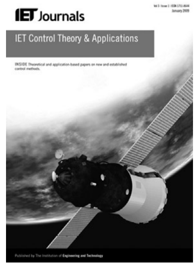
www.ietdl.org

Published in IET Control Theory and Applications
Received on 5th May 2014
Revised on 2nd August 2014
Accepted on 20th August 2014
doi: 10.1049/iet-cta.2014.0511

Special Issue on Co-operative Multi-Agent Systems with Engineering Applications



ISSN 1751-8644

Multi-agent model of group polarisation with biased assimilation of arguments

Qipeng Liu[1,2,3], *Jiuhua Zhao*[2,3], *Xiaofan Wang*[2,3]

[1] *Institution of Complexity Science, Qingdao University, Qingdao 266071, People's Republic of China*
[2] *Department of Automation, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China*
[3] *Key Laboratory of System Control and Information Processing, Ministry of Education of China, Shanghai 200240, People's Republic of China*
E-mail: sjtu_lqp@sjtu.edu.cn

Abstract: This study studies the opinion dynamics with group polarisation. The authors say a process of opinion formation is polarising if the final opinions after group discussion are more divergent than individuals' initial opinions. To capture this phenomenon, they develop a multi-agent model of opinion formation in which each agent forms their opinion according to all arguments available to them, and during interacting with other agents, they assimilate the arguments of neighbours in a biased manner which can be represented by a modified Polya's urn process. They find that biased assimilation of arguments can result in group polarisation effectively, especially under the condition when the agents are relatively isolated in the network.

1 Introduction

Collective behaviour is ubiquitous in nature and man-made systems, such as fish schooling, bird flocking, robots formation and sensor fusion [1–7]. Owing to its amazing appearance and manifold potential applications, collective behaviour has attracted significant attention and gradually become a hot research field in past decades.

As an important research area with respect to collective behaviour in social networks, and also because of its mathematical representativeness for a large class of collective behaviours, opinion dynamics has recently attracted wide interest from the field of sociology, economics, physics, engineering and so forth. Multi-agent models of opinion formation provide us an effective way to investigate how consensus is formed among individuals who have quite different initial opinions [8–11], how (mis)information is spread over social networks [12], and how homophily principle (i.e. greater interaction between like-minded individuals) affects opinion formation and social learning [13–15], just to name a few.

In most of the existing models of opinion formation, agents are set to update their opinions by adopting the weighted averages of their neighbours' opinions (e.g. DeGroot model [8, 16, 17] and Hegselmann–Krause model [13, 14]), in which weights are all non-negative. For a detailed discussion of the assumption of averaging in social-influence models see the recent book by Friedkin and Johnsen [18]. Under such a mechanism, the difference of opinions in the group is inclined to decrease to zero if a consensus is reached or to a non-zero constant if persistent disagreement is established. As stated in [18], 'disagreeing individual positions and collective consensual positions rarely fall outside the range of a group's initial positions'. These models can represent the process of aggregation of opinions in the society to an extent while cannot capture one well-known social phenomenon – group polarisation which means the final opinions in the group are more divergent than individuals' initial opinions. In the real world, the tendency of group polarisation can usually be found on controversial issues such as abortion, homosexual marriage and leader election [19–21], where the discussion among individuals might result in increased divergence of opinions. Since putting non-negative weights on neighbours' opinions can reduce difference between agents, a natural way to cause polarisation is taking into account negative weights which captures dislike/distrust interactions. Indeed, this relax on weights can effectively result in polarisation, and most of the studies have quite solid theoretical foundations from structural balance theory on signed networks [22, 24–26].

However, involving negative weights (or adverse interactions) is not the only way to model group polarisation. We have more inspirations from sociology. For instance, in a recent 'PNAS' article Dandekar *et al.* propose a variant of DeGroot model by accounting for biased assimilation of others' opinions – individuals are more inclined to accept confirming opinions while reject disconfirming opinions [27]. Biased assimilation is a classical social psychology theory which can be traced back to the original work of Lord *et al.* [28]. Meanwhile, biased assimilation can also be regarded as a deduction from Heider's well-known work on balance theory [29] which points out the tendency to like similar persons. An important feature that distinguishes Dandekar's model from other opinion formation models with positive weights is that the agents use a non-linear updating rule to form their new opinions and might update their opinions to the outside of the convex hull formed by the opinions of their neighbours. This mechanism plays a crucial role in the sense that allows opinions to be more divergent than their initial values. Dandekar *et al.* test their model in homophilous networks, and show that their biased opinion formation process can result in polarisation when individuals in the network are sufficiently biased. Similar to [27], the updating rule in [30] also contains certain non-linearity such that the opinions after group discussion fall outside the range of the initial opinions. In [31], Mäs *et al.* propose a model of argument-communication theory of bi-polarisation where the opinion of each agent is decided by a set of arguments in her mind, and during discussion in the group, the agents exchange their arguments rather than opinions. The assumption that the opinion is a function of related arguments can be traced back to classical formalised theories from psychology [32]. By randomly receiving a new argument that supports her initial opinion, one agent might form a more extreme opinion, and therefore, in the group level, opinions might be more divergent after group discussion. Similar to [27], Mäs *et al.* also test their model in homophilous networks and show that polarisation can arise when homophily is sufficiently strong.

In this paper, along the line of the above studies, we propose an opinion formation model accounting for biased assimilation of arguments where, rather than receive and accept randomly chosen arguments like that in [31], the agents tend to accept those that confirm their own arguments. The mechanism of biased assimilation of arguments can be described by a modified Polya's urn process. Both theoretical analyses and simulations show that group polarisation happens under biased assimilation of arguments, and that networks with relatively isolated individuals lead to group polarisation easily.

The remainder of this paper is organised as follows: in Section 2, our opinion formation model is presented. In Section 3, theoretical analyses are provided. Simulations are given in Section 4. The paper ends by concluding remarks and open questions.

2 Model description

2.1 Network structure

Consider a social network as an undirected graph $G = (V, E)$, where $V = \{1, 2, \ldots, n\}$ is the node set and $E \subset V \times V$ is the edge set. Each node in V represents an agent, and an edge connecting i and j, denoted by the unordered pair $(i, j) \in E$, captures the fact that information can flow between agent i and agent j, and the two agents are called neighbours. The set of neighbours of agent i is denoted by $N_i = \{j \in V : (i, j) \in E\}$.

2.2 Arguments and opinion formation

Each agent $i \in V$ holds a numerically valued opinion $O_{i,t}$, which represents her stance on the issue under discussion at time t. Inspired by Mäs and Flache [31], here we also assume that the opinion of agent i is decided by her arguments related to the issue. This is a plausible and quite realistic setting in the sense that in the real world almost all of our opinions are established based on certain evidences or arguments that we have observed or been informed. Therefore introducing arguments into the process of opinion formation is a natural extension of the existing models. Suppose that the number of arguments of any agent is m which captures the argument diversity. The value of each argument is set to be either 1 (a pro argument) or -1 (a con argument), and we assume that all arguments are unlabelled which means that any two pro (or con) arguments are equivalent from any agent's point of view. Thus, each agent $i \in V$ only needs to maintain the number of pro arguments, $n^{\mathrm{p}}_{i,t}$, and the number of con arguments, $n^{\mathrm{c}}_{i,t}$. The assumption of unlabelled arguments also implies that hearing an argument a second time does not have a weaker effect. In [9], DeMarzo *et al.* use the definition of 'persuasion bias' to capture the situation that individuals update their opinions without accurately accounting for which information they receive is new and which is repetition. Based on these settings, the opinion of any agent i can be described as follows

$$O_{i,t} = \frac{1}{m}\left(n^{\mathrm{p}}_{i,t} - n^{\mathrm{c}}_{i,t}\right) \tag{1}$$

It is easy to see that the value of any opinion is located inside the interval of $[-1, 1]$. The opinion with 1 means that the agent extremely supports the issue while opinion with -1 means extremely opposes. The opinion updating process can be divided into two steps: the agents first update their arguments according to a certain rule which is stated in the following, and then, the opinions are formed based on new arguments according to (1).

2.3 Biased assimilation of arguments

As for the process of arguments updating, at each time step t one agent i is randomly selected from the whole group to participate in the updating. Then, one of her neighbours, say j, is chosen to discuss with her. During the interaction, agent i randomly receives an argument from the set of arguments of agent j. Rather than accepting the argument directly like that in [31], agent i further needs to randomly pick one argument from her own set of arguments. If the two arguments are equal (both pro or both con), then agent i accepts the argument, that is, adds a new argument of the same bias into her set of arguments, and randomly drops one argument from the set. This ends one step updating. The process can last as long as needed.

To obtain an intuitive impression of the process of biased assimilation of arguments, we can describe the above setting by a modified Polya's urn process. A basic Polya's urn model can be simply described as follows: one urn contains red balls and blue balls; one ball is drawn randomly from the urn and its colour observed; it is then put back in the urn; an additional ball of the same colour is added to the urn, and the selection process is repeated [33]. Here we modify the basic model by considering an urn for each agent and that whether adding a ball of a certain colour to one urn is also closely related to the ball drawn from one of other urns. More precisely, imagine that each agent has an urn containing m colourful balls, each being red or blue. If agent i and her neighbour j are chosen to participate in the updating at present time step, agent i first randomly selects a ball from j's urn. Without loss of generality, say one red ball is drawn, observed and put back. Then agent i chooses one ball from her own urn. Only if this ball is also red, she adds a new ball of red colour into her urn. Then, a ball in i's urn is randomly chosen and dropped to keep the number of balls in i's urn constant. If the ball drawn from i's urn is blue, she will keep the balls in her urn unchanged. From the above description, we can obtain an impression that the modified

doi: 10.1049/iet-cta.2014.0511

urn model has certain bias on adding balls. If an urn has more red balls, it has greater possibility to add a red ball, and the same statement can be applied to an urn with more blue balls. This captures the idea of biased assimilation of arguments in the process of argument communication, which is the main feature that distinguishes our model from other opinion formation models.

3 Theoretical analysis

In this section, we provide theoretical analysis of the model. Our theoretical analysis mainly focuses on the number of pro arguments, and leaves its opposite as a straightforward deduction.

Suppose at time t agent i with $n^{\mathrm{p}}_{i,t}$ is chosen to update her opinion, and meanwhile, one of her neighbours, say agent $j \in N_i$ with $n^{\mathrm{p}}_{j,t}$, is chosen to communicate with. Then, we can list all the situations with corresponding probabilities agent i might run into as follows:

(a) One pro argument is added, that is, $n^{\mathrm{p}}_{i,t+1} = n^{\mathrm{p}}_{i,t} + 1$, with probability

$$P(n^{\mathrm{p}}_{i,t+1} = n^{\mathrm{p}}_{i,t} + 1 | n^{\mathrm{p}}_{i,t}, n^{\mathrm{p}}_{j,t}) = \frac{n^{\mathrm{p}}_{j,t}}{m}\frac{n^{\mathrm{p}}_{i,t}}{m}\frac{m - n^{\mathrm{p}}_{i,t}}{m+1}$$

(b) One pro argument is removed, that is, $n^{\mathrm{p}}_{i,t+1} = n^{\mathrm{p}}_{i,t} - 1$, with probability

$$P(n^{\mathrm{p}}_{i,t+1} = n^{\mathrm{p}}_{i,t} - 1 | n^{\mathrm{p}}_{i,t}, n^{\mathrm{p}}_{j,t}) = \frac{m - n^{\mathrm{p}}_{j,t}}{m}\frac{m - n^{\mathrm{p}}_{i,t}}{m}\frac{n^{\mathrm{p}}_{i,t}}{m+1}$$

(c) The number of pro (and also con) arguments is unchanged, that is, $n^{\mathrm{p}}_{i,t+1} = n^{\mathrm{p}}_{i,t}$, with probability

$$P(n^{\mathrm{p}}_{i,t+1} = n^{\mathrm{p}}_{i,t} | n^{\mathrm{p}}_{i,t}, n^{\mathrm{p}}_{j,t}) = 1 - \frac{m n^{\mathrm{p}}_{i,t} - (n^{\mathrm{p}}_{i,t})^2}{m(m+1)}$$

Note that if at time T agent i has all arguments being pro (or equivalently, being con), her arguments along with her opinion will stay unchanged for any $t \geq T$. These situations can be covered with special values of $n^{\mathrm{p}}_{i,t}$.

3.1 Expected value and variance of the number of pro arguments

Based on the above analysis, we have the expected number of pro arguments after updating given its present value

$$\begin{aligned} E(n^{\mathrm{p}}_{i,t+1} | n^{\mathrm{p}}_{i,t}, n^{\mathrm{p}}_{j,t}) &= \frac{n^{\mathrm{p}}_{j,t}}{m}\frac{n^{\mathrm{p}}_{i,t}}{m}\frac{m - n^{\mathrm{p}}_{i,t}}{m+1}(n^{\mathrm{p}}_{i,t} + 1) \\ &\quad + \frac{m - n^{\mathrm{p}}_{j,t}}{m}\frac{m - n^{\mathrm{p}}_{i,t}}{m}\frac{n^{\mathrm{p}}_{i,t}}{m+1}(n^{\mathrm{p}}_{i,t} - 1) \\ &\quad + \left[1 - \frac{m n^{\mathrm{p}}_{i,t} - (n^{\mathrm{p}}_{i,t})^2}{m(m+1)}\right] n^{\mathrm{p}}_{i,t} \\ &= n^{\mathrm{p}}_{i,t} + \frac{2n^{\mathrm{p}}_{j,t} - m}{m}\frac{n^{\mathrm{p}}_{i,t}(m - n^{\mathrm{p}}_{i,t})}{m(m+1)} \end{aligned} \tag{2}$$

From (2), we know that if agent i's neighbour, agent j, has a neutral opinion, that is, $n^{\mathrm{p}}_{j,t} = n^{\mathrm{c}}_{j,t} = m/2$, then the number of pro arguments of agent i is independent of her neighbour's property. Equation (2) specialises to the following form

$$E(n^{\mathrm{p}}_{i,t+1} | n^{\mathrm{p}}_{i,t}, n^{\mathrm{p}}_{j,t}) = n^{\mathrm{p}}_{i,t} \tag{3}$$

which implies that the expected number of pro arguments of one agent is its present value if the agent chooses a neutral neighbour to communicate with.

Except the special case of $n^{\mathrm{p}}_{j,t} = m/2$, the set of arguments of agent i is influenced by her neighbours' in the interaction. The level of influence is captured by the weight $[\{n^{\mathrm{p}}_{i,t}(m - n^{\mathrm{p}}_{i,t})\}/\{m(m+1)\}]$. Clearly, the influence is greater if agent i's opinion is moderate, that is, $n^{\mathrm{p}}_{i,t}$ close to $m/2$, and correspondingly, the influence is smaller if agent i has extreme opinions, that is, $n^{\mathrm{p}}_{i,t}$ close to m or zero. This can be interpreted as: the agents with extreme opinions are more difficult to be affected by their neighbours (or other information sources) than those who have moderate opinions.

For convenience, let

$$X = \frac{n^{\mathrm{p}}_{i,t}}{m}\frac{m - n^{\mathrm{p}}_{i,t}}{m+1} \quad \text{and} \quad Y = \frac{2n^{\mathrm{p}}_{j,t} - m}{m}\frac{n^{\mathrm{p}}_{i,t}(m - n^{\mathrm{p}}_{i,t})}{m(m+1)}$$

Then, the variance of the number of pro arguments of agent i at time $t+1$ can be computed as follows

$$\begin{aligned} \mathrm{Var}(n^{\mathrm{p}}_{i,t+1} | n^{\mathrm{p}}_{i,t}, n^{\mathrm{p}}_{j,t}) &= \frac{n^{\mathrm{p}}_{j,t}}{m} X (Y-1)^2 \\ &\quad + \frac{m - n^{\mathrm{p}}_{j,t}}{m} X (Y+1)^2 + (1 - X)Y^2 \\ &= Y^2 + 2\left(1 - \frac{2n^{\mathrm{p}}_{j,t}}{m}\right) XY + X \end{aligned} \tag{4}$$

For the special case with $n^{\mathrm{p}}_{j,t} = n^{\mathrm{c}}_{j,t} = m/2$, we have

$$\mathrm{Var}(n^{\mathrm{p}}_{i,t+1} | n^{\mathrm{p}}_{i,t}, n^{\mathrm{p}}_{j,t}) = \frac{n^{\mathrm{p}}_{i,t}}{m+1}\left(1 - \frac{n^{\mathrm{p}}_{i,t}}{m}\right) \tag{5}$$

which is independent of neighbours' properties.

3.2 Convergence to extreme opinions

We next prove that, with biased assimilation of arguments, the opinion of each agent will eventually converge to one of the two extremes, that is, 1 or −1, with probability one.

Theorem 1: Suppose that each agent in the network has at least one neighbour to communicate with, and all agents update their opinions by biasedly assimilating the arguments of neighbours as described in Section 2.3. Then, the opinions of all agents converge to extremes with probability one.

Proof: Without loss of generality, we focus our analysis on any agent $i \in V$. First of all, we deal with two trivial cases. The first one is that the agent has all arguments with a common bias, that is, $n^{\mathrm{p}}_{i,t} = 0$ or $n^{\mathrm{p}}_{i,t} = m$ at a certain time step t. From (1), we know that the agent has an extreme opinion and will keep it for any $\tau \geq t$. The second trivial case is that, although the agent has both of pro and con arguments, all of her neighbours have all arguments with a common bias. An elementary computation shows that the agent will reach an extreme opinion same to her neighbours with probability one.

Now we consider more general cases. Since updates in the model are asynchronous, we only need to consider the time steps when agent i is selected to update, and rearrange the corresponding time steps as $\{t_{i,1}, t_{i,2}, \ldots\}$. In fact, under the new time scale the updating of $n^{\rm p}_{i,t}$ can be viewed as a random walk with two absorbing boundaries, as illustrated in the following:

Except for the two trivial cases mentioned above, both of the probabilities of adding one pro argument and removing one pro argument are strictly positive. We bound the probability as follows

$$\begin{aligned} P(n^{\rm p}_{i,t_{i,k+1}} = n^{\rm p}_{i,t_{i,k}} + 1 | n^{\rm p}_{i,t_{i,k}}) &\geq \frac{1}{|N_i|}\frac{1}{m}\frac{n^{\rm p}_{i,t_{i,k}}}{m}\frac{m - n^{\rm p}_{i,t_{i,k}}}{m+1} \\ &\geq \frac{1}{N-1}\left(\frac{1}{m}\right)^2\frac{m-1}{m+1} \triangleq p_\sigma \end{aligned}$$

By the same deduction, we have

$$P(n^{\rm p}_{i,t_{i,k+1}} = n^{\rm p}_{i,t_{i,k}} - 1 | n^{\rm p}_{i,t_{i,k}}) \geq p_\sigma$$

From Fig. 1, we know that, for any position of $n^{\rm p}_{i,t_{i,k}}$, it needs at most $m/2$ steps to reach an absorbing boundary, and therefore the probability of reaching an absorbing boundary is greater or equal to $p_\sigma^{m/2} > 0$. Accordingly, the probability that the process will not reach any boundary in $m/2$ steps is less than $1 - p_\sigma^{m/2} < 1$, in m steps less than $(1 - p_\sigma^{m/2})^2$ etc.

Since the probability of not reaching any boundary in n steps is monotonically decreasing, as $t \to \infty$, we have that

$$\left(1 - p_\sigma^{m/2}\right)^t \to 0 \tag{6}$$

which means that the set of arguments of agent i will eventually contain only pro or con arguments.

The above deduction can be applied to any agent $i \in V$. By (1), we have that the opinions of all agents converge to extremes with probability one. □

In the model of Mäs *et al.* [31], all the arguments are labelled such that any argument is distinguishable from any other ones. Under this assumption, if agent i receives an argument that already exists in her set of arguments, she simply rearranges the order of these arguments according to their 'recency'. The new opinion is based on the same set of arguments as that in the previous time step. Consensus on moderate opinions can be achieved if all agents base their opinions on the same set of arguments such that their opinions are identical and argument communication will not change their sets of arguments.

Once we only count the number of pro and con arguments, the agent will always update her arguments until all the arguments have identical bias. Mathematically, the opinion dynamics of any agent will turn into an one-dimensional (1D) random walk with absorbing boundaries. The number of pro arguments might be 1 to $m-1$ in transient state, but will eventually reach 0 or m with probability one. Therefore reaching an extreme is an inherent property of the proposed model. We regard this property as a result of the Polya-like self-reinforcement process. In order to model consensus on moderate opinions, some types of feedback mechanism should be introduced to counter the self reinforcement tendency. We leave this for our future work. In [34] an opinion formation model involving persuasive arguments exchange is proposed. Although the dynamics is quite different from ours, they obtain a similar result that the opinion of any agent will reach an extreme value after group discussion. However, they further find that consensus in either positive or negative extreme opinion is eventually achieved. This is quite different from our model where both consensus on one extreme opinion and bi-polarisation are possible.

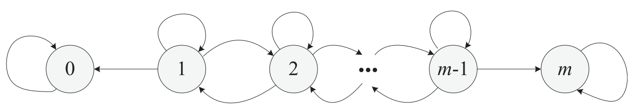


Fig. 1 *Illustration of the updating of pro arguments by a random walk with absorbing boundaries*

3.3 Convergence speed of the heterogeneity of the arguments: a simplified case study

In the above section, we have proven that the opinion of each agent will reach one of the extreme values. However, it is quite difficult to figure out the convergence speed, since it is closely related to the topology of the network and the choice of interaction pairs at each time step. In the following, we investigate the convergence speed from the perspective of heterogeneity of any agent i's arguments in a simplified case. Let the definition of heterogeneity for agent i be

$$H_{i,t} = \frac{n^{\rm p}_{i,t}}{m+1}\left(1 - \frac{n^{\rm p}_{i,t}}{m}\right)$$

The heterogeneity $H_{i,t}$ for any $t \geq 0$ has its minimal value at zero when $n^{\rm p}_{i,t} = 0$ or $n^{\rm p}_{i,t} = m$, and maximal value at $[m/\{4(m+1)\}]$ when $n^{\rm p}_{i,t} = n^{\rm c}_{i,t} = m/2$. This definition of heterogeneity conforms to our intuition.

The following theorem shows that, in a totally neutral environment, the agent who updates her opinion by biasedly assimilating arguments will achieve an extreme opinion exponentially fast.

Theorem 2: Suppose that agent i has all neighbours with $n^{\rm p}_{j,t} = m/2$ for any t and $j \in N_i$. Then, the expected heterogeneity of agent i's arguments decays at rate of $[1/\{m(m+1)\}]$.

Proof: For any $i \in V$ and any time $t \geq 0$, we have the following relationship between the expected value of heterogeneity at the next step and its present value.

$$\begin{aligned} &E\left(H_{i,t+1}\right) \\ &= \frac{1}{m+1}\left[E\left(n^{\rm p}_{i,t+1}\right) - \frac{1}{m}E(n^{\rm p}_{i,t+1})^2\right] \\ &= \frac{1}{m+1}\left[n^{\rm p}_{i,t} + \frac{2n^{\rm p}_{j,t} - m}{m}H_{i,t} - \frac{1}{m}\left(H_{i,t} + E^2(n^{\rm p}_{i,t+1})\right)\right] \\ &= \frac{n^{\rm p}_{i,t}}{m+1} + \frac{2n^{\rm p}_{j,t} - m}{m(m+1)}H_{i,t} - \frac{H_{i,t}}{m(m+1)} - \frac{(n^{\rm p}_{i,t})^2}{m(m+1)} \\ &\quad - 2n^{\rm p}_{i,t}\frac{2n^{\rm p}_{j,t} - m}{m^2(m+1)}H_{i,t} - \frac{H^2_{i,t}}{m(m+1)}\left(\frac{2n^{\rm p}_{j,t} - m}{m}\right)^2 \\ &= H_{i,t}\left[1 - \frac{1}{m(m+1)} - \frac{2n^{\rm p}_{j,t} - m}{m(m+1)}\right. \\ &\quad \left.\times\left(\frac{2n^{\rm p}_{i,t} - m}{m} + \frac{2n^{\rm p}_{j,t} - m}{m^2}H_{i,t}\right)\right] \end{aligned} \tag{7}$$

IET Control Theory Appl., 2015, Vol. 9, Iss. 3, pp. 485–492
doi: 10.1049/iet-cta.2014.0511

Here we have $n^{\mathrm{p}}_{j,t} = m/2$ which implies that agent i is facing a neighbour with a neutral opinion. Under this condition, (7) specialises to the following form

$$E\left(H_{i,t+1}\right) = H_{i,t}\left[1 - \frac{1}{m(m+1)}\right]$$

By induction, we have that, for a large enough m,

$$E\left(H_{i,t}\right) = H_{i,0}\left[1 - \frac{1}{m(m+1)}\right]^{t} \simeq H_{i,0}e^{-t/[m(m+1)]} \tag{8}$$

which implies that the expected heterogeneity of agent i's arguments decays at rate of $[1/\{m(m+1)\}]$. □

The result that biased assimilation of arguments causes individuals to form more extreme opinions even being exposed to neutral environments conforms to experimental evidences found in sociology (e.g. [35–38]). Furthermore, (6) and (8) show that a smaller number of arguments might result in a faster convergence to extreme opinions, which also conforms to previous works, such as [31, 39]. This result is further verified by simulations in the next section, in which we focus on the qualitative influence of the network structure and some key parameters on the performance of the model.

4 Simulations

To test our model on a well-controlled simulation environment, we choose a class of computer-generated networks – Newman–Watts small-world networks (small-world networks for short) [40]. A 1D small-world network model is defined as a chain with n nodes and periodic boundary condition. Initially, each node is connected to k closest neighbours. Therefore each node has an initial degree of k. Then, for any pair of nodes in the network, a shortcut is added between them with probability p. Illustration of small-world networks is depicted in Fig. 2.

4.1 Verification of group polarisation

We first test our model on a specific small-world network to obtain an intuition of the group polarisation established by the model. Simulation settings are as follows. Let the number of nodes in the network $n = 100$, the initial degree of each node $k = 4$, the probability of adding random edges $p = 0.02$ and the number of arguments $m = 50$. Initially, an element in any agent's set of arguments has the same probability to be 1 (pro argument) or -1 (con argument). At each time step, biased assimilation of arguments is performed following the statement in Section 2.3, and the opinion of each agent is calculated by (1). The opinion evolution of the whole group is illustrated in Fig. 3, from which we can see that the initial opinions are distributed close to zero because of the equal probability of each argument to be 1 or -1. Then, after about 2.25×10^5 simulation steps, the whole group is divided into two opposing extremes, that is, group polarisation is achieved. This also implies that at the end of simulation, the set of arguments of any agent is either full of 1 or full of -1.

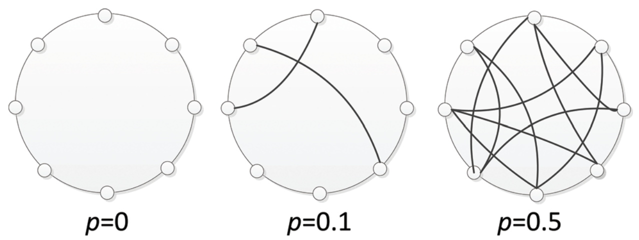


Fig. 2 *Illustration of small-world networks with different probability of adding random edges*

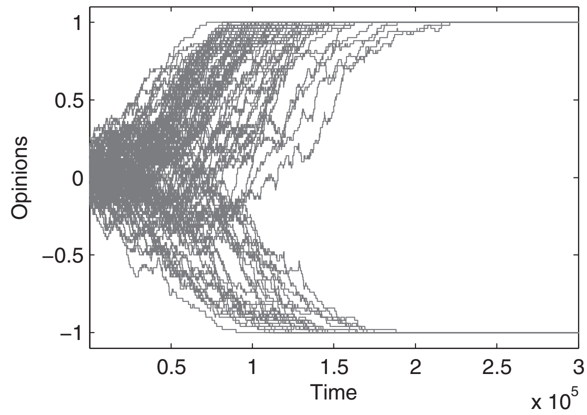


Fig. 3 *Group polarisation in a small-world network ($N = 100$, $k = 4, p = 0.02$ and $m = 50$)*

4.2 Influence of network structure

To investigate how the network structure affects the performance of opinion dynamics in our model, we let $n = 100$, $k = 4$, $m = 50$ and p vary from 0 to 1 to establish different network conditions. A small p corresponds to a network in which the agents are relatively isolated from each other, and accordingly, a large p corresponds to wide communication.

From Theorem 1, we know that the opinion of each agent must converge to one of the extremes. All of the opinions might reach a consensus on one extreme, or split into two groups eventually converging to opposite extremes. As the definition of group polarisation in this paper, that is, opinions through group discussion are more divergent than initial opinions, we only regard the latter as a successful group polarisation.

We define a variable Z to capture the population difference of the two opposite extremes. Technically

$$Z = 1 - \frac{|n_{+1} - n_{-1}|}{n} \tag{9}$$

where n_{+1} is the number of agents with opinions of 1 at the end of opinion evolution and accordingly, n_{-1} is the number with -1. By this definition, we know that the variable Z can capture the level of group polarisation from a consensus on one extreme, that is, $Z = 0$, to a perfect split with equal number of opinions on two opposite extremes, that is, $Z = 1$.

First of all, we study how the variable Z changes as p varies from 0 to 1. The result is shown in Fig. 4*a*. Note that the horizontal axis is logarithmic. This is because if the curve is plotted on a linear scale, the change of Z as p varies from 10^{-3} to 10^{-2} will be compressed which makes the tendency very difficult to judge. We can see that the curve monotonically decreases as p increases, especially at the interval of $10^{-2} \le p \le 10^{-1}$. This means that group polarisation happens easily in the case that the agents are

doi: 10.1049/iet-cta.2014.0511

relatively isolated in the network. Meanwhile, we study how the time to polarisation varies as p changes, and the result is shown in Fig. 4*b*. It is well known that the small-world effect can effectively promote the spread of information, the transmission of disease, reaching a consensus among individuals and so forth [41–43]. Similarly, we find that the small-world effect can also speed up opinion convergence, either to a consensus on one extreme value or to group polarisation.

The phenomenon shown in Fig. 4 might be explained as follows: when agents are relatively isolated in the network, they mainly interact with their local neighbours which leads to the emergence of some local strong supporters whose arguments are all pro or all con. Strong supporters can spread their arguments to the reset of the population while do not change their own arguments. Since the network contains very few long-range random links, this spread is very slow. Conversely, when there exist a great number of long-range links which connect otherwise separated subgroups, it might be not so easy to reach polarisation because of diversity influence from neighbours, but once strong supporters appear, the spread of their arguments will be very fast.

Here we can compare our findings to the results of [44], in which the authors also find that links between subgroups, which play a role similar to long-range random links, can prevent subgroup polarisation. As for convergence time, however, they show that more links between subgroups leads to overall consensus in longer time, which is contrary to our finding. We attribute this to the different methods of argument communication in the two models. In our model, agents biasedly assimilate arguments in the sense that if an agent with more pro (con) argument receives a con (pro) argument from her neighbour, she accepts it with a relatively low probability. However, in [44] an agent accepts any argument she receives from her neighbour and let it have the highest 'recency'. This essential difference in argument communication leads to quite different even opposite opinion dynamics.

4.3 Influence of the number of arguments

In our model, the number of arguments is an adjustable parameter. In the following, we investigate its influence on the process of group polarisation. Let $n = 100$, $k = 4$, $p = 0.02$ and the number of arguments m increase from 5 to 100 at five intervals. Under each condition, the simulation repeats for 100 times. From Fig. 5*a*, we can see that as the number of arguments increases, the level of group polarisation decreases. This tendency can be interpreted as that group polarisation is hard to achieve if the issue discussed has a great number of related arguments. We can also study how the time to polarisation changes under different number of arguments, and the result is shown in Fig. 5*b*. We can see that as the number of arguments increases, the time to polarisation also increases. This tendency conforms to the result shown by (6) and (8), and also the results obtained in previous works, such as [31, 39].

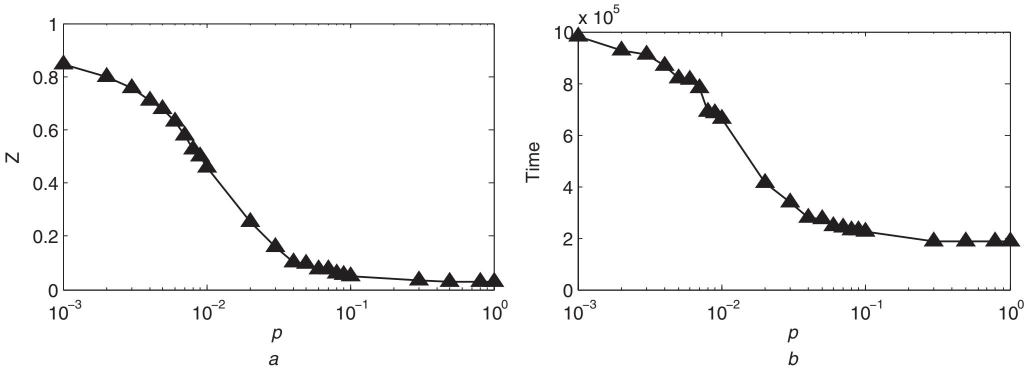


Fig. 4 *Influence of the probability of adding random edges on the opinion dynamics (n = 100, k = 4 and m = 50)*

a Level of group polarisation as a function of p
b Time to convergence as a function of p

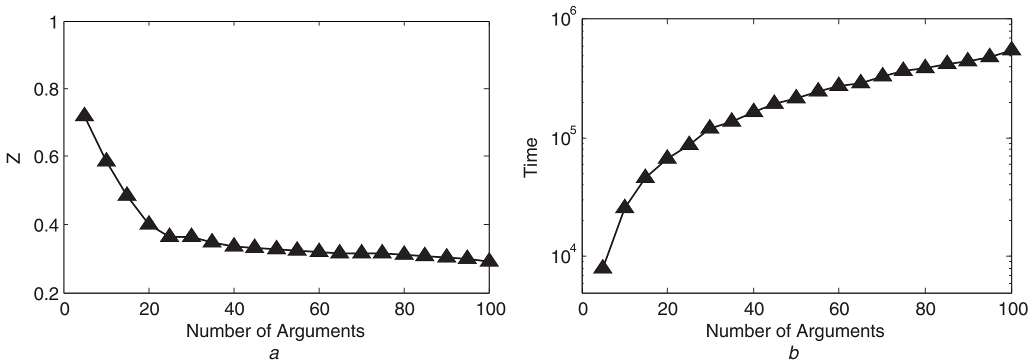


Fig. 5 *Influence of the number of arguments on the opinion dynamics (n = 100, k = 4 and p = 0.02)*

a Level of group polarisation as a function of the number of arguments
b Time to convergence as a function of the number of arguments

IET Control Theory Appl., 2015, Vol. 9, Iss. 3, pp. 485–492
doi: 10.1049/iet-cta.2014.0511

Intuitively, an agent has a relatively high probability to form either pro polarisation or con polarisation if m is small. As the number of arguments increases, it becomes harder for the agents to collect only pro or con arguments, and thus, the time to polarisation becomes longer. Once a strong supporter of one extreme opinion appears, she has the so-called first-mover advantage, and she is more likely to spread her opinion/arguments to the rest of the network, which results in an uneven population distribution of the supports of the two extreme opinions.

5 Conclusion

In this paper, we have proposed an opinion formation model with group polarisation. By introducing biased assimilation of arguments in the process of argument communication, we show that agents in the group tend to form two clusters with opposing extreme opinions, that is, the process of opinion formation is polarising, especially when agents are relatively isolated in the network. In future work, we will perform more in-depth theoretical analysis and simulations in different types of networks besides small-world networks to investigate how the network structure affects group polarisation in the model. Furthermore, the leader selection problem is also important which has attracted wide interest from both sociology and engineering fields. In particular, from the perspective of control theory, if we can control or influence a fixed number of agents, how to identify these target agents such that the overall opinions in the system reach a desired state under the proposed opinion updating rule. This is also an open question for our future work.

6 Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant nos. 61374176 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).

7 References

1 Reynolds, C.W.: 'Flocks, herds, and schools: a distributed behavioral model', *Comput. Graph.*, 1987, **21**, (4), pp. 25–34

2 Couzin, I.D., Krause, J., Franks, N.R., Levin, S.A.: 'Effective leadership and decision-making in animal groups on the move', *Nature*, 2005, **433**, pp. 513–516

3 Ballerini, M., Cabibbo, N., Candelier, R., Cavagna, A.: 'Interaction ruling animal collective behavior depends on topological rather than metric distance: evidence from a field study', *Proc. Natl. Acad. Sci. USA*, 2008, **105**, (4), pp. 1232–1237

4 Xue, D., Yao, J., Chen, G., Yu, Y.L.: 'Formation control of networked multi-agent systems', *IET Control Theory Appl.*, 2010, **4**, (10), pp. 2168–2176

5 Wang, Y., Yan, W., Li, J.: 'Passivity-based formation control of autonomous underwater vehicles', *IET Control Theory Appl.*, 2012, **6**, (4), pp. 518–525

6 Chen, Y.Y., Tian, Y.P.: 'Coordinated adaptive control for three-dimensional formation tracking with a time-varying orbital velocity', *IET Control Theory Appl.*, 2013, **7**, (5), pp. 646–662

7 Zhang, H., Yan, H., Yang, F., Chen, Q.: 'Distributed average filtering for sensor networks with sensor saturation', *IET Control Theory Appl.*, 2013, **7**, (6), pp. 887–893

8 DeGroot, M.H.: 'Reaching a consensus', *J. Am. Stat. Assoc.*, 1974, **69**, (345), pp. 118–121

9 DeMarzo, P.M., Vayanos, D., Zwiebel, J.: 'Persuasion bias, social influence, and unidimensional opinons', *Q. J. Econ.*, 2003, **118**, (3), pp. 909–968

10 Golub, B., Jackson, M.O.: 'Naïve learning in social networks: convergence, influence, and the wisdom of crowds', *Am. Econ. J. Microecon.*, 2009, **2**, (1), pp. 112–149

11 Molavi, P., Jadbabaie, A., Rad, K.R., Tahbaz-Salehi, A.: 'Reaching consensus with increasing information', *IEEE J. Sel. Top. Signal Process.*, 2013, **7**, (2), pp. 358–369

12 Acemoglu, D., Ozdaglar, A., ParandehGheibi, A.: 'Spread of (mis)information in social networks', *Games Econ. Behav.*, 2010, **70**, (2), pp. 194–227

13 Hegselmann, R., Krause, U.: 'Opinion dynamics and bounded confidence models, analysis, and simulation', *J. Artif. Societies Soc. Simul.*, 2002, **5**, (3)

14 Lorenz, J.: 'Continuous opinion dynamics under bounded confidence: a survey', *Int. J. Mod. Phys. C*, 2007, **18**, (12), pp. 1819–1838

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

16 Berger, R.L.: 'A necessary and sufficient condition for reaching a consensus using DeGroot's method', *J. Am. Stat. Assoc.*, 1981, **76**, (374), pp. 415–418

17 Gilardoni, G.L., Clayton, M.K.: 'On reaching a consensus using DeGroot's iterative pooling', *Ann. Stat.*, 1993, **21**, (1), pp. 391–401

18 Friedkin, N.E., Johnsen, E.C.: 'Social influence network theory' (Cambridge University Press, New York, 2011)

19 Abramowitz, A.I., Saunders, K.L.: 'Is polarization a myth', *J. Polit.*, 2008, **70**, (2), pp. 542–555

20 Fiorina, M.P., Abrams, S.J.: 'Political polarization in the American public', *Annu. Rev. Polit. Sci.*, 2008, **11**, pp. 563–588

21 Levendusky, M.S.: 'The microfoundations of mass polarization', *Polit. Anal.*, 2009, **17**, (2), pp. 162–176

22 Flache, A., Macy, M.W.: 'Small world and cultural polarization', *J. Math. Sociol.*, 2011, **35**, pp. 146–176

23 Altafini, C.: 'Dynamics of opinion forming in structural balanced social networks', *PLoS One*, 2012, **7**, (6), p. e38135

24 Centola, D., González-Avella, J.C., Eguíluz, V.M., Miguel, M.S.: 'Homophily, cultural drift, and the co-evolution of cultural groups', *J. Confl. Resolut.*, 2007, **51**, (6), pp. 905–929

25 Altafini, C.: 'Consensus problems on networks with antagonistic interactions', *IEEE Trans. Autom. Control*, 2013, **58**, (4), pp. 935–946

26 Hu, J., Zheng, W.X.: 'Bipartite consensus for multi-agent systems on directed signed networks', Proc. IEEE. Conf. Decision and Control, Florence, Italy, December 2013, pp. 3451–3456

27 Dandekar, P., Goel, A., Lee, D.T.: 'Biased assimilation, homophily, and the dynamics of polarization', *Proc. Natl. Acad. Sci. USA*, 2013, **110**, (15), pp. 5791–5796

28 Lord, C.G., Ross, L., Lepper, M.R.: 'Biased assimilation and attitude polarization: the effects of prior theories on subsequently considered evidence', *J. Pers. Soc. Psychol.*, 1979, **37**, (11), pp. 2098–2109

29 Heider, F.: 'Attitudes and cognitive organization', in Fishbein, M. (Ed.): 'Readings in Attitude Theory and Measurement' (Wiley, New York, 1967), pp. 39–41

30 Shin, J.K., Lorenz, J.: 'Tipping diffusivity in information accumulation systems: more links, less consensus', *J. Stat. Mech.*, 2010, P06005

31 Mäs, M., Flache, A.: 'Differentiation without distancing: explaining bi-polarization of opinions without negative influence', *PLoS One*, 2013, **8**, (11), p. e74516

32 Fishbein, M., Ajzen, I.: 'Belief, attitude, intention and behavior: an introduction to theory and research (Addison-Wesley, Reading, MA, 1975)

33 Mahmoud, H.M.: 'Polya urn models' (Taylor & Francis Group, Boca Raton, 2009)

34 La Rocca, C.E., Braunstein, L.A., Vazquez, F.: 'The influence of persuasion in opinion formation and polarization', *EPL*, 2014, **106**, p. 40004

35 Miller, A.G., McHoskey, J.W., Bane, C.M., Dowd, T.G.: 'The attitude polarization phenomenon: role of response measure, attitude extremity, and behavioral consequences of reported attitude change', *J. Pers. Soc. Psychol.*, 1993, **64**, (4), pp. 561–574

36 Munro, G.D., Ditto, P.H., Lockhart, L.K., Fagerlin, A., Gready, M., Peterson, E.: 'Biased assimilation of sociopolitical arguments: evaluating the 1996 U.S. presidential debate', *Basic Appl. Soc. Psychol.*, 2002, **24**, (1), pp. 15–26

37 Taber, C.S., Lodge, M.: 'Motivated skepticism in the evaluation of political beliefs', *Am. J. Pol. Sci.*, 2006, **50**, (3), pp. 755–769

38 Metz, T., Lorenz, J.: 'Become who you are: the homing pattern in partisanship as a self-reinforcing stochastic process'. The 3rd Annual General Conf. of the European Political Science Association. Available at SSRN: http://www.ssrn.com/abstract=2224740, accessed June 2013

39 Axelrod, R.: 'The dissemination of culture: a model with local convergence and global polarization', *J. Confl. Resolut.*, 1997, **41**, (2), pp. 203–226

40 Newman, M.E.J., Watts, D.J.: 'Renormalization group analysis of the small-world network model', *Phys. Lett. A*, 1999, **263**, (11), pp. 341–346

41 Moore, C., Newman, M.E.J.: 'Epidemics and percolation in small-world networks', *Phys. Rev. E*, 2000, **61**, (5), pp. 5678–5682

42 Lago-Fernández, L.F., Huerta, R., Corbacho, F., Sigüenza, J.A.: 'Fast response and temporal coherent oscillations in small-world networks', *Phys. Rev. Lett.*, 2000, **84**, (12), pp. 2758–2761

43 Olfati-Saber, R.: 'Ultrafast consensus in small-world networks'. Proc. American Control Conf., Portlan, USA, June 2005, pp. 2371–2378

44 Mäs, M., Flache, A., Takács, K., Jehn, K.: 'In the short term we divide, in the long term we unite: demographic crisscrossing and the effects of fault lines on subgroup polarization', *Org. Sci.*, 2013, **24**, (3), 716–736

doi: 10.1049/iet-cta.2014.0511