynamical property of the system under certain conditions might be quite subtle and unpredictable. Therefore, many studies on the bounded-confidence models usually resort to computer simulations. In this section, we theoretically explain the positions of Zone I and Zone II in the $(T, \epsilon)$-plane, while leave the positions of Zone III and Zone IV unsolved. Considering the high complexity of the model, especially the involvement of endogenous leaders, our results, although not quite complete, are still meaningful, and might shed a light on future work.

We first explain the position of Zone I where the upper bound of $\alpha$ can be one.

*Proposition 2.* In the case of $\alpha = 1$, the condition of $\epsilon \geq \frac{3}{5}\max\{T, 1-T\}$ is necessary and sufficient for a correct consensus.

**Proof.** Note that the region of $\epsilon \geq \frac{3}{5}\max\{T, 1-T\}$ is symmetric with respect to $T = 1/2$. Thus, we only focus on the half part with $T \leq 1/2$, and the result on the other part is straightforward. In addition, we only discuss the evolution of opinions inside the interval $[T, 1]$, and the analysis can be applied to the interval $[0, T]$ directly.

Let us start at the initial time period. Due to $\alpha = 1$, any agent $i$ satisfying $x_i(0) - T \leq \epsilon$ must converge to $T$ immediately at the next time step, i.e., $x_i(1) = T$. For the agent $i$ with the smallest opinion satisfying $x_i(0) - T > \epsilon$, if $1 - x_i(0) \geq \epsilon$, her opinion does not change at the next time step, i.e., $x_i(1) = x_i(0)$, since the average of the opinions of her neighbors is exactly her current opinion. According to the order-preserving property of the updating rule (1) we have proved in Proposition 1, all the opinions of her neighbors at $t = 1$ are greater than hers, and thus, all the agents with opinion greater than $x_i(1)$ are separated from the influence of truth for all $t \geq 1$.

Thus a correct consensus fails. To avoid this situation, the smallest opinion $x_i(0)$ satisfying $x_i(0) - T > \epsilon$ must be greater than $1 - \epsilon$, and thus, it moves towards $T$ rather than stays unchanged. Under this situation, we also have $T + 2\epsilon < 1$.

We next investigate the distribution of the opinions at $t = 1$. For simplicity, we index the agents based on the order of their initial opinions, i.e., $x_1(0) \geq x_2(0) \geq x_3(0) \geq \cdots \geq x_n(0)$. Due to the uniform distribution of the initial opinions, $x_1(1) = 1 - \epsilon/2$ [1]. Furthermore, we have

$$\begin{aligned} x_2(1) &= 1 - \frac{1}{n-1} - (\epsilon - \frac{1}{n-1})/2 = 1 - \frac{\epsilon}{2} - \frac{1}{2(n-1)} \\ x_3(1) &= 1 - \frac{2}{n-1} - (\epsilon - \frac{2}{n-1})/2 = 1 - \frac{\epsilon}{2} - \frac{2}{2(n-1)} \\ &\vdots \\ x_k(1) &= 1 - \frac{k-1}{n-1} - (\epsilon - \frac{k-1}{n-1})/2 = 1 - \frac{\epsilon}{2} - \frac{k-1}{2(n-1)} \end{aligned}$$

where $k = [1 - (T + \epsilon)](n - 1) + 1$ is the agent with the smallest opinion that does not converge to $T$ at $t = 1$. Thus,

$$x_k(1) = \frac{1}{2}(1 + T).$$

It is obvious that $\{x_1(1), x_2(1), \cdots, x_k(1)\}$ is uniformly distributed, and the distance between any two sequential opinions is $\frac{1}{2(n-1)}$. More importantly, we have

$$x_1(1) - x_k(1) = 1 - \frac{\epsilon}{2} - \frac{1}{2}(1 + T) = \frac{1}{2}(1 - T - \epsilon) < \frac{1}{2}\epsilon$$

which means that any pair of the $k$ agents are neighbors. We also have that all other agents are converge to $T$ at $t = 1$. Therefore, agents $\{1, 2, \cdots, k\}$ compose a cluster with no neighbors outside. At the time step $t = 2$, any agent $i$ with $x_i(1) - T \leq \epsilon$ converges to $T$, and other agents belonging to $\{1, 2, \cdots, k\}$ reach a consensus value

$$x^* = \frac{1}{2}[1 - \frac{\epsilon}{2} + \frac{1}{2}(1 + T)] = \frac{1}{4}(3 - T - \epsilon). \quad (4)$$

All the agents converge to the truth if and only if $x^* - T \leq \epsilon$. Thus, we have $\epsilon \geq \frac{3}{5}(1 - T)$.

By applying the similar argument to the opinion interval $[0, T]$, we have $\epsilon \geq \frac{3}{5}T$. Since $T \leq \frac{1}{2}$, we have $\epsilon \geq \frac{3}{5}(1 - T) \geq \frac{3}{5}T$. Therefore, we obtain the final range of $\epsilon$ as $\epsilon \geq \frac{3}{5}(1 - T)$.

As for the rest part of the $(T, \epsilon)$-plane, i.e., $T \geq \frac{1}{2}$, the proof is similar, and thus omitted.

Next we prove why the lower bound of $\alpha$ in the region of $\epsilon \geq \max\{|T - 1/2|, \epsilon_c\}$ shown in Fig. 3 is close to zero.

*Proposition 3.* For $\epsilon \geq \max\{|T - 1/2|, \epsilon_c\}$, there exits no lower bound for $\alpha$ which can lead the whole group converging to $T$.

**Proof.** To prove the statement, we only need to consider the extreme situation when $\alpha$ is small enough such that its effect can be neglected for a large enough number of steps. Then the whole group will evolve simply as if there exists no influence of $T$. Since $\epsilon \geq \epsilon_c$, the whole group must converge to a common value (not necessarily $T$) without the influence of $T$. In addition, due to the uniform distribution of initial opinions, the final common value must be 1/2. Obviously, after the single opinion cluster is established, it can be attracted by and converge to $T$ if and only if $\epsilon \geq |T - 1/2|$. Recalling that this result is based on $\epsilon \geq \epsilon_c$, we have $\epsilon \geq \max\{|T - 1/2|, \epsilon_c\}$.

In Fig. 3 we can see a concave on the bottom of Zone II around $T = 1/2$ . In fact this bottom line is very close to the critical bound of confidence, at which the dynamical system is quit sensitive to small changes of parameters. The opinions of the group might switch between division and consensus even the value of $\alpha$ changes slightly. Basically, the influence of $\alpha$ is quite irregular when the bound of confidence is near $\epsilon_c$, as we can see from the fluctuation of the boundary between Zone III and Zone IV. Even though the whole group converge to a consensus because of the influence of $T$, the common value generally is very close to the central value of initial opinions, which is 1/2 in our context. The final consensus opinion converges to the truth only when $T$ is close enough to 1/2; otherwise, a correct consensus fails.

At last, for Zone III and Zone IV, as we have mentioned above, their boundary is quite difficult to identify, and simulation is our only tool. As it is shown in Fig. 1 and Fig. 2, we can always find a suitable $\alpha$ to reach a correct consensus in Zone III, while in Zone IV the bound of confidence is too small that the truth only has local influence on the opinions of agents, which cannot change the division of the whole group.

## 4. CONCLUSIONS AND FUTURE WORKS

We have studied an opinion formation model with truth seeking and endogenous leaders. We mainly focus on the conditions under which all agents converge to the truth. More precisely, for any pair of bound of confidence and position of truth, we provide a possible range of the strength of the attraction of truth that can guarantee the whole group converging to the truth. We find that as long as the whole group reach a consensus in the absence of the influence of truth, there always exists a suitable strength of the attraction of truth resulting in a correct consensus, and its range of value changes with respect to the bound of confidence and the position of the truth. Although a part of our results is obtained by simulations, it is still instructive and shows the possibility of driving multiple agents to a desired state without any constraint on connectivity.

As for the important aspects of future work, on the one hand, a complete theoretical analysis is necessary; on the other hand, extending the model to the two-dimensional case is also very meaningful. It is worth pointing out that the two-dimensional HK model can be regarded as a rendezvous problem. Thus, extending our endogenous-leader model to the two-dimensional situation might provide some insight into the research on consensus tracking problems in the cooperative control of multi-agent systems.

[1] One might notice that the accurate value of $x_1(1)$ should be between $1 - \epsilon/2$ and $1 - (\epsilon - \frac{1}{n-1})/2$. Ignoring the influence of $\frac{1}{n-1}$ here does not affect our result apparently, and it releases the analysis from tedious details.

## REFERENCES

D. Acemoglu, A. Ozdaglar, and A. ParandehGheibi. Spread of (mis)information in social networks. *Games and Economic Behavior*, 70(2):194–227, 2010.

V. D. Blondel, J. M. Hendrickx, and J. N. Tsitsiklis. On Krause's multi-agent consensus model with state-dependent connectivity. *IEEE Transaction on Automatic Control*, 54(11):2586–2597, 2009.

T. Carletti, D. Fanelli, S. Grolli, and A. Guarino. How to make an efficient propaganda. *Europhysics Letters*, 74 (2):222–228, 2006.

Y. Chen, J. Lü, X. Yu, and D. J. Hill. Multi-agent systems with dynamical topologies: consensus and applications. *IEEE Circuits and Systems Magazine*, 13(3):21–34, 2013.

P. Clifford and A. Sudbury. A model for spatial conflict. *Biometrika*, 60(3):581–588, 1973.

G. Deffuant, D. Neau, F. Amblard, and G. Weisbuch. How can extremism prevail? A study based on the relative agreement interaction model. *Journal of Artificial Societies and Social Simulation*, 5(4), 2002.

M. DeGroot. Reaching a consensus. *Journal of the American Statistical Association*, 69(345):118–121, 1974.

B. Golub and M. Jackson. Naïve learning in social networks: convergence, influence, and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–149, 2010.

R. Hegselmann and U. Krause. Truth and cognitive division of labour: First steps towards a computer aided social epistemology. *Journal of Artifical Societies and Social Simulation*, 9(3), 2006.

R. A. Holley and T. M. Liggett. Ergodic theorems for weakly interacting infinite systems and the voter model. *The Annals of Probability*, 3(4):643–663, 1975.

U. Krause. A discrete nonlinear and non-autonomous model of consensus formation. In S. Elaydi, G. Ladas, J. Popenda, and J. Rakowski, editors, *Proc. of Communications in Difference Equations*, pages 227–236, London: Gordon and Breach, 2000.

Q. Liu and X. Wang. Social learning with bounded confidence and heterogeneous agents. *Physica A*, 392 (10):2368–2374, 2013.

K. Malarz. Truth seekers in opinion dynamics models. *International Journal of Modern Physics C*, 17(10): 1521–1524, 2006.

P. Molavi and A. Jadbabaie. Aggregate observational distinguishability is necessary and sufficient for social learning. In *Proc. of the 50th IEEE Conference on Decision and Control*, pages 1513–1519, Orlando, FL, USA, December 2011. IEEE.

R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1):215–233, 2007.

W. Ren. Multi-vehicle consensus with a time-varying reference state. *Systems and Control Letters*, 56(7–8): 474–483, 2007.

W. Ren. Consensu tracking under directed interaction topologies: Algorithms and experiments. In *Proc. of the American Control Conference*, pages 742–747, Seattle, Washington, USA, June 2008.

F. Slanina and H. Lavicka. Analytical results for the Sznajd model of opinion formation. *The European Physical Journal B*, 35(2):279–288, 2003.

K. Sznajd-Weron and J. Sznajd. Opinion evolution in closed community. *International Journal of Modern Physics C*, 11(6):1157–1165, 2000.

B. Touri and A. Nedić. On ergodicity, infinite flow, and consensus in random models. *IEEE Transactions on Automatic Control*, 56(7):1593–1605, 2011.