Modeling the Emergence of Echo Chambers through the Interaction between Beliefs and Opinions

Hiro Kataoka
(Master's Program in Computer Science)

Advised by Koji Hasebe and Jérôme Euzenat

Submitted to the Graduate School
of Science and Technology
Degree Programs in Systems and
Information Engineering
in Partial Fulfillment of the Requirements
for the Degree of Master of Engineering
at the
University of Tsukuba

March 2026

Abstract

The occurrence of echo chambers is a well-known phenomenon in social networks. Echo chambers are groups of agents satisfying three properties: homogeneity, i.e., sharing the same opinion, segregation, i.e., listening primarily to agents of the same group, and reinforcement, i.e., increasing opinion similarity over time. Opinion dynamics models have been used to explain how echo chambers occur through agents revising their opinions. However, social network users also exchange beliefs supporting their opinions. Here, the aim of this study is to identify the effect of propagating beliefs together with opinions on the formation of echo chambers. For that purpose, an existing opinion dynamics model is complemented with classical belief merge operations. This allows agents to propagate and update their opinions and beliefs. Beliefs and opinions may also be maintained coherent guided by values that agents share. In addition, new measures, which apply to either beliefs and opinions, to identify echo chambers based on the three properties are introduced. This model is used to simulate opinion and belief propagation with varying initial conditions and agent tolerance for different opinions/beliefs. Results confirm the occurrence of echo chambers observed in previous opinion dynamics studies in a different setting and with a more strict measure. They also demonstrate the occurrence of echo chambers with belief propagation. Finally, they show that connecting opinions to beliefs and maintaining their coherence decreases the number of echo chambers.

Note

Following the double degree program between University of Tsukuba and Université Grenoble Alpes (UGA, France) in the field of computer science, an earlier version of this thesis has been submitted to UGA in June 2025 with the title "Exchanging and updating opinions and beliefs reduces echo chambers" as my master report. This version subsumes the earlier version.

This document is compiled on: March 12, 2026

Contents

List of Figures

Chapter 1

Introduction

1.1 Background

Echo chambers are well-known phenomena in social networks. They denote a situation in which a group of persons only interacts with people who have opinions close to theirs, thus reinforcing their opinions [45]. Overall, literature supports that they exist [34].

Opinion dynamics studies how agents' opinions are formed through their interactions [28]. Bounded confidence models are one of the most well-known and classical models in this field. They have been applied to explain how echo chambers arise because they allow agents to ignore other opinions which are too different from theirs. For example, it is shown that, similarly to existing social networks, agents following bounded confidence models and modifying the network to stop listening to others with different opinions form echo chambers [55].

However, these models do not take into account the effect of agents' beliefs. People exchange and revise not only their opinions but also their beliefs through interactions. For example, beliefs can be communicated to justify opinions: "*I am for constructing tram lines* (opinion) because *I believe that this would reduce traffic jams* (justification by beliefs)." Moreover, people can change their opinions with respect to their beliefs and vice versa. Such interactions have not been considered.

The way in which beliefs can propagate has been studied in the field of belief change. In addition to the social propagations, opinions and beliefs should be connected within a single agent. This can be described based on values agents share.

1.2 Approach

Based on these observations, a model defining interactions between opinions and beliefs as well as interactions between opinions and between beliefs is introduced. Hereafter, opinions and beliefs are different: the opinions of an agent denote how positive or negative it is toward the corresponding topic, while beliefs represent what it considers as true based on logical reasoning. In this model, opinions and beliefs are propagated through existing models [55, 57] and a value-based mechanism is used to maintain their coherence. The novelty of this model is that it connects opinions and symbolic beliefs by generalizing existing ones.

The aim of this study is to investigate whether allowing agents to interact based on beliefs leads to more or less echo chambers. To achieve this, simulations with the *SOBA* simulator [11] I implemented are performed to test the following hypotheses:

H0 If opinions evolve independently of beliefs, opinion echo chambers should arise.

H1 If beliefs evolve independently of opinions, belief echo chambers should arise.

H2 Considering opinions and beliefs when agents filter their neighbors creates more echo chambers than considering opinions or beliefs.

H3 Connecting opinions and beliefs and keeping them coherent increases the number of echo chambers.

Various measures have been provided to identify echo chambers in online social networks and simulations [45, 38, 34]. However, such measures do not necessarily cover all the characteristics of echo chambers. Thus, new measures to capture echo chambers based on their definition in existing work are introduced. They compute the number of strongly connected components that have the characteristics of echo chambers. They allow to identify echo chambers directly.

1.3 Results

Multi-agent simulations of the proposed model confirm that agents form echo chambers either based on their opinions, as in [55], or their beliefs. In other words, the results show that it is a generalization of existing models. In addition, agents form less echo chambers if they evolve their opinions and beliefs together from the viewpoint of opinions, but they form more echo chambers from the viewpoint of beliefs. Finally, they form less echo chambers if they synchronize their opinions and beliefs than if they do not. These results are obtained by comparing the means of the measures that are defined in this thesis.

1.4 Outline

The remainder of this thesis is organized as follows. Chapter 2 shows how agents behave in an intuitive scenario. This illustrates that following procedures are required: (1) a procedure to update opinions based on other agents' opinions, (2) to update beliefs based on other agents' beliefs, and (3) to synchronize them. Thus, in Chapter 3, existing models relevant to this problem and especially to process (1) and (2) are presented. Based on the existing models, a new model is defined in Chapter 4. In this model, the processes (1) and (2) are based on existing models, and the process (3) is provided. Chapter 5 introduces hypotheses to be tested with the model in this thesis, a measure to identify echo chambers, and experimental settings. Chapter 6 presents and discusses the results from the experiments planned in the previous chapter. Finally, Chapter 7 concludes this thesis and presents future work.

Chapter 2

Intuitive Scenario

Before reviewing related models and presenting the proposed model, how agents evolve their opinions and beliefs in the model is presented through an intuitive scenario.

Alice, Bob, Celia, and David are discussing whether the city they are living in should build tram lines or not. The network they form can be visualized in Figure 2.1 (left). Alice is in favor of the construction plan because she believes that this would reduce traffic jams. On the contrary, Bob is against the plan because he believes that this will lead to raise taxes. Celia may be negative toward the plan because she may think that if taxes should be kept low, then trams cannot be built. Finally, David may be neutral toward it because he may believe that either 'having tram lines and less traffic jams' or 'doing without trams and less taxes' can happen, but has no idea about which one actually happens.

Their discussion starts from listening to each other's opinions and beliefs. Whether a person takes into account what another explains depends on how close their opinions and beliefs are. Here, assume that Bob and Alice have close enough opinions and beliefs and thus Alice listens to Bob, while Alice ignores Celia because theirs are too different. Then, Alice listens to what Bob explains and updates her opinion (i.e., how positive or negative she considers the topic) and beliefs (i.e., what she believes). For example,

(A) She may become slightly less in favor of trams.

(B) She may believe that if their city constructs the tram lines, then there will be less traffic jams

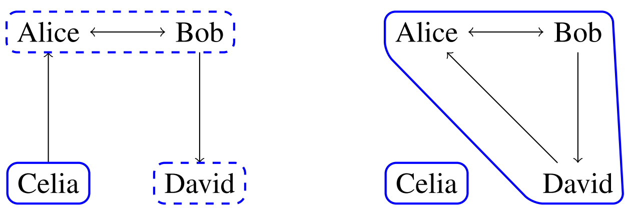


Fig. 2.1: Visualization of the network in the intuitive scenario before (left) and after (right) interactions. The arrow from a person p to another person q means that q listens to p (i.e., the arrow shows the direction of an opinion transmission). Blue dashed lines circumscribe strongly connected components; blue solid lines are echo chambers.

and higher taxes.

(C) If she considers higher taxes as bad, or worse than traffic jams, based on her values, she will be more negative about the topic.

(D) She may also feel uncomfortable and change her beliefs about how much traffic jams will be reduced to better support her new opinion [51, 65].

In addition, Alice may decide to stop listening to Celia and start listening to someone else with closer opinions and beliefs than Celia, e.g., David, instead. This can be achieved by stopping someone with the farthest opinions and beliefs (here Celia) and starting listening to someone who is the closest to Alice (David here). Thus, the network is updated from the left to the right in Figure 2.1. If she decides to do so, the next time she updates her opinions and beliefs, she will take Bob's and David's opinions and beliefs into consideration.

If they continue interacting in this way, the four people may be split in two groups such that people listen to other people in the same group but ignore those in different groups. People in the same group will share the same (or similar) opinions and beliefs. Such a state is called an "echo chamber".

Considering the story above, some questions can be asked:

(Q1) Does exchanging and updating opinions allow agents to form echo chambers as it is discussed in [55]?

(Q2) Does the same happen if they exchange beliefs?

(Q3) Moreover, does introducing beliefs help fighting against echo chambers, i.e., their number decreases?

(Q4) In addition, does keeping opinions and beliefs coherent affect this?

In order to answer these questions, a model of opinion and belief joint propagation and measures to precisely identify echo chambers are provided.

Chapter 3

State-of-the-Art

To answer the questions raised at the end of Chapter 2, it is needed to model agents' behaviors (A) to (D) in the chapter. Especially, opinion and belief propagation ((A) and (B), respectively) have been considered from the standpoint of opinion dynamics and distributed belief change. Thus, both fields as well as the definition of echo chambers and how they are detected are reviewed.

In this chapter, first, general concepts shared by various state-of-the-art models and the proposed model are presented (Section 3.1). Then, the definition of echo chambers and how they are detected in the existing work are reviewed (Section 3.2). After that, several classical opinion dynamics models (Section 3.3) and their application for modeling echo chambers (Section 3.4) as well as distributed belief change frameworks (Section 3.5) are presented. Finally, former work that tried to connect opinions and beliefs are presented (Section 3.6) before concluding this chapter (Section 3.7).

3.1 Preliminaries

Let A be a set of agents that correspond to people in our society. Each agent $a \in A$ has a non-empty set of neighbors $N_a^t \subseteq A \setminus \{a\}$ at time t. The directed network that agents form at time t is denoted as $G^t = \langle A, N^t \rangle$ such that $N^t = \bigcup_{a \in A} \{\langle a, a' \rangle; a' \in N_a^t\}$.

Each agent $a \in A$ also has a cognitive state $\langle O_a^t, B_a^t \rangle$ at time t made of its opinion O_a^t toward a topic ϕ and beliefs B_a^t. Only single topic is considered in this thesis. Opinions are represented as numbers between 0 and 1, i.e., $O_a^t \in [0, 1] \subseteq \mathbb{R}$. Similarly, beliefs are represented as a set of propositional formulae, i.e., $B_a^t \subseteq \mathcal{L}_P$ such that $\mathcal{L}_P$ is the set of propositional formulae constructed from the set P of atoms using the usual connectives. Moreover, the set of models of beliefs $B \subseteq \mathcal{L}_P$ is noted as $\mathcal{M}(B)$.

In the propositional settings, beliefs B can be replaced by a single formula $\bigvee_{\psi \in B} \psi \in \mathcal{L}_P$. Hence, hereafter, beliefs are also represented as the corresponding formula for simplicity.

In addition to $\mathcal{M}(\cdot)$, the beliefs (as a single formula) whose set of models is $M \subseteq \mathcal{M}(\top)$ are denoted as $[M]$. However, it is impossible to obtain unique beliefs from $[\cdot]$ because several beliefs can share the same set of models. For example, if $p, q \in P$, then $[\{m \in \mathcal{M}(\top); m \models p \wedge m \models q\}]$ can be $p \wedge q$, $q \wedge p$, $(p \wedge q) \vee (p \wedge q)$, etc. Hence, to consider the operation $[\cdot]$ as a function which maps set of models to the corresponding single formula, in this thesis, without loss of generalities, it is assumed that each element of $\mathcal{L}_P$ is represented as full disjunctive normal form (FDNF). To simplify the notation, the formula which is satisfied by all interpretations is denoted as $\top$. In other word, $\mathcal{M}(\top)$ is the set of all interpretations. Note that in this thesis, FDNF contains no redundancy and commutativity. Hence, there are $2^{|\mathcal{M}(\top)|}$ formulae in FDNF including the contradiction.

The topic ϕ is also represented as a set of propositional formulae ($\phi \subseteq \mathcal{L}_P$) or the corresponding formula $\phi \in \mathcal{L}_P$ in FDNF.

Example 1 (basic settings)**.** *Consider the four agents, Alice, Bob, Celia, and David, in Chapter 2. In the example, Bob and Celia are Alice's neighbors before they start interacting ($t = 0$), i.e., $N^0_{Alice} = \{Bob, Celia\}$.*

Each agent's initial beliefs can be represented as belief sets, for example, $B^0_{Alice} = \{p \to q\}$, $B^0_{Bob} = \{p \to \neg r\}$, $B^0_{Celia} = \{r \to \neg p\}$, and $B_{David} = \{(p \wedge q) \vee (\neg p \wedge r)\}$, where the three atoms p, q and r mean that 'a tram be built', 'traffic jams are low', and 'taxes are low', respectively. Only these atoms are used in this example, i.e., $P = \{p, q, r\}$. Hence, their beliefs can be represented as equivalent formulae in FDNF as follows:

$$\begin{aligned}
B^0_{Alice} =&(p \wedge q \wedge r) \vee (p \wedge q \wedge \neg r) \vee (\neg p \wedge q \wedge r) \vee (\neg p \wedge q \wedge \neg r) \vee \\
&(\neg p \wedge \neg q \wedge r) \vee (\neg p \wedge \neg q \wedge \neg r), \\
B^0_{Bob} =&(p \wedge q \wedge \neg r) \vee (p \wedge \neg q \wedge \neg r) \vee (\neg p \wedge q \wedge r) \vee (\neg p \wedge q \wedge \neg r) \vee \\
&(\neg p \wedge \neg q \wedge r) \vee (\neg p \wedge \neg q \wedge \neg r), \\
B^0_{Celia} =&(p \wedge q \wedge \neg r) \vee (p \wedge \neg q \wedge \neg r) \vee (\neg p \wedge q \wedge r) \vee (\neg p \wedge q \wedge \neg r) \vee \\
&(\neg p \wedge \neg q \wedge r) \vee (\neg p \wedge \neg q \wedge \neg r), \\
B^0_{David} =&(p \wedge q \wedge r) \vee (p \wedge q \wedge \neg r) \vee (\neg p \wedge q \wedge r) \vee (\neg p \wedge \neg q \wedge r).
\end{aligned}$$

Similarly, their opinions can be represented as $O^0_{Alice} = 1.0$, $O^0_{Bob} = 0.6$, $O^0_{Celia} = 0.4$, and $O^0_{David} = 0.5$. Finally, their discussion topic ϕ can be represented using the atoms above as $\{p\}$, or

$$\phi = (p \wedge q \wedge r) \vee (p \wedge q \wedge \neg r) \vee (p \wedge \neg q \wedge r) \vee (p \wedge \neg q \wedge \neg r)$$

in FDNF.

The distance between two opinions O and O' is measured by the absolute difference between their values:

$$d_O(O, O') = |O - O'|. \tag{3.1}$$

Similarly, the distance between two beliefs B and B' is defined as the Hamming distance between their models:

$$d_B(B, B') = |\mathcal{M}(B) \setminus \mathcal{M}(B')| + |\mathcal{M}(B') \setminus \mathcal{M}(B)|. \tag{3.2}$$

Agents receive opinions and beliefs from their neighbors only.

At each time t, agent a can determine its set C^t_a of concordant neighbors as those neighbors having close opinions and beliefs by:

$$C^t_a = \{a' \in N^t_a; d_O(O^t_a, O^t_{a'}) \leq \varepsilon \wedge d_B(B^t_a, B^t_{a'}) \leq \delta\} \tag{3.3}$$

such that $\varepsilon \in [0, 1]$ and $\delta \in \{0, 1, \ldots, |\mathcal{M}(\top)|\}$ are thresholds for opinions and beliefs, respectively.

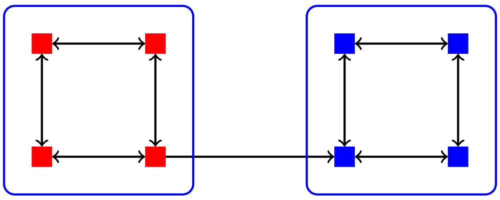


Fig. 3.1: Visualization of echo chambers. In this plot, there are two echo chambers (surrounded by blue lines). Different colors show different opinions and arrow from p to q means opinions of p are transmitted to q.

Example 2 (concordant neighbors)**.** *Assume that* $\varepsilon = 0.4$ *and* $\delta = 3$ *and consider agents in Example 1. Bob is a concordant neighbor of Alice, as* $d_O(O^0_{Alice}, O^0_{Bob}) = 0.4 \leq \varepsilon$ *and* $d_B(B^0_{Alice}, B^0_{Bob}) = 2 \leq \delta$*. However, Celia is not because* $d_O(O^0_{Alice}, O^0_{Celia}) = 0.6 > \varepsilon$ *and* $d_B(B^0_{Alice}, B^0_{Celia}) = 2 \leq \delta$*. Hence,* $C^0_{Alice} = \{Bob\}$*.*

3.2 Echo chambers in social networks

Results reported by existing literature support that echo chambers exist, while some does not [34]. Echo chambers in social networks describe a situation in which a group of persons strengthen their viewpoints by interacting only with persons who have opinions similar to theirs. Precisely, '*An echo chamber is [...] characterized by like-minded users predominantly interacting with each other. Within these echo chambers, users express and reinforce their beliefs on specific issues, thereby amplifying their shared viewpoints*' [45]. Typically, echo chambers can be visualized as Figure 3.1. Note that the term 'echo chamber' is used following the definition above and not as something inherently bad [25]. Sometimes other phrases, such as 'filter bubble', are used as synonyms of echo chambers [38].

Following this definition, echo chambers are groups of people satisfying three properties:

homogeneity means that group members share the same or very close opinions, or beliefs;

segregation means that group members mainly listen to people in the same group;

reinforcement means that group members' opinions or beliefs do not become more distant over time.

Concerning the last feature, there are other ways of interpreting "reinforcement". For example, it can be test whether agents believe that everyone believes the same things as their beliefs to see whether their beliefs are reinforced. Moreover, it is possible to test whether agents become more confident about their beliefs to see it. However, in this thesis, another approach is used: details will be discussed in Section 4.2.2.

There are various ways to attempt at measuring echo chambers [34]:

centrality measures which are used to detect influential agents in their network. This can be obtained by computing the degree of each node;

content-based metrics which compare nodes attributes (e.g., their beliefs or opinions);

homophily and polarization metrics which represent the tendency for agents to connect to others in the same group as them. For example, an agent's homophily is measured by computing the ratio of its neighbors who are in the same group as it [38];

graph clustering and partitioning methods measure how segregated the network is with respect to a given partition. Modularity [49] and random walk controversy [32] are the most well-known measures.

Among the three properties above, homogeneity can be assessed by a content-based metrics, but segregation is rather complementary to homophily and reinforcement is barely considered.

Hence, to identify echo chambers according to this definition, measures taking into account the three properties are proposed.

3.3 Opinion dynamics models

How echo chambers arise has been explained using mathematical models such as opinion dynamics models. Before reviewing such an application, classical models are introduced firstly.

Opinion dynamics studies how agents' opinions evolve through interactions [28]. The basic idea of such models is that agents' opinions are affected by other agents' opinions. The modality (A) in the intuitive scenario is an example of behavior which follows this principle. The "opinions" can be interpreted in various ways: not only subjective attitudes toward something, but also preferences or knowledge [14, 61].

There are various models that follow the idea. For example, in the DeGroot model [27], one of the simplest and most well-known classical models, opinions are updated as follows:

$$O_a^{t+1} = \sum_{a' \in A} w_{a,a'} O_a^t$$

where $w_{a,a'} \in \mathbb{R}$ is the weight of agents a on agent a' such that for all $a \in A$, $\sum_{a' \in A} w_{a,a'} = 1$ and for all $a, a' \in A$, $w_{a,a'} \geq 0$. In this model, the condition for agents to reach a consensus is proven theoretically. Here, consensus means that all agents' opinions are converged, i.e., there exists a number $O^* \in [0, 1]$ such that $\forall a \in A, \lim_{t \to \infty} O_a^t = O^*$.

Other classical models are the bounded confidence models [26, 36]. In the DW model [26], two agents a and a' revise their opinions if the difference between their opinions, i.e., $d_O(O_a^t, O_{a'}^t)$, is equal to or less than a threshold $\varepsilon \in \mathbb{R}$. The largest difference between this model and the DeGroot model is that agents can ignore others' opinions. The updating rule of the model can be described as:

$$\begin{aligned} O_a^{t+1} &= O_a^t + \mu(O_{a'}^t - O_a^t), \\ O_{a'}^{t+1} &= O_{a'}^t + \mu(O_a^t - O_{a'}^t), \end{aligned}$$

such that μ is a constant between 0 and 0.5 [28]. Based on this model, it is shown that low tolerance for dissimilar opinions, i.e., a small ε, prevents agents' opinions from converging.

The model by Hegselmann and Krause (as known as the HK model in these days) [36] is based on the similar idea. The main difference between the DW model and this model is that the latter enables agents to take opinions from various agents. The opinion updating process of this model is described as:

$$O_a^{t+1} = \frac{1}{|A_a^t|} \sum_{a' \in A_a^t} O_{a'}^t$$

where A_a^t is the set of agents with enough similar opinions with respect to a threshold ε, i.e., $A_a^t = \{a' \in A; d_O(O_a^t, O_{a'}^t) \leq \varepsilon\}$.

The process to update opinions can be synchronous, i.e., all agents update their opinions at the same time, or asynchronous, i.e., some agents can preserve current opinions without performing updating processes. For example, the DeGroot model is a synchronous model, while other models introduced so far are asynchronous.

Sometimes, opinions are represented as discrete values such as integers. For example, the voter model [24, 37] uses the opinion space $\{1, -1\}$ [35]. In this model, when agent a wants to update its opinions, it chooses one neighbor a' randomly and adopts its opinions, i.e., $O_a^{t+1} = O_{a'}^t$.

Classical models have been extended so that they can represent agents with opinions toward several topics. In this case, agent a's opinions are represented as a n-dimensional vector

$$(x_a(1), x_a(2), \ldots, x_a(n))^\top \in \mathbb{R}^n$$

where n is the number of topics and $x_a(i)$ is the a's opinions toward the i-th topic. Here, opinions are updated by considering (1) other agents' opinions toward the same topic, (2) other agents' opinions toward a different topic, and (3) the same agent's opinions toward a different topic [52, 15, 54].

With the advance of large language models (LLMs), some models consider LLM-based agents (e.g., [21, 66, 20]). In these models, agents have their personas and exchange their opinions and reasons to support their opinions. How their opinions are updated depends on the outputs of LLMs, hence it is difficult to show the process in a similar way to classical models.

3.4 Opinion dynamics models to explain echo chambers

Nowadays, opinion dynamics models based on classical bounded confidence models have been proposed to explain how echo chambers arise [60, 22]. A model with opinion threshold, timelines, and network rewiring has been proposed to model online social networks, and particularly Facebook, realistically [55]. In this model, each agent follows at least another agent. At each iteration, one agent (a) is chosen randomly from A and performs the following:

1. retrieve those posts from a's timeline which opinion is within the threshold (ε) from a's own opinion;

2. update a's opinions with opinions selected at step 1 as:

$$O_a^t = \begin{cases} \mu O_a^{t-1} + (1-\mu)\frac{\sum_{O \in T_a^t} O}{|T_a^t|} & \text{if } T_a^t \neq \emptyset, \\ O_a^{t-1} & \text{otherwise,} \end{cases}$$

   where $\mu \in [0, 1]$ and T_a^t is the set of opinions selected at the previous step;

3. remove an agent $a' \in N_a \setminus C_a$ who sent a discordant post from N_a and add another agent $a'' \in A \setminus (N_a \cup \{a\})$ to it if $N_a \setminus C_a$ and $A \setminus (N_a \cup \{a\})$ are not empty; otherwise N_a is kept as is. Intuitively, this means that a "unfollows" a' and "follows" a'' if possible. Various algorithms to choose a'' is considered: random (choose a'' randomly), repost (choose one of the authors of reposts contained in the timeline), and recommendation (choose a'' with close enough opinions);

4. submit a new post or repost an existing one in a's timeline.

The steps 3 and 4 are executed probabilistically. It is observed that the smaller ε is, the more echo chambers are created. Moreover, it is reported that network rewiring, i.e., unfollowing and following, based on recommendation or repost accelerates their appearance than simple rewiring algorithm based on random selection.

Using an extension of this model, Nagura and Akiyama [47] show that the fewer topics are discussed by agents and the more important they are, the more echo chambers agents form.

LLM-based opinion dynamics models are also applied to explain the phenomena [50, 64, 53]. It is shown that LLM-based agents can form echo chambers through their interactions.

3.5 Distributed belief change

In the field of belief change [29], many operators describe how to revise or merge beliefs. The modality (B) in Chapter 2 illustrates this. The most straightforward way to achieve this is to concatenate beliefs with $\wedge$. However, this does not work in general since such beliefs can be contradictory. Thus, various ways of merging beliefs are proposed to always obtain consistent ones.

Belief revision games (BRGs) [57] define procedures, similar to those of opinion dynamics, for merging beliefs. These operators merge several propositional beliefs into new beliefs. They can be described using the following components [42]:

- pseudo-distance $d\colon \mathcal{M}(\top) \times \mathcal{M}(\top) \to \mathbb{R}_{\geq 0}$;
- aggregation function $f\colon \mathbb{R}^n_{\geq 0} \to \mathbb{R}^m_{\geq 0}$;
- total preorder $\leq$ over $\mathbb{R}^m_{\geq 0}$;
- integrity constraints $B_0 \in \mathcal{L}_P$,

such that $n, m \in \mathbb{Z}_{>0}$. The pseudo-distance d can be extended to a function which maps beliefs $B \in \mathcal{L}_P$ and an interpretation $I \in \mathcal{M}(\top)$ to $\mathbb{R}_{\geq 0}$ as follows:

$$d(B, I) = \operatorname*{argmin}_{I' \in \mathcal{M}(B)} d(I, I').$$

Then, for given d, f, and $\leq$, the corresponding BRG operator $R^{d,f,\leq}$ is defined as [57]:

$$R^{d,f,\leq}(B_0; B_1, B_2, \ldots, B_n) = \left[\operatorname*{argmin}_{m \in \mathcal{M}(B_0)} (f(d(B_1, m), d(B_2, m), \ldots, d(B_n, m)), \leq) \right].$$

Note that in the initial definition, the syntax of beliefs that are the results of merging is not specified because the operator is designed to be syntax-independent [57] to satisfy the postulate **(IC3)** [43]. Moreover, the merged results by this operator are always consistent if B_0 is consistent, which is required by the postulate **(IC0)** [43].

One of the concrete examples of the BRG operators uses the Hamming distance d^H between interpretations defined as:

$$d^H(m, m') = |\{p \in P; (m \models p \wedge m' \not\models p) \vee (m \not\models p \wedge m' \models p)\}|,$$

summation, usual relation $\leq$ over $\mathbb{R}_{\geq 0}$, and $\top$ as the pseudo-distance, the aggregation function, the total preorder, and the integrity constraints, respectively. The operator is defined equivalently as [57]:

$$R(B_1, B_2, \ldots, B_n) = \left[\operatorname*{argmin}_{m \in \mathcal{M}(\top)} \sum_{k=1}^{n} \min_{m' \in \mathcal{M}(B_k)} d^H(m, m') \right]. \tag{3.4}$$

The operator is denoted as R instead of the notation in [57] to make the notation simpler because no other operators from BRG will be used. In BRG, some properties, including belief stability, are proved theoretically. Note that the property of belief stability, i.e., agents' beliefs are not changed after enough interactions, is called "convergence" in [57]. Such property is called as 'stability' in this thesis to distinguish from the DeGroot-wise convergence.

These BRG operators are special cases of merging operators, but do not satisfy the AGM revision postulates [12] in general [57]. In particular, the operator R defined above is not a revision operator because it does not satisfy the AGM properties [57].

Belief Flow Networks (BFNs) [59] is another framework which enables beliefs to be updated asynchronously. It uses existing improvement operators which are generalizations of belief revision operators [44, 41] to revise agents' beliefs. Examples of the operators are the restrained revision operator [16] and the one-improvement operator [41]. They map a total preorder over $\mathcal{M}(\top)$ and a beliefs $B \in \mathcal{L}_P$ to a total preorder over $\mathcal{M}(\top)$ [59]. It is shown that agents reach a consensus if the network they form is strongly connected. However, the definition of "consensus" in BFN is different from the definition in the DeGroot model because in BFN, all agents' beliefs are not necessarily the same as each other even they reach a consensus [59].

Dynamic doxastic-epistemic logics [63] can model agents exchanging knowledge and beliefs through, public and/or private, announcements. Such logics have been designed to perform belief revision when announced statements contradict an agent's beliefs [13].

None of these procedures allow agents to modify the network. Network rewiring has been considered in the network knowledge base approach on a simpler belief description language [31]. These work have not considered the impact of belief change on the formation of echo chambers.

One of the application of distributed belief change is cultural knowledge evolution [18, 19]. In this area, agents have their knowledge represented as description logics and exchange them through working on common tasks. Through this framework, it is shown that agents can make the same decision while they do not have to share the same ontology (i.e., knowledge or beliefs).

3.6 Toward distinguishing opinions and beliefs

In spite of socio-psychological evidence of their relation [51, 65], there has been few models explicitly connecting changing opinions and beliefs as defined in opinion dynamics and belief propagation. Some models constrain Boolean public opinions with common integrity constraints expressed logically [17]. Other work has used an influence matrix between opinions in order to enforce their dependencies [30, 48]. However, these constraints do not have the status of beliefs and do not change.

In [62], the distinction is made between beliefs and experience. Experience is modeled as data incompletely sampling the environment. From samples, agents learn beliefs, under the form of a classifier based on feature correlation. Then they explicitly exchange both data and classifier with their neighbors. They adopt neighbor classifiers only if it is more accurate on their own experience. This procedure is called 'epistemic rationality', meaning that agents try to maintain as much coherence as possible between their beliefs and their experience. It is shown theoretically that agents converge towards the same beliefs but that this can be suboptimal in terms of accuracy. Although, this work does not deal with opinions, only belief and data, it provides the notion of internal coherence of agents' beliefs and experience.

3.7 Conclusion

In this chapter, relevant work in echo chambers, opinion dynamics, and distributed belief change is reviewed. From this, no model clearly distinguishes opinions from symbolic beliefs. However, it is required to model the typical interactions between people to answer the questions raised at the end of Chapter 2: when people express their opinions, they usually add the reasons to support them as presented in Example 1. Hence, in the next chapter, a new model is introduced.

Chapter 4

A model for opinion and belief propagation and integration

As reviewed in the previous chapter, there are no models which connect opinions and beliefs with clear distinction between them. However, such models are necessarily to analyze the dynamics of agents with opinions and beliefs to justify them. Hence, in this chapter, a model with this distinction is introduced. However, the difference between opinions and beliefs is sometimes vague in the literature. How they are distinguished is discussed in Section 4.1.

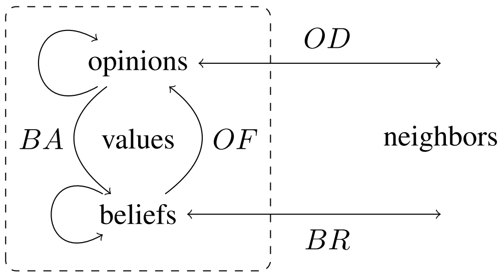


Fig. 4.1: The four procedures (BR, OD, OF, and BA) for interactions between opinions and beliefs.

As discussed in Chapter 2, agents can change their opinions based on their beliefs and vice versa, as well as based on opinions and beliefs from their neighbors. Hence, four procedures OD, BR, OF, and BA corresponding to the four possible actions (A) to (D) of agents in the intuitive scenario (Chapter 2) are defined. The procedures are visualized in Figure 4.1. These functions map an agent's cognitive state to a new state. They are separated into two types of functions: belief and opinion propagation (Section 4.2) and coherence restoration (Section 4.3). Here, coherence means that an agent has opinions that are obtained based on their beliefs and vice versa. More precise definitions will be presented in Section 4.3.2.

Note that the model is defined following classical mathematical modeling approach, not using LLM-based agents. This is because how LLM-based agents change their opinions and beliefs are opaque, hence this would not allow us to understand which factor(s) contribute to the obtained results.

4.1 Distinction between opinions and beliefs

As discussed before (Section 3.3), there are several interpretations of the term "opinions". In addition, sometimes the words "opinions" and "beliefs" are used as synonyms. For example, subjective logic, an extension of probabilistic logic that can represent the confidence of agents about whether something is true [40], is applied to represent opinions [33]. Moreover, in some papers in opinion dynamics (e.g., [26]), the term beliefs is used to show opinions. The same thing can happen in belief change (e.g., [58]). Hence, it is important to clarify the meaning of opinions and beliefs.

In this thesis, agent a's *beliefs* are what a regards as true about the environment; agent a's *opinions* toward a topic ϕ are its judgment about how desirable if ϕ becomes true in the future. The higher and lower opinions are, the more positive and negative agents are toward the topic, respectively. In other word, beliefs describe the current understanding of an agent about the environment, while opinions show agents' attitudes toward possible environments in the future. Note that opinions and beliefs are not necessarily coincide: agents who believe that there are no tram lines can consider building trams as something good. If they should be coincide, opinions should be interpreted as confidence about something for example.

Such distinctions allow us to investigate how interactions between opinions and beliefs creates phenomena in our society including echo chambers. However, this approach can restrict the generalities of the proposed model: for example, opinion dynamics models can be used to explain how agents update their confidence about something [30].

Here, the word beliefs, not knowledge, is used in this thesis. This is because knowledge means true beliefs with respect to the environment (i.e., facts), while in the proposed model, there are no facts defined by the environment. For example, both of "University of Tsukuba is located in Japan" and "University of Tsukuba is located in France" can be beliefs of agents, while only the former is knowledge as the latter is false.

4.2 Propagation

Before introducing the cognitive operations based on the distinction of opinions and beliefs discussed in the previous section, first, the propagation processes are defined. In these processes, agents can revise their opinions by taking other agents' opinions into account: this corresponds to classical opinion dynamics models. Similarly, beliefs can be revised based on others' beliefs: this process is formalized by classical belief change framework. Hence, existing procedures are simply used to define them. This means that this model is a generalization of existing frameworks.

4.2.1 Opinion dynamics

In the opinion dynamics process (OD), agents update their current opinions using the current opinions of their concordant neighbors. Here, the opinion updating rules of the model proposed by [55] is used because this is a simple and enough realistic model. Thus, OD is defined as:

$$OD(\langle O_a, B_a\rangle) = \begin{cases} \left\langle \mu O_a + (1-\mu) \times \frac{\sum_{a' \in C_a^t} O_{a'}}{|C_a^t|}, B_a \right\rangle & \text{if } C_a^t \neq \emptyset \\ \langle O_a, B_a\rangle & \text{if } C_a^t = \emptyset \end{cases} \tag{4.1}$$

such that $\mu \in [0, 1]$. Note that filtering with thresholds (ε and δ) is achieved while computing C_a^t from N_a^t.

Here, continuous opinions are used, i.e., the range of opinions is $[0, 1]$, not discrete opinions. This is because the former can entail the latter, and thus the former is more expressive than the other. Note that the range has been changed from [55] to this model ($[-1, 1]$ to $[0, 1]$) considering the synchronizing process OF (Section 4.3.2), but this does not raise problems here. To represent opinions taken from the interval $[-1, 1]$, changing scale is enough.

Example 3 (opinion dynamics)**.** *Consider agents in Example 1. Assume that $\mu = 0.5$. If Alice performs OD, since $C^0_{\mathit{Alice}} = \{\mathit{Bob}\}$ (Example 2), her opinions after performing OD are*

$$0.5 \cdot O^0_{\mathit{Alice}} + (1 - 0.5)O^0_{\mathit{Bob}} = \frac{1.0 + 0.6}{2} = 0.8.$$

Hence, after OD, her cognitive state is changed from $\langle O^0_{\mathit{Alice}}, B^0_{\mathit{Alice}}\rangle$ to $\langle 0.8, B^0_{\mathit{Alice}}\rangle$.

4.2.2 Belief revision

In the belief revision process (BR), agents update their beliefs using the current beliefs of their concordant neighbors. Here, one of the operators in BRG [57] and in particular the operation R (Equation 3.4) is used as this is a simple model. Thus, BR is defined as:

$$BR(\langle O_a, B_a\rangle) = \langle O_a, R(B_a, B_{a_1}, \ldots, B_{a_n})\rangle \tag{4.2}$$

such that $C_a^t = \{a_1, \ldots, a_n\}$.

Here, beliefs are represented as a single propositional formula. This is because it is the simplest among the logical formulae (including formulae in first-order logic or in modal logics) and the process of revising such beliefs is defined constructively. Especially, simplicity is the most important because this model is built to perform experiments.

However, this propositional approach restricts the interpretation of "reinforcement" in the three features of echo chambers. For example, as discussed in Section 3.2, "reinforcement", one of the characteristics of echo chambers, can be tested by checking whether agents consider that everyone's beliefs are the same as theirs. More expressive logics than the propositional one that allows us to represent the 'second-order' beliefs are required to perform such tests.

Example 4 (belief revision)**.** *Consider again agents in Example 1. If Alice performs BR, since $C^0_{\mathit{Alice}} = \{\mathit{Bob}\}$ (Example 2), her beliefs after performing BR are*

$$R(B^0_{\mathit{Alice}}, B^0_{\mathit{Bob}}) = \{p \to (q \land \neg r)\},$$

or $(p \land q \land \neg r) \lor (\neg p \land q \land r) \lor (\neg p \land q \land \neg r) \lor (\neg p \land \neg q \land r) \lor (\neg p \land \neg q \land \neg r)$ in FDNF. Hence, after BR, her cognitive state is changed from $\langle O^0_{\mathit{Alice}}, B^0_{\mathit{Alice}}\rangle$ to $\langle O^0_{\mathit{Alice}}, \{p \to (q \land \neg r)\}\rangle$.

4.2.3 How to obtain the existing models

Our model is a generalization of the existing models [55, 57]. The existing opinion dynamics model [55] can be obtained from the model by setting $\delta = |\mathcal{M}(\top)|$ and performing OD because beliefs are not taken into account when obtaining C_a^t. Similarly, setting $\varepsilon = 1$ and performing BR corresponds to the existing belief revision model [57].

4.3 Synchronization

After propagation, agents' opinions and beliefs may not correspond to each other. This loss of coherence may be uncomfortable for agents as they may not be able to justify their opinions from their beliefs. For example, if an agent believes that constructing tram lines lead to increase taxes and it regards higher taxes as bad, it cannot have positive opinions toward the construction because opinions do not support beliefs and vice versa.

In order to restore coherence (or 'epistemic rationality' in [62]), two synchronization operations are defined: opinion formation from beliefs (OF, Section 4.3.2) and belief alignments with opinions (BA, Section 4.3.3). Because opinions and beliefs are not direct translations of each others, these operations are based on the values that agents may hold (Section 4.3.1).

4.3.1 Values

Values (or cultural values) help agents to determine whether something is 'good' or 'bad' [56, 46]. They are used here to determine preferences between different possible states of the world. This is rendered as a map V from interpretations in $\mathcal{M}(\top)$ to real numbers between 0 and 1. Intuitively, the values toward an interpretation I is close to 0 (resp. 1), if the agent considers I as bad (resp. good). For example, Table 4.1 represents values V which consider worlds consistent with p, then q, then r, as good. These may express preferences between 'reducing traffic jams' and 'reducing taxes' in Example 1.

Such values can be defined based on the strict total order $<$ over P ($r < q < p$ in the Table 4.1). Here, the order $<$ works as a preference order over P. Then, the strict total order $\ll$ over $\mathcal{M}(\top)$ is defined with $<$ as follows:

$$m \ll m' \iff \exists p \in P, \begin{cases} \forall q \in P \text{ s.t. } p < q, m(q) = m'(q) \\ m(p) = \text{f} \wedge m'(p) = \text{t} \end{cases}$$

with $m(p) = \text{t}$ if $m \models p$ and $m(p) = \text{f}$ otherwise. Then, values V can be defined as:

$$V(m) = \frac{|\{m' \in \mathcal{M}(\top); m' \ll m\}|}{|\mathcal{M}(\top)| - 1}.$$

This approach enables us to focus on the preferences between atoms. This is useful because this is simpler and easier than considering $\ll$ directly, especially when experiments are designed so that they represent specific situations. This can help us to interpret the results with respect to the values.

However, one of the backwards of this definition is that it forces for the interval between values of a model and its successor to remain constant ($1/(\mathcal{M}(\top) - 1)$). This prevents us from testing the

Table 4.1: A score V is provided to each Boolean assignment of three atoms.

model	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
p	t	t	t	t	f	f	f	f
q	t	t	f	f	t	t	f	f
r	t	f	t	f	t	f	t	f
V	$\frac{7}{7}$	$\frac{6}{7}$	$\frac{5}{7}$	$\frac{4}{7}$	$\frac{3}{7}$	$\frac{2}{7}$	$\frac{1}{7}$	$\frac{0}{7}$

Table 4.2: One example of scoring V' with atoms with different importance (p is very important but the effect of q and r is very tiny).

p	t	t	t	t	f	f	f	f
q	t	t	f	f	t	t	f	f
r	t	f	t	f	t	f	t	f
V	$\frac{100}{100}$	$\frac{99}{100}$	$\frac{98}{100}$	$\frac{97}{100}$	$\frac{3}{100}$	$\frac{2}{100}$	$\frac{1}{100}$	$\frac{0}{100}$

effect of the importance of the atoms on the final opinions/beliefs or echo chambers. For example, if the agents consider that the environment with p and $\neg p$ is extremely good and bad for them, respectively, the values V' may be defined as Table 4.2. The two values V' and V share the same preference order ($\ll$) between interpretations. Here, if $m \models p$, then $V'(m)$ is very close to 1 and if otherwise, $V'(m)$ is very close to 0; the effect of other atoms q and r is tiny compared to the effect of p. Freezing the interval does not allow to use such values and thus prevents us from testing whether all agents can agree (in the sense of BFN [59]) that p is true.

Such values allow agents to form their opinions from beliefs and vice versa as discussed hereafter.

4.3.2 Opinion formation

In the opinion formation process, an agent revises its opinions based on its beliefs. It consists first in generating an opinion from the current beliefs (ϕ is the topic on which to form an opinion):

$$v(B) = \frac{\sum_{m \in \mathcal{M}(R(B,\phi))} V(m)}{|\mathcal{M}(R(B,\phi))|}. \tag{4.3}$$

This means that an agent forms its opinion about the topic ϕ from its beliefs B based on how valuable the models of the revision of B by ϕ, i.e., the state when ϕ becomes true from B, are. Note that the divisor of $v(B)$ is never 0 since R returns consistent beliefs [57].

Based on the function v defined above, the coherence of a cognitive state can be defined formally: opinions O and beliefs B are *coherent* if $O = v(B)$.

The generated opinion is then aggregated to the agent's current opinion. The function OF is thus defined as:

$$OF(\langle O_a, B_a \rangle) = \langle \alpha O_a + (1-\alpha) \times v(B_a), B_a \rangle \tag{4.4}$$

with $\alpha \in [0,1]$ controlling the inertia to the changing process. This process is a simple way that allows us to control the willingness to changing opinions. However, this process is not idempotent if $\alpha \in]0,1[$.

The function v may be defined in a different way:

$$v'(B) = \begin{cases} \frac{\sum_{m \in \mathcal{M}(B \wedge \phi)} V(m)}{|\mathcal{M}(B \wedge \phi)|} & \text{if } B \text{ and } \phi \text{ are consistent,} \\ 0 & \text{otherwise.} \end{cases}$$

In fact, v and v' are equivalent if B and ϕ are consistent since $R(B,\phi) \equiv B \wedge \phi$ in this case [57]. However, v' is not used in this thesis instead of v because v' always returns 0 if ϕ and B are inconsistent. If v is used, opinions which are larger than 0 can always obtained in general, and thus it is possible to analyze the process of evolution of agents' opinions in more details.

Example 5 (opinion formation)**.** *Consider Alice after performing BR in Example 4. Now her cognitive state is $\langle 1.0, \{p \to (q \wedge \neg r)\}\rangle$. To adapt her opinions towards the topic $\phi = p$ from her new beliefs, she first revises her current beliefs by ϕ: $R(p \to (q \wedge \neg r), \phi) = p \wedge q \wedge \neg r$. Assume that her values V are those defined in Table 4.1, she can compute the mean of the values $V(m)$ where $m \in \mathcal{M}(R(p \to (q \wedge \neg r), \phi))$, hence $v(p \to (q \wedge \neg r)) = \frac{6}{7}$. Based on this beliefs-based opinions, with $\alpha = 0.5$, she obtains her new opinions by:*

$$\alpha \cdot 1.0 + (1-\alpha) \cdot \frac{6}{7} = \frac{13}{14} \approx 0.93.$$

Hence, her new cognitive state after aligning opinions and beliefs by OF is $\langle \frac{13}{14}, \{p \to (q \wedge \neg r)\}$.

4.3.3 Belief alignment

In the belief alignment process, agents revise their beliefs based on their opinions. This is achieved by filtering beliefs in two steps. First, agents select the beliefs which minimize the distance between their current opinions O_a^t and the opinions based on beliefs from the function v of Equation 4.3. Second, agents adopt those beliefs which minimize the distance between their current beliefs B_a^t and the beliefs selected in the previous step. This step is introduced to reduce the randomness of this process.

Thus, the function BA is defined as:

$$BA(\langle O_a, B_a\rangle) = \langle O_a, \operatorname*{argmin}_{B \in \mathcal{L}_P(O_a)} d_B(B, B_a)\rangle \tag{4.5}$$

such that

$$\mathcal{L}_P(O) = \operatorname*{argmin}_{B \in \mathrm{FDNF}^*(\mathcal{L}_P)} d_O(O, v(B))$$

and $\mathrm{FDNF}^*(\mathcal{L}_P)$ is the set of consistent FDNF formulae. This is because the belief merging operator R is not defined for inconsistent beliefs, hence v is also not defined in this case. If several beliefs minimize $d_B(B, B_a)$ in Equation 4.5, the aligned result is chosen randomly among them.

Example 6 (belief alignment)**.** *Consider Alice after performing OD (Example 3) with the cognitive state $\langle 0.8, \{p \to q\}\rangle$. When she performs belief alignment, she first computes the set of beliefs that yield coherent and the closest opinions to the current one, i.e., $\mathcal{L}_P(0.8)$. Here, the set is:*

$$\mathcal{L}_P(0.8) = \{B \in \mathcal{L}_P; m_1 \models B, m_2 \models B, m_3 \not\models B, m_4 \models B\}$$

where m_1, m_2, m_3, and m_4 are defined in Table 4.1. Then, she chooses the closest beliefs to her current one among $\mathcal{L}_P(0.8)$. In this case, $\{(p \wedge r) \rightarrow q\}$ or

$$(p\wedge q\wedge r)\vee(p\wedge q\wedge\neg r)\vee\underline{(p\wedge\neg q\wedge\neg r)}\vee(\neg p\wedge q\wedge r)\vee(\neg p\wedge q\wedge\neg r)\vee(\neg p\wedge\neg q\wedge r)\vee(\neg p\wedge\neg q\wedge\neg r)$$

in FDNF minimizes the distance to $\{p \rightarrow q\}$ as the distance is 1 (the underlined disjunct is the difference). Hence, her new cognitive state is $\langle 0.8, \{(p \wedge r) \rightarrow q\}\rangle$.

4.4 Protocol

The processes to update agents' cognitive states are defined in the previous sections. In this section, how agents behave in each interactions including the update of their states is introduced. These processes are adapted from [55] (see Section 3.4).

At time t, the set $A^t \subseteq A$ of activated agents is determined by drawing agents in A with probability p_{activate}. Each non-active agent ($a \notin A^t$) does nothing at time t; active agents will update their opinions and beliefs, and possibly their set of neighbors.

Four experimental processes are distinguished to test the questions in Chapter 2: (odonly) (resp. (bronly)) only performs opinion dynamics (resp. belief revision); (both) performs both operations and (coherence) applies, in addition, opinion formation and belief alignment to synchronize them:

$$\langle O^t_a, B^t_a\rangle = OD(\langle O^{t-1}_a, B^{t-1}_a\rangle), \quad \text{(odonly)}$$
$$\langle O^t_a, B^t_a\rangle = BR(\langle O^{t-1}_a, B^{t-1}_a\rangle), \quad \text{(bronly)}$$
$$\langle O^t_a, B^t_a\rangle = BR(OD(\langle O^{t-1}_a, B^{t-1}_a\rangle)), \quad \text{(both)}$$
$$\langle O^t_a, B^t_a\rangle = OF(BR(BA(OD(\langle O^{t-1}_a, B^{t-1}_a\rangle)))). \quad \text{(coherence)}$$

All functions are executed synchronously among the set of active agents. For example, if agents follow the process (both), agents revise beliefs based on their beliefs after performing OD.

After the cognitive states are updated, as in [55], with the probability p_{rewire}, agent $a \in A^t$ tries to modify its set of neighbors. This process is based on an order $\prec^t_a$ over A which is defined based on the distance between agents' internal states. Since such states contain opinions and beliefs, four kinds of such order $a' \prec^t_a a''$ defined as below are considered:

$$d_O(O^t_a, O^t_{a'}) < d_O(O^t_a, O^t_{a''}) \quad \text{(oponly)}$$
$$d_B(B^t_a, B^t_{a'}) < d_B(B^t_a, B^t_{a''}) \quad \text{(belonly)}$$
$$\begin{cases} d_O(O^t_a, O^t_{a'}) < d_O(O^t_a, O^t_{a''}), \text{or} \\ d_O(O^t_a, O^t_{a'}) = d_O(O^t_a, O^t_{a''}) \wedge d_B(B^t_a, B^t_{a'}) < d_B(B^t_a, B^t_{a''}) \end{cases} \quad \text{(opbel)}$$
$$\begin{cases} d_B(B^t_a, B^t_{a'}) < d_B(B^t_a, B^t_{a''}), \text{or} \\ d_B(B^t_a, B^t_{a'}) = d_B(B^t_a, B^t_{a''}) \wedge d_O(O^t_a, O^t_{a'}) < d_O(O^t_a, O^t_{a''}) \end{cases} \quad \text{(belop)}$$

In other words, Equation (opbel) and Equation (belop) are the lexicographical orders considering opinions and then beliefs, and beliefs and then opinions, respectively.

Firstly, a chooses randomly one agent among its neighbors a' which is maximal with respect to $\prec_a^t$:

$$a' = \operatorname*{argmax}_{\prec_a^t} N_a^{t-1}.$$

Secondly, a chooses randomly one agent $a'' \in A \setminus (N_a^{t-1} \cup \{a\})$ among the non-connected agents which are minimal with respect to $\prec_a^t$:

$$a'' = \operatorname*{argmin}_{\prec_a^t} A \setminus (N_a^{t-1} \cup \{a\}).$$

Then, a updates its set of neighbors if excluding a' from it and including a'' to it helps it to reduce the distances among its neighbors:

$$N_a^t = \begin{cases} (N_a^{t-1} \setminus \{a'\}) \cup \{a''\} & \text{if } a'' \prec_a^t a', \\ N_a^{t-1} & \text{otherwise.} \end{cases}$$

Note that for each $a \in A$, the size of N_a^t remains constant over t. If there are several minimal or maximal elements with respect to $\prec_a^t$, one of them are selected randomly.

This rewiring algorithm is similar to the recommendation algorithm (i.e., choosing one concordant agents among $A \setminus (N_a^{t-1} \cup \{a\})$) in [55]. However, in the algorithm, a'' is not necessarily concordant to a. This allows agents to make their neighbors' opinions and beliefs as close to theirs as possible even if no agents except for themselves are concordant. Moreover, in the recommendation algorithm in [55], the opinions of a'' are not necessarily the closest to those of a, while (oponly) and (opbel) enforce this.

Example 7 (network rewiring)**.** *Assume that Alice performs network rewiring after performing OD and BA in this order and only Alice is activated. Recall that Alice's, Bob's, Celia's, and David's cognitive states are now $\langle 0.8, \{(p \wedge r) \to q\}\rangle$ (see Example 6), $\langle 0.6, \{p \to \neg r\}\rangle$, $\langle 0.4, \{r \to \neg p\}\rangle$, and $\langle 0.5, \{(p \wedge q) \vee (\neg p \wedge r)\}\rangle$ (see Example 1), respectively. If Alice uses the order $\prec_{Alice}^0$ defined as Equation (opbel), $a' = Celia$ because the distance between Alice's and Bob's opinions is smaller than it from Alice to Celia. Moreover, $a'' = David$ as $A \setminus (N_{Alice}^0 \cup \{Alice\}) = \{David\}$. Finally, David $\prec_{Alice}^0$ Celia as David has closer opinions to Alice than Celia. Hence, in this case, Alice unfollows Celia and follows David instead of Celia. This means that she changes the set of her neighbors by:*

$$N_{Alice}^1 = \{Bob, Celia\} \setminus \{Celia\} \cup \{David\} = \{Bob, David\}.$$

The network after this rewiring is shown in Figure 2.1 (right).

4.5 Conclusion

In this chapter, the model (visualized as Figure 4.1) which clearly distinguishes opinions from beliefs is defined (Section 4.1). In this model, agents propagate their opinions (Equation 4.1) and beliefs (Equation 4.2). Moreover, they can restore the coherence between their opinions and beliefs through the values they have (Equation 4.4 and Equation 4.5).

In the next chapter, the hypotheses based on the questions in Chapter 2, new measures to capture echo chambers, and the experimental plan to test the hypotheses with the measures are explained.

Chapter 5

Experimental plan

In the previous chapter, the model with the dual propagation of opinions and beliefs is defined. In this chapter, the experimental plan to answer questions raised in Chapter 2 is introduced. More specifically, in Section 5.1, the hypotheses which correspond to the questions are stated. In Section 5.2, new measures to capture echo chambers following their definition are introduced (Section 3.2). Then, in Section 5.3, how the parameters are set for each experiment is shown. In Section 5.4, the hypotheses are reformulated with the measures. Finally, how the simulator used to perform the experiments is implemented is described in Section 5.5.

5.1 Hypotheses

To answer the questions that were raised at the end of the Chapter 2 within the framework, four hypotheses are considered.

First, the protocol is based on updates of opinions and beliefs based on bounded confidence models and network rewiring. In the previous work, it is shown that protocols which contain both creates echo chambers [55]. Because of this, it is expected that echo chambers can be observed in the proposed model if opinions are updated independently of beliefs. Thus, the first hypothesis is:

H0 If opinions evolve independently of beliefs, opinion echo chambers should arise.

Using a similar procedure, it is expected that, the same will occur to beliefs. Thus, the second hypothesis is:

H1 If beliefs evolve independently of opinions, belief echo chambers should arise.

Next, if agents evolve their opinions and beliefs without synchronizing them, it is expected that agents will form more echo chambers than before. It is observed that less tolerant agents will form more polarized opinions in the context of opinion dynamics [55, 26]. Here, the condition to filter neighbors (i.e., $d_O(O, O') \leq \varepsilon \wedge d_B(B, B') \leq \delta$ which appears in the definition of C_a^t; see Equation 3.3) is as strict as or more strict than the condition $d_O(O, O') \leq \varepsilon$ or $d_B(B, B') \leq \delta$ since this considers opinions and beliefs. Thus, the third hypothesis is:

H2 Considering opinions and beliefs when agents filter their neighbors creates more echo chambers than considering opinions or beliefs.

Finally, keeping opinions and beliefs coherent is expected to tighten the links between them and to induce more echo chambers than without keeping the coherence. This is because this helps agents to resist adopting their opinions and beliefs to opinions and beliefs from other agents. In other word, it is expected that coherent cognitive states work as attractors of agents' opinions and beliefs, hence agents' final states are discretized. This helps agents to connect to others who already have enough similar cognitive states before opinions and beliefs are fused through the interactions. Thus, the final hypothesis is:

H3 Connecting opinions and beliefs and keeping them coherent increases the number of echo chambers.

These hypothesis **H0** to **H3** correspond to the questions (Q1) to (Q4) in Chapter 2, respectively.

5.2 Measures

To test the hypotheses presented in the previous section, it is necessary to identify echo chambers.

Different measures from those of [55] are used (hence the importance of **H0**). In [55], the emergence of echo chambers is shown visually using plots of the evolution of message diversity, agents' opinions, and the network of agents over time. These measures cover the three requirements of echo chambers introduced in Section 3.2. However, these plots are generated from the results of one single experiment. While it is possible to aggregate the opinion diversity in such plots by computing their mean, this is impossible for the other two measures. Thus, in this thesis their measures are not used.

In the same paper, the effects of ε on the number of final opinion peaks and final opinion distance are presented based on 20 experiments which share the same parameters. These measures show how opinions are polarized and the network segregated. Although polarization may be thought of as a way to measure homogeneity, reinforcement is ignored, so this should not allow us to conclude on echo chambers.

As discussed in Section 3.2, other typical measures used in literature are not necessarily based on the characteristics of echo chambers. Hence, new measures are introduced and it is tested whether our measure can allow us to reach the same conclusion as existing work.

The way to identify echo chambers is to partition the network into groups and to count those satisfying the properties of homogeneity, segregation and reinforcement. Of course, it is not practicable to test all partitions, the measures are based on strongly connected components to identify echo chambers. Basing measurements on strongly connected components makes sense because: (1) echo chambers are expected to be connected, (2) they are maximal, and (3) reinforcement is usually obtained through feedback, which can be obtained in a strongly connected component. Moreover, strongly connected components can be obtained deterministically and basing on strongly connected components is a simpler way than recently proposed community detection algorithms (see [39] for example). Hereafter, *components* denote strongly connected components. Let $\mathcal{S}^t$ be the set of components.

Other ways of defining components (in the sense the groups of agents in G^t) are possible. For example, defining components based on modularity [23] is a common way to identify communities. This algorithm is not used in this thesis because it is based on undirected network, while the network G^t agents form is directed. Yet the undirected one based on it can be obtained by $\langle A, N^t \cup \{\langle a', a\rangle; \langle a, a'\rangle \in N^t\}\rangle$, this loses the information about the direction of the edges. Moreover, the novelty of the measures proposed here is not the introduction of strongly connected components. It is to count the number of components with the three features of echo chambers.

5.2.1 Segregation

To assess how *segregated* the network is, the ratio of the edges whose sources are in a component C and whose targets [1] are not in C to the edges whose sources are in C is computed:

$$L^t(C) = \frac{|\{\langle a, a'\rangle \in N^t; a \in C \wedge a' \notin C\}|}{|\{\langle a, a'\rangle \in N^t; a \in C\}|}.$$

Hence, the lower this measure, the more segregated the network.

5.2.2 Homogeneity

To assess how *homogeneous* agents in a component $C \in \mathcal{S}^t$ are, the maximal distance between opinions (resp. beliefs) of agents which belong to the component is measured:

$$M^t_O(C) = \max_{a,a' \in C} d_O(O^t_a, O^t_{a'}),$$
$$M^t_B(C) = \max_{a,a' \in C} d_B(B^t_a, B^t_{a'}).$$

The lower these measures, the more homogeneous the component. The homogeneity is not measured in edge level by computing the difference of opinions and beliefs of a and a' where $\langle a, a'\rangle \in N^t$ because community level homogeneity defined above is stronger than edge level ones.

5.2.3 Reinforcement

To assess if a component C satisfies *reinforcement*, whether its homogeneity measure (maximal distance between agents' opinions or beliefs) decreases during its lifespan is tested. Let $[t_C, t]$ be the maximal window during which a component $C \in \mathcal{S}^t$ exists, then, the measure is defined as:

$$D^t_O(C) \equiv \forall s \in [t_C, t[, M^s_O(C) \geq M^{s+1}_O(C)$$
$$D^t_B(C) \equiv \forall s \in [t_C, t[, M^s_B(C) \geq M^{s+1}_B(C)$$

If different systems are used to represent opinions or beliefs, other tests can be possible. For example, if beliefs were represented in multi-agent dynamic doxastic logic, it is possible to test whether each agent believes that everyone has the same beliefs as its beliefs to conclude that agents' beliefs are reinforced.

[1] For an edge $e = \langle a, a'\rangle \in N^t$, a and a' are the source and the target of e, respectively.

5.2.4 Echo chambers

Based on the above measures, the number of echo chambers at time t can be computed as those components that are segregated and whose opinions and beliefs are homogeneous and reinforcing:

$$
\begin{aligned}
eo^t &= |\{C \in \mathcal{S}^t; \overbrace{L^t(C) \leq \theta}^{\text{segregation}} \wedge \overbrace{M_O^t(C) \leq 10^{-3}}^{\text{homogeneity}} \wedge \overbrace{D_O^t(C)}^{\text{reinforcement}}\}|,\\
eb^t &= |\{C \in \mathcal{S}^t; L^t(C) \leq \theta \wedge \quad M_B^t(C) = 0 \quad \wedge D_B^t(C)\}|,
\end{aligned}
$$

such that $\theta \in \,]0,1[$. The measures are made independent of ε and δ, which vary in experiments, by choosing more strict thresholds used to test the homogeneity (10^{-3} and 0). In the following, $\theta = 0.5$ is used. Here, $L^t(C) \leq 0.5$ is equivalent to the situation in which there are more edges whose sources and targets are in C than edges whose targets are not included in C. Note that this automatically disqualifies singleton (components C such that $|C| = 1$) as echo chambers. This is because such components $\{a\}$ do not satisfy the segregation property since $L^t(\{a\}) = 1 > \theta$. In complement, the proportion of agents involved in echo chambers is computed.

5.3 Parameters

In order to test each of the four hypotheses of Section 5.1, six experiments under specific experimental conditions described hereafter are performed.

P is freezed to $\{p, q, r\}$ considering the diversity of beliefs and computational cost. In addition, the topic ϕ is always $\{p\}$ to allow us to interpret the results easily by assigning the meaning of each atom. Moreover, $|A|$ and $|N^0|$ are freezed to 100 and 400 to follow the same conditions as those of [55]. In addition, $p_{\text{active}} = p_{\text{rewire}} = 0.5$ and $\mu = 0.5$ to perform experiments under the same conditions as the experiments used to generate Figure 3 in [55]. The number of interactions is always $T = 3000$.

α, the parameter in OF (Section 4.3.2), is set to 0.5 so that the effect of an agent's beliefs on its opinions is the same as the effect of other agents' opinions on it. Finally, the values V presented in Table 4.1 are always used. For the sake of simplicity, it is assumed that V is shared among the agents and does not change over time.

The initial state is generated by drawing randomly $|N^0|$ edges among $|A|$ agents, so that each agent has at least one neighbor. Agents are assigned random beliefs and opinions. To control the randomness of both the initial network and the propagation process, 20 different seeds are used. Before starting interacting with each other, all agents perform $BA \circ OF$ until their opinions and beliefs become stable, i.e., each agent's beliefs do not change and the difference between opinions before and after the operation is less than or equal to 10^{-5}. Hence, agents are in a coherent state at the beginning.

δ is taken from the set $\{1, 2, \ldots, |\mathcal{M}(\top)|\} = \{1, 2, \ldots, 8\}$ (since $|P| = 3$). ε is taken from the set $\{0.05, 0.1, \ldots, 0.5\}$ so that the experiments correspond to those performed in [55]. Table 5.1 shows how each controlled parameter is set in each setting.

Hence, there are

$$\underbrace{10}_{\varepsilon} \times \underbrace{20}_{\text{seed}} = 200$$

Table 5.1: Configuration of the parameters.

Setting	protocol	ε	δ	$\prec_a^t$
S0	(odonly)	$\{0.05, 0.1, \ldots, 0.5\}$	8	(oponly)
S1	(bronly)	1	$\{1, 2, \ldots, 7\}$	(belonly)
S2	(both)	$\{0.05, 0.1, \ldots, 0.5\}$	$\{1, 2, \ldots, 7\}$	(opbel)
S3	(both)	$\{0.05, 0.1, \ldots, 0.5\}$	$\{1, 2, \ldots, 7\}$	(belop)
S4	(coherence)	$\{0.05, 0.1, \ldots, 0.5\}$	$\{1, 2, \ldots, 7\}$	(opbel)
S5	(coherence)	$\{0.05, 0.1, \ldots, 0.5\}$	$\{1, 2, \ldots, 7\}$	(belop)

experiments in S0,

$$\underbrace{7}_{\delta} \times \underbrace{20}_{\text{seed}} = 140$$

experiments in S1, and

$$\underbrace{10}_{\varepsilon} \times \underbrace{7}_{\delta} \times \underbrace{20}_{\text{seed}} = 1400$$

experiments in other settings.

5.4 Hypotheses revisited

With the measures and settings defined above, the hypotheses stated in Section 5.1 can be tested as follows:

Test T0 for H0 $eo^T_{\mathrm{S0}} > 0$.

Test T1 for H1 $eb^T_{\mathrm{S1}} > 0$.

Test T2 for H2 $eo^T_{\mathrm{S2}} > eo^T_{\mathrm{S0}}$ and $eb^T_{\mathrm{S3}} > eb^T_{\mathrm{S1}}$.

Test T3 for H3 $eo^T_{\mathrm{S4}} > eo^T_{\mathrm{S2}}$ and $eb^T_{\mathrm{S5}} > eb^T_{\mathrm{S3}}$.

Here, eo^T_X and eb^T_X mean the measure eo^T and eb^T obtained from the experiment with setting X. As a complement, the effect of changing the binary relation $\prec_a^t$ over A is investigated. This can be achieved by comparing the results from S2 and S3, and S4 and S5.

Hence, the correspondence between hypotheses **H2** and hypotheses **H3** and experiments used to test them is visualized as Figure 5.1.

5.5 Implementation of the simulator

A software environment, called *SOBA* [11], is implemented to perform the experiments. In the simulator, each opinion is represented as a number and each belief is represented as the string that corresponds to the set of its models, not its syntax. For example, the beliefs $B = \{\neg p \wedge q\}$ with $P = \{p, q, r\}$ is represented as `00001100`, where the i-th character from the left shows whether

Fig. 5.1: The correspondence between hypotheses and experiments used to test them. 'CO' refers complementary analysis.

$$\begin{tikzcd}
S0 \arrow[r, leftrightarrow, "H2"] & S2 \arrow[r, leftrightarrow, "H3"] \arrow[d, leftrightarrow, dashed, "CO"'] & S4 \arrow[d, leftrightarrow, dashed, "CO"'] \\
S1 \arrow[r, leftrightarrow, "H2"'] & S3 \arrow[r, leftrightarrow, "H3"'] & S5
\end{tikzcd}$$

m_i in Table 4.1 satisfies the beliefs or not (`1` if $m_i \models B$ and `0` otherwise). This is a simple and direct representation of beliefs in FDNF presented in Section 3.1. Moreover, this allows us to reduce the duration of simulations because this approach can avoid converting a concrete syntax of beliefs to their models and vice versa.

5.6 Conclusion

In this chapter, (1) the hypotheses (Section 5.1), (2) the measures (eo^t and eb^t, Section 5.2), (3) how to set parameters (Section 5.3), and (4) how to test each hypothesis using the measures (Section 5.4) are introduced. Moreover, the implementation of the *SOBA* simulator [11] is also described (Section 5.5). Hence, what is required to perform experiments is described so far.

In the next chapter, the results from the experiments are presented and whether the hypotheses are supported is discussed.

Chapter 6

Results and discussions

In this chapter, the results from the six experiments (Table 5.1) are presented. The results corresponding to the hypotheses **H0**, **H1**, **H2**, and **H3** are presented in Section 6.1, Section 6.2, Section 6.3, and Section 6.4, respectively. Finally, the obtained results are compared with other results reported in [3] briefly in Section 6.5 before concluding (Section 6.6).

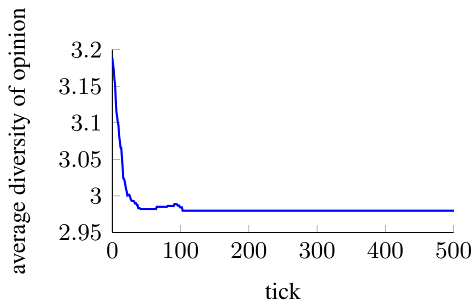
This chapter contains several plots from the experiments. The colors of the nodes in the visualizations correspond to opinions or beliefs (depending on the plots). The same color shows the same opinions or beliefs. When the nodes are colored with respect to opinions, the colors blue, purple, and red correspond to the opinions 0, 0.5, and 1, respectively. Moreover, the directed edge $\langle a, a' \rangle$ means that $a' \in N_a$.

6.1 Hypothesis H0

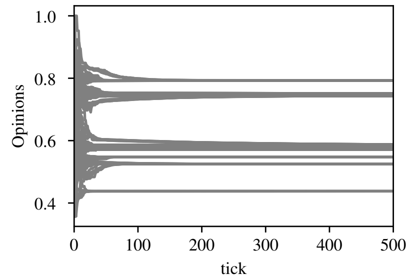
In each simulation with S0, $eo^T > 0$, i.e., at least one echo chamber is created. Thus, the results support Hypothesis **H0**. This result coincides with what is reported in the field of opinion dynamics and in particular in [55]. However, there are no experiments such that agents form single echo chambers and other components are singletons. The plots corresponding to Figure 3 in [55] are shown in Figure 6.1. The same tendency as it can be observed.

Figure 6.2 (left plot, solid green curve) shows the decrease of echo chambers with the increase of tolerance (higher ε). Hence, although a stronger measure is applied, results confirm those of [55].

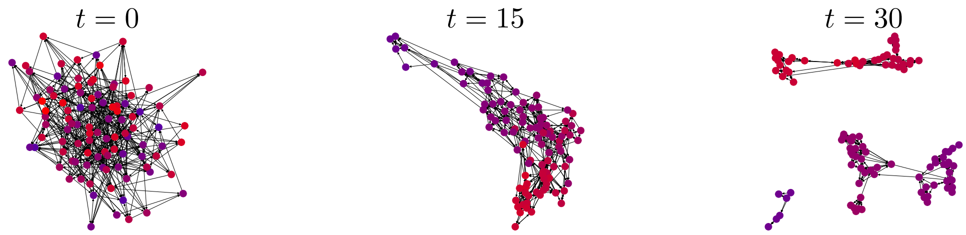
Figure 6.3(a) (the initial network) and Figure 6.3(c) (the final network) shows what happens on a specific case. In the initial network, there are one large component and four singletons. After agents' interactions, there are 7 weakly connected components (WCCs) and each WCC contains some components. Note that different echo chambers can share the same (or quite similar) opinions. For example, agents in the two echo chambers at the bottom of the plot of the final network share the opinion 0.4378.



(a) Average message diversity measured by Shannon entropy.

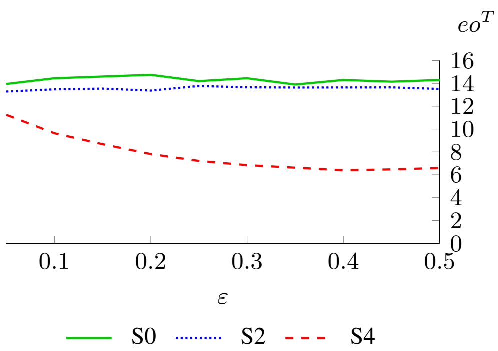


(b) Evolution of agents' opinions

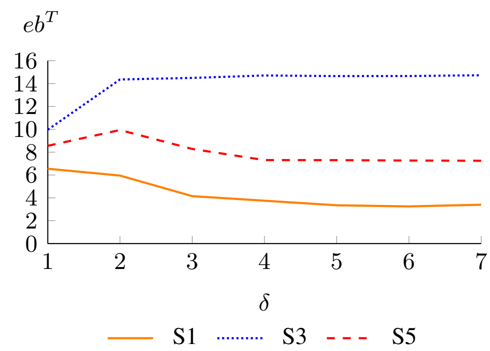


(c) Evolution of the network colored with respect to agents' opinions

Fig. 6.1: Evolution of agents' opinions and the network. These plots correspond to Figure 3 in [55].

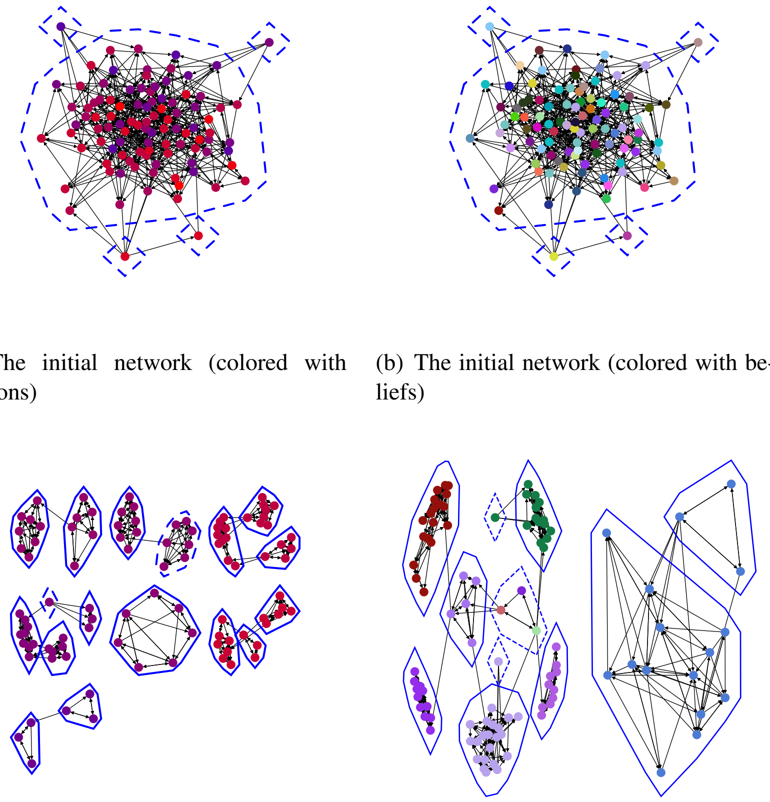


(a) Effect of ε on eo^T. More tolerant agents and connecting opinions and beliefs lead to less echo chambers.



(b) Effect of δ on eb^T. More tolerant agents and connecting opinions and beliefs lead to less echo chambers. However, synchronizing opinions and beliefs makes less echo chambers.

Fig. 6.2: Effect of thresholds (ε and δ) on the number of echo chambers.



(a) The initial network (colored with opinions)

(b) The initial network (colored with beliefs)

(c) Final network from setting S0 (colored with opinions)

(d) Final network from setting S1 (colored with beliefs)

Fig. 6.3: Networks at T from experiments in settings S0 and S1, and using $\varepsilon = 0.2$, $\delta = 2$, and the same seed.

6.2 Hypothesis H1

In each simulation with S1, $eb^T > 0$. These results show that belief echo chambers are observed. Thus, Hypothesis **H1** is supported.

Similarly as for opinions, Figure 6.2 (right plot, solid orange curve) shows the decrease of echo chambers with the increase of tolerance (higher δ).

Figure 6.3(b) (the initial network) and Figure 6.3(d) (the final network) shows the network scattered into less echo chambers compared to the setting S0 (Figure 6.3(c)). However, beliefs are more diverse in the society: there are seven different beliefs which are shared in beliefs echo chambers, while there are (roughly) only three different opinions in S0.

Table 6.1: Results of the pairwise comparison on the measure eo^T between S0 and S2.

condition	#simulations
$eo^T_{\text{S0}} > eo^T_{\text{S2}}$	684 (48.9%)
$eo^T_{\text{S0}} = eo^T_{\text{S2}}$	385 (27.5%)
$eo^T_{\text{S0}} < eo^T_{\text{S2}}$	331 (23.6%)

Table 6.2: Results of the pairwise comparison on the measure eb^T between S1 and S3.

condition	#simulations
$eb^T_{\text{S1}} > eb^T_{\text{S3}}$	9 (0.6%)
$eb^T_{\text{S1}} = eb^T_{\text{S3}}$	5 (0.4%)
$eb^T_{\text{S1}} < eb^T_{\text{S3}}$	1386 (99.0%)

6.3 Hypothesis H2

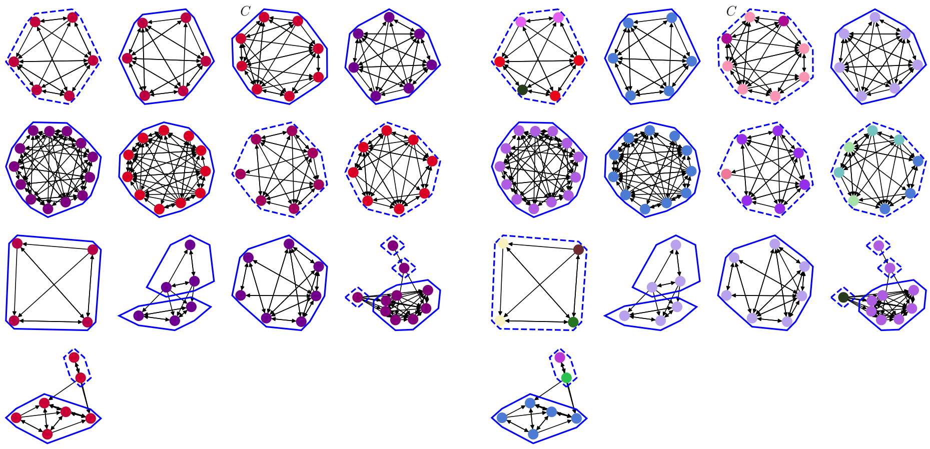
To investigate the effect of evolving opinions and beliefs together without synchronizing them on opinions (resp. beliefs) echo chambers, the results from S0 and S2 (resp. from S1 and S3) are compared.

The effect on opinions echo chambers Table 6.1 shows the results of pairwise comparisons between eo^T_{S0} and eo^T_{S2}. It shows that there are no direct relations between them. Moreover, in $(684+385=)1069$ simulations out of 1400, eo^T does not increase by evolving them together. Figure 6.2(a) shows that on average, eo^T_{S2} (in blue dotted line) is smaller than eo^T_{S0} (in green solid line). This means that evolving opinions and beliefs can reduce the number of opinions echo chambers. Hence, these results do not support the hypothesis.

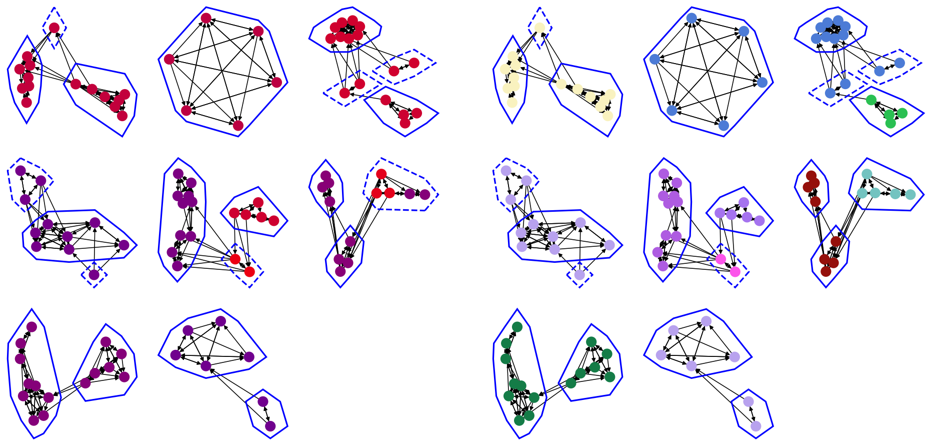
The effect on beliefs echo chambers Table 6.2 shows the results of pairwise comparisons between eb^T_{S1} and eb^T_{S3}. It shows that in all experiments except for $(9+5=)14$ out of 1400, evolving them together creates more beliefs echo chambers. Moreover, Figure 6.2(b) shows on average, eb^T_{S3} (in blue dotted line) is larger than eb^T_{S1} (orange solid line). This means that evolving opinions and beliefs at the same time makes more beliefs echo chambers. These results support the hypothesis.

Figure 6.4 shows the final networks in setting S2 and S3. It shows that opinions echo chambers and beliefs echo chambers do not necessarily coincide. For example, the component C in Figure 6.4(a) and Figure 6.4(b) is an opinions echo chamber while is not a beliefs echo chamber.

Complement analysis As a complement, the effect of changing the binary relation over A (i.e., $\prec^t_a$) is tested. From the results of the pairwise comparisons (Table 6.3 and Table 6.4), there are no direct relations between them. Moreover, from the statistical tests (the Wilcoxon signed-rank test with the significance threshold 0.01; the obtained data is not normal in general), the distribution of eo^T and eb^T can be different if ε or δ are small. For example, if $\delta \in \{1, 2\}$, the distribution of eb^T differs depending on $\prec^t_a$ regardless of ε. However, the obtained data does not show clear relations between them.



(a) Setting S2 (colored with opinions) (b) Setting S2 (colored with beliefs)



(c) Setting S3 (colored with opinions) (d) Setting S3 (colored with beliefs)

Fig. 6.4: Networks at T from experiments in settings S2 and S3, and using $\varepsilon = 0.2$, $\delta = 2$, and the same seed as Figure 6.3.

Table 6.3: Results of the pairwise comparison on the measure eo^T between S2 and S3.

condition	#simulations
$eo^T_{\text{S2}} > eo^T_{\text{S3}}$	471 (33.6%)
$eo^T_{\text{S2}} = eo^T_{\text{S3}}$	274 (19.6%)
$eo^T_{\text{S2}} < eo^T_{\text{S3}}$	655 (46.8%)

Table 6.4: Results of the pairwise comparison on the measure eb^T between S2 and S3.

condition	#simulations
$eb^T_{\text{S2}} > eb^T_{\text{S3}}$	272 (19.4%)
$eb^T_{\text{S2}} = eb^T_{\text{S3}}$	225 (16.1%)
$eb^T_{\text{S2}} < eb^T_{\text{S3}}$	903 (64.5%)

Table 6.5: Results of the pairwise comparison on the measure eo^T between S2 and S4.

condition	#simulations
$eo^T_{\mathrm{S2}} > eo^T_{\mathrm{S4}}$	1213 (86.6%)
$eo^T_{\mathrm{S2}} = eo^T_{\mathrm{S4}}$	80 (5.7%)
$eo^T_{\mathrm{S2}} < eo^T_{\mathrm{S4}}$	107 (7.6%)

Table 6.6: Results of the pairwise comparison on the measure eo^T between S3 and S5.

condition	#simulations
$eb^T_{\mathrm{S3}} > eb^T_{\mathrm{S5}}$	1274 (91.0%)
$eb^T_{\mathrm{S3}} = eb^T_{\mathrm{S5}}$	61 (4.4%)
$eb^T_{\mathrm{S3}} < eb^T_{\mathrm{S5}}$	65 (4.6%)

6.4 Hypothesis H3

Finally, whether the hypothesis **H3** is supported is tested.

First, the effect of synchronizing opinions and beliefs on opinions echo chambers is presented. Table 6.5 shows the results of the pairwise comparisons of eo^T_{S2} and eo^T_{S4}. It shows that in most of the cases (86.6%, 1213 runs out of 1400), synchronizing opinions and beliefs makes less opinions echo chambers. From the Figure 6.2(a), on average, such synchronization always makes less echo chambers (eo^T_{S2} and eo^T_{S4} are shown in blue dotted line and red dashed line, respectively).

The same tendency can be observed for beliefs echo chambers, from Table 6.6 and Figure 6.2(b) (eb^T_{S3} and eb^T_{S5} are shown in blue dotted line and red dashed line, respectively), such synchronization creates less beliefs echo chambers.

These results do not support the hypothesis: synchronizing them allows agents to form less opinions echo chambers and beliefs ones.

It is worth noting that on average, $eo^T_{\mathrm{S0}} > eo^T_{\mathrm{S4}}$ regardless of ε in Figure 6.2(a). This tendency can be observed in most of the cases if the two values are compared pairwise. This means that exchanging opinions and beliefs and keeping coherence between them makes fewer opinions echo chambers than just exchanging opinions following existing opinion dynamics models.

Figure 6.5 shows the different effects of S4 and S5: agents form less echo chambers than before (Figure 6.4). Note that opinions echo chambers and beliefs echo chambers do not necessarily coincide: in some cases, agents in the same opinions echo chambers can have different opinions. In fact, in 574 in S4 and 377 in S5 simulations out of 1400, $eo^T \neq eb^T$. Note that even in these settings with the synchronizations, opinions echo chambers and beliefs echo chambers do not necessarily coincide. Hence, the proposed synchronization algorithms are not the best one: considering AGM-style postulates for such algorithms would help to design alternative ones.

As a complement, eo^T and eb^T obtained from the settings S4 and S5 are compared. Table 6.7 and Table 6.8 show the results of pairwise comparisons. For eo^T, there is no direct relations between them. For eb^T, in 960 simulations out of 1400, following Equation (belop) makes more beliefs echo chambers. Strong relations between them are not observed in the statistical tests: in some cases there are significant differences with respect to the significant threshold 0.01, while in some cases the differences are not significant.

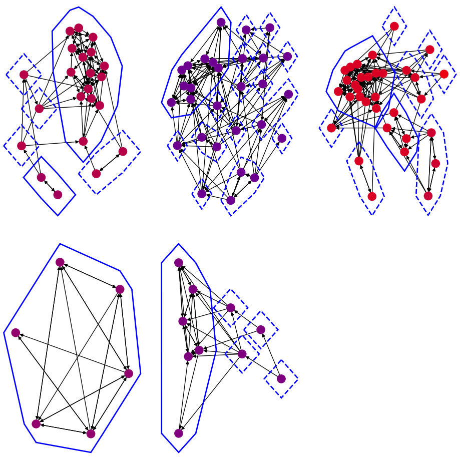
Note that these analyses so far are conducted by aggregating results by computing the mean. In fact, as the results of the pairwise comparisons show, the results are not conclusive. This makes it difficult to analyze how changing initial conditions affects the echo chambers. It is still required to investigate how the final states including the number of echo chambers can be characterized by the initial states to have more solid conclusions.

Table 6.7: Results of the pairwise comparison on the measure eo^T between S4 and S5.

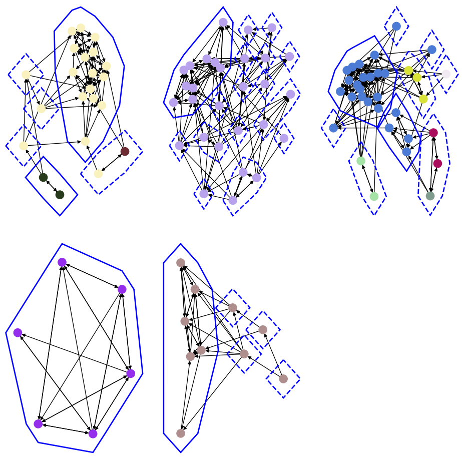
condition	#simulations
$eo^T_{\text{S4}} > eo^T_{\text{S5}}$	496 (35.4%)
$eo^T_{\text{S4}} = eo^T_{\text{S5}}$	218 (15.6%)
$eo^T_{\text{S4}} < eo^T_{\text{S5}}$	585 (49.0%)

Table 6.8: Results of the pairwise comparison on the measure eb^T between S4 and S5.

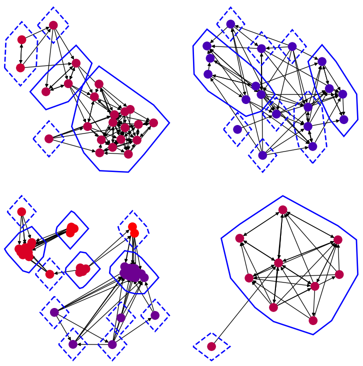
condition	#simulations
$eb^T_{\text{S4}} > eb^T_{\text{S5}}$	261 (18.6%)
$eb^T_{\text{S4}} = eb^T_{\text{S5}}$	179 (12.8%)
$eb^T_{\text{S4}} < eb^T_{\text{S5}}$	960 (68.6%)



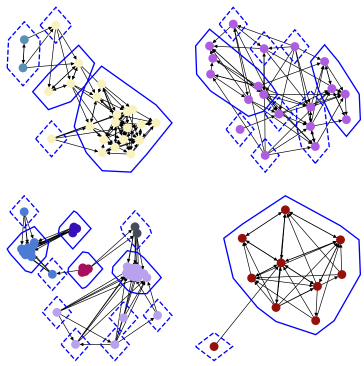
(a) Setting S4 (colored with opinions)



(b) Setting S4 (colored with beliefs)

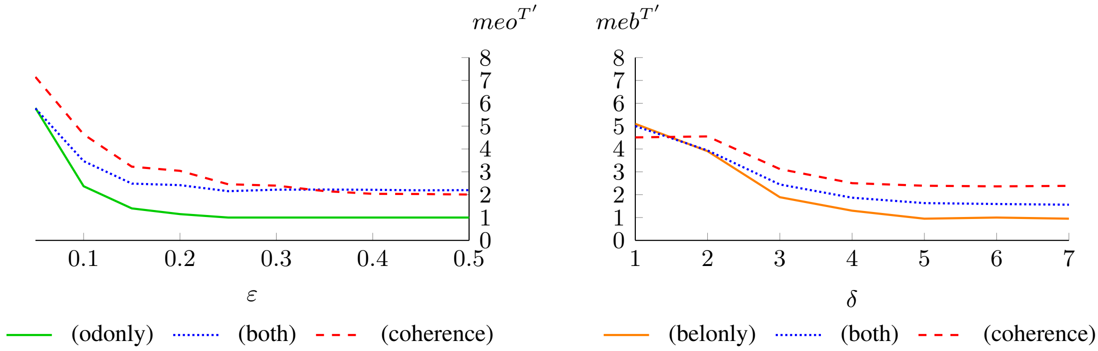


(c) Setting S5 (colored with opinions)



(d) Setting S5 (colored with beliefs)

Fig. 6.5: Networks at T from experiments in settings S4 and S5, and using $\varepsilon = 0.2$, $\delta = 2$, and the same seed as Figure 6.3.



(a) Effect of ε on $meo^{T'}$.

(b) Effect of δ on $meb^{T'}$.

Fig. 6.6: Effect of thresholds (ε and δ) on the number of echo chambers. These plots are obtained by reanalyzing the results obtaied in [3].

6.5 Remark: comparisons with other results

The reported results so far do not coincide with other results reported in [3]. In [3], different rewiring method is used: unfollwed agent a' is one of the agents with different opinions or beliefs (i.e., $a' \in N_a^{t-1} \setminus C_a^{t-1}$) and followed agent a'' is one of the non-neighbors (i.e., $a'' \in A \setminus (\{a\} \cup N_a^{t-1})$). If such agents exist, in other word, if $N_a^{t-1} \setminus C_a^{t-1}$ and $A \setminus (\{a\} \cup N_a^{t-1})$ are not empty, N_a^t is obtained by excluding a' from N_a^{t-1} and adding a'' to it, i.e., $N_a^t = (N_a^{t-1} \setminus \{a'\}) \cup \{a''\}$. Otherwise, the set of neighbors is kept as is ($N_a^t = N_a^{t-1}$). Other components of the proposed model and the measures are the same as what is presented in this thesis. Moreover, the values of frozen and controlled parameters are the same.

In these configurations, it is observed that keeping the coherence between opinions and beliefs could create more echo chambers than without keeping it or without the dual propagation of opinions and beliefs. In [3], there results are obtained using the same measures $eo^{T'}$ and $eb^{T'}$ as what are introduced in Section 5.2 where $T' = 5000$. However, they are based on strongly connected components that could be affected by the tiny modifications to the networks, while in some cases the networks agents form never stabilizes because it is possible that for some agents, there could be less concordant agents than their neighbors. Such constantly changing networks can change the conclusions.

To reduce the effect of such infinity rewiring, the results obtained in [3] are reanalyzed using the average number of opinions/beliefs echo chambers (meo^t and meb^t, respectively) defined as:

$$meo^t = \frac{1}{W} \sum_{t'=t-W+1}^{t} eo^{t'}, \text{ and } meb^t = \frac{1}{W} \sum_{t'=t-W+1}^{t} eb^{t'},$$

where $W = 25$ is the width of time window during which eo and eb are computed. The effects of ε and δ on $meo^{T'}$ and $meb^{T'}$, respectively, are displayed in Figure 6.6. The figure shows the same conclusions as [3]. Hence, the results reported in [3], i.e., the results obtained with the random rewiring, contradict with what is reported in this thesis.

These results suggest that how the network is modified also affects the formation of echo chambers. The difference in the two rewiring methods is that how the candidates of unfollowed and followed agents are selected: in [3] they are neither a itself nor a's neighbors, while in this thesis they are the minimal one among non-neighbors with respect to $\prec_a^t$. Hence, this result suggests that depending on the characteristics of the rewiring algorithms, agents can form less echo chambers, which contradicts with what is reported in existing work (e.g., [55]). Investigating under which condition(s) such results are obtained is left for future work.

6.6 Conclusion

In this chapter, results from the six experiments are presented. They can be summarized as follows:

Hypothesis H0 is supported;

Hypothesis H1 is supported;

Hypothesis H2 is supported on average for beliefs echo chambers and not supported for opinions echo chambers;

Hypothesis H3 is not supported.

Moreover, changing the definition of $\prec_a^t$ (Equation opbel or Equation belop) does not affect the results in general, while in some cases (e.g., eb_{S2}^T and eb_{S3}^T if $\delta \in \{1, 2\}$, see Section 6.3) it affects the final states. Finally, it is shown that depending on the rewiring algorithm, the number of echo chambers can increase.

In the next chapter, the conclusions and future work are presented.

Chapter 7

Conclusion and Future Work

7.1 Conclusion

In social networks, beliefs are propagated as well as opinions typically to justify them. However, existing models do not take such beliefs into account. In order to test their influence on echo chambers, a model in which agents exchange and update their opinions (resp. beliefs) using opinions (resp. beliefs) from others by extending existing models [55, 57] is introduced. Moreover, agents use their (cultural) values for maintaining the coherence between opinions and beliefs.

In addition, new measures (eo^t and eb^t) are introduced to capture the definition of echo chambers based on their definition [45]. They enable us to determine how many echo chambers are in the network directly.

Experimental results show that agents always form echo chambers if they exchange and update opinions only. The same can be observed for beliefs. Moreover, they also show that joining opinions and beliefs updates generates less opinions echo chambers and more beliefs echo chambers on average. In addition, synchronizing opinions and beliefs creates less echo chambers.

Hence, exchanging beliefs and opinions and keeping the coherence between them help to reduce echo chambers from the viewpoint of opinions. However, from the standpoint of beliefs, only synchronizing them allows this.

The raw results from the experiments and notebooks used to analyze them in settings S0, S1, S2, S3, S4, and S5 can be retrieved from [4], [5], [6], [7], [8], and [9], respectively. The notebook corresponding to the additional analysis is available at [10].

7.2 Future Work

In this thesis, only whether connecting opinions and beliefs contributes to create more echo chambers is investigated. Other effects are not studied. For example, it can be asked whether agents can create echo chambers more resilient to tiny modifications for edges (e.g., removing or adding edges to the network agents form) or cognitive states (e.g., changing agents' cognitive states independently from the workflow). This can help us to find a way to 'break' echo chambers, i.e., make some of the components do not satisfy the three features.

In addition, only the social aspect of opinion and belief propagation is considered: whether beliefs are true with respect to the environment is not modeled. This does not allow us to address the problem of misinformation in social networks these days, because agents cannot determine whether the beliefs they obtain from their neighbors are true in the proposed model: to achieve this, it is required to compare such beliefs with what is true in the environment. Moreover, this prevents us from investigating how opinions associated to beliefs accelerate the spread of misinformation. Hence, it is worth asking whether keeping the coherence between opinions and beliefs contributes to spread misinformation. This can be achieved by allowing some agents to observe (possibly some of) the truth from the environment and revise their beliefs.

Acknowledgements

I would like to express my sincere gratitude to my supervisors, Associate Professor Koji Hasebe (University of Tsukuba, UT) and Dr. Jérôme Euzenat (Inria), for their enthusiastic discussions. Moreover, I am grateful to members of the Multi-Agent Systems Laboratory at UT and mOeX team at Inria for their advice. Finally, I would like to thank my family for their support.

I have partially supported by the École Française de l'Intelligence Artificielle - MIAI Institute (Efelia, ANR-22-CMAS-0005), by the gratification for master internship by Inria, and by the Echo action of the NTU-UT-UGA trilateral center.

This work is supported by the French National Research Agency n°ANR-22-CMAS-0005 France 2030

Bibliography

My Publications

[1] Hiro Kataoka, Jérôme Euzenat, and Koji Hasebe. “Exchanging and updating opinions and beliefs reinforces echo chambers”. In: *Proc. 39th annual conference of the Japanese society for artificial intelligence (JSAI), Osaka (JP)*. 2025, 8 pages. DOI: 10.11517/pjsai.JSAI2025.0_3K1IS302.

[2] Hiro Kataoka, Jérôme Euzenat, and Koji Hasebe. “How can beliefs alter opinions? Joint opinion and belief evolution”. In: *Proc. 18th International Conference on Agents and Artificial Intelligence (ICAART), Marbella (ES)*. 2026, pp. 15–26. DOI: 10.5220/0014290200004052.

[3] Hiro Kataoka, Jérôme Euzenat, and Koji Hasebe. “Coherent belief and opinion propagation produces more echo chambers”. In: *Proceedings of the 25th International Conference on Autonomous Agents and Multiagent Systems (AAMAS), Paphos (CY)*. to appear (accepted as full paper). 2026, 9 pages.

Repository with the Analyzed Data

[4] Hiro Kataoka. *20250527-BROD Experiment description*. https://sake.re/20250527-BROD/. 2025.

[5] Hiro Kataoka. *20250528-BROD Experiment description*. https://sake.re/20250528-BROD/. 2025.

[6] Hiro Kataoka. *20250529-BROD Experiment description*. https://sake.re/20250529-BROD/. 2025.

[7] Hiro Kataoka. *20250530-BROD Experiment description*. https://sake.re/20250530-BROD/. 2025.

[8] Hiro Kataoka. *20250531-BROD Experiment description*. https://sake.re/20250531-BROD/. 2025.

[9] Hiro Kataoka. *20250532-BROD Experiment description*. https://sake.re/20250532-BROD/. 2025.

[10] Hiro Kataoka. *20251230-BROD Experiment description*. https://sake.re/20251230-BROD/. 2025.

Software

[11] *SOBA*. https://github.com/tsukuba-mas/soba. 2025.

Other References

[12] Carlos Alchourrón, Peter Gärdenfors, and David Makinson. “On the logic of theory change: partial meet contraction and revision functions”. In: *Journal of symbolic logic* 50.2 (1985), pp. 510–530.

[13] Alexandru Baltag and Sonja Smets. “A qualitative theory of dynamic interactive belief revision”. In: *Logic and the Foundations of Game and Decision Theory*. Ed. by Giacomo Bonanno, Wiebe van der Hoek, and Michael Wooldridge. Vol. 3. Amsterdam university press, Amsterdam (NL), 2008, pp. 22–58. DOI: 10.1007/978-3-319-20451-2_39.

[14] Carmela Bernardo et al. “Bounded confidence opinion dynamics: A survey”. In: *Automatica* 159 (2024), p. 111302. DOI: 10.1016/j.automatica.2023.111302.

[15] Anastasia Bizyaeva, Alessio Franci, and Naomi Ehrich Leonard. “Nonlinear Opinion Dynamics With Tunable Sensitivity”. In: *IEEE Transactions on Automatic Control* 68.3 (2023), pp. 1415–1430. DOI: 10.1109/TAC.2022.3159527.

[16] Richard Booth and Thomas Meyer. “Admissible and Restrained Revision”. In: *Journal of Artificial Intelligence Research* 26 (June 2006), pp. 127–151. DOI: 10.1613/jair.1874.

[17] Sirin Botan, Umberto Grandi, and Laurent Perrussel. “Multi-Issue Opinion Diffusion under Constraints”. In: *Proceedings of the 18th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*. Montréal (CA), May 2019, pp. 828–836. URL: https://www.ifaamas.org/Proceedings/aamas2019/pdfs/p828.pdf.

[18] Yasser Bourahla, Manuel Atencia, and Jérôme Euzenat. “Knowledge Improvement and Diversity under Interaction-Driven Adaptation of Learned Ontologies”. In: *Proceedings of the 20th international conference on Autonomous Agents and Multi-Agent Systems (AAMAS), Online*. 2021, pp. 242–250. URL: http://www.ifaamas.org/Proceedings/aamas2021/pdfs/p242.pdf.

[19] Yasser Bourahla, Manuel Atencia, and Jérôme Euzenat. “Knowledge Transmission and Improvement Across Generations do Not Need Strong Selection”. In: *Proceedings of the 21st international conference on Autonomous Agents and Multi-Agent Systems (AAMAS), Online*. 2022, pp. 163–171. URL: http://www.ifaamas.org/Proceedings/aamas2022/pdfs/p163.pdf.

[20] Erica Cau et al. “Selective agreement, not sycophancy: investigating opinion dynamics in LLM interactions”. In: *EPJ Data Science* 14.1 (Aug. 2025). ISSN: 2193-1127. DOI: 10.1140/epjds/s13688-025-00579-1.

[21] Yun-Shiuan Chuang et al. *Simulating Opinion Dynamics with Networks of LLM-based Agents*. 2024. arXiv: 2311.09618 [physics.soc-ph]. URL: https://arxiv.org/abs/2311.09618.

[22] Federico Cinus et al. “The Effect of People Recommenders on Echo Chambers and Polarization”. In: *Proceedings of the International AAAI Conference on Web and Social Media* 16 (May 2022), pp. 90–101. DOI: 10.1609/icwsm.v16i1.19275.

[23] Aaron Clauset, M. E. J. Newman, and Cristopher Moore. “Finding community structure in very large networks”. In: *Phys. Rev. E* 70 (6 Dec. 2004), p. 066111. DOI: 10.1103/PhysRevE.70.066111.

[24] Peter Clifford and Aidan Sudbury. “A model for spatial conflict”. In: *Biometrika* 60.3 (Dec. 1973), pp. 581–588. DOI: 10.1093/biomet/60.3.581.

[25] David Coady. “Stop Talking about Echo Chambers and Filter Bubbles”. In: *Educational Theory* 74.1 (2024), pp. 92–107. DOI: 10.1111/edth.12620.

[26] Guillaume Deffuant et al. “Mixing beliefs among interacting agents”. In: *Advances in complex systems* 3.1-4 (2000), pp. 87–98. DOI: 10.1142/S0219525900000078.

[27] Morris DeGroot. “Reaching a consensus”. In: *Journal of the American Statistical Association* 69.345 (1974), pp. 118–121. DOI: 10.1080/01621459.1974.10480137.

[28] Yucheng Dong et al. “A survey on the fusion process in opinion dynamics”. In: *Information Fusion* 43 (2018), pp. 57–65. DOI: 10.1016/j.inffus.2017.11.009.

[29] Eduardo Fermé and Sven Ove Hansson. *Belief change: introduction and overview*. Springer, Cham (CH), 2018.

[30] Noah Friedkin et al. “Network science on belief system dynamics under logic constraints”. In: *Science* 354.6310 (2016), pp. 321–326. DOI: 10.1126/science.aag2624.

[31] Fabio R. Gallo et al. “Local Belief Dynamics in Network Knowledge Bases”. In: *ACM Trans. Comput. Logic* 23.1 (Oct. 2021). DOI: 10.1145/3477394.

[32] Kiran Garimella et al. “Quantifying Controversy in Social Media”. In: *Proceedings of the Ninth ACM International Conference on Web Search and Data Mining*. San Francisco, California, USA, 2016, pp. 33–42. DOI: 10.1145/2835776.2835792.

[33] Zhen Guo, Jaber Valinejad, and Jin-Hee Cho. “Effect of Disinformation Propagation on Opinion Dynamics: A Game Theoretic Approach”. In: *IEEE Transactions on Network Science and Engineering* 9.5 (2022), pp. 3775–3790. DOI: 10.1109/TNSE.2022.3181130.

[34] David Hartmann et al. “A systematic review of echo chamber research: comparative analysis of conceptualizations, operationalizations, and varying outcomes”. In: *Journal of Computational Social Science* 8.2 (Apr. 2025). DOI: 10.1007/s42001-025-00381-z.

[35] Hossein Hassani et al. “Classical dynamic consensus and opinion dynamics models: A survey of recent trends and methodologies”. In: *Information fusion* 88 (2022), pp. 22–40. DOI: 10.1016/j.inffus.2022.07.003.

[36] Rainer Hegselmann and Ulrich Krause. “Opinion Dynamics and Bounded Confidence: Models, Analysis and Simulation”. In: *Journal of artificial societies and social simulation* 5.3 (2002). URL: https://www.jasss.org/5/3/2.html.

[37] Richard A. Holley and Thomas M. Liggett. “Ergodic Theorems for Weakly Interacting Infinite Systems and the Voter Model”. In: *The annals of probability* 3.4 (1975), pp. 643–663. DOI: 10.1214/aop/1176996306.

[38] Ruben Interian et al. "Network polarization, filter bubbles, and echo chambers: an annotated review of measures and reduction methods". In: *International Transactions in Operational Research* 30.6 (2023), pp. 3122–3158. DOI: 10.1111/itor.13224.

[39] Di Jin et al. "A Survey of Community Detection Approaches: From Statistical Modeling to Deep Learning". In: *IEEE Transactions on Knowledge and Data Engineering* 35.2 (2023), pp. 1149–1170. DOI: 10.1109/TKDE.2021.3104155.

[40] Audun Jøsang. *Subjective Logic*. Springer International Publishing, 2016. DOI: 10.1007/978-3-319-42337-1.

[41] Sébastien Konieczny, Mattia Medina Grespan, and Ramón Pino Pérez. "Taxonomy of Improvement Operators and the Problem of Minimal Change". In: *Proceedings of the Twelfth International Conference on Principles of Knowledge Representation and Reasoning (KR'10)*. 2010, pp. 161–170.

[42] Sébastien Konieczny, Jérôme Lang, and Pierre Marquis. "DA^2 merging operators". In: *Artificial Intelligence* 157.1–2 (Aug. 2004), pp. 49–79. DOI: 10.1016/j.artint.2004.04.008.

[43] Sébastien Konieczny and Ramón Pino Pérez. "Merging Information Under Constraints: A Logical Framework". In: *Journal of Logic and Computation* 12.5 (Oct. 2002), pp. 773–808. DOI: 10.1093/logcom/12.5.773. eprint: https://academic.oup.com/logcom/article-pdf/12/5/773/3852854/120773.pdf.

[44] Sébastien Konieczny and Ramón Pino Pérez. "Improvement Operators". In: *Proceedings of the Eleventh International Conference on Principles of Knowledge Representation and Reasoning (KR'08)*. 2008, pp. 177–187.

[45] Amin Mahmoudi, Dariusz Jemielniak, and Leon Ciechanowski. "Echo Chambers in Online Social Networks: A Systematic Literature Review". In: *IEEE Access* 12 (2024), pp. 9594–9620. DOI: 10.1109/ACCESS.2024.3353054.

[46] Rijk Mercuur, Virginia Dignum, and Catholijn Jonker. "The value of values and norms in social simulation". In: *Journal of artificial societies and social simulation* 22.1 (2019). DOI: 10.18564/jasss.3929.

[47] Takuya Nagura and Eizo Akiyama. "Emergence of echo chambers in social media communication involving multiple topics". In: *Information, Communication & Society* (2025). Accepted for publication, pp. 1–22. DOI: 10.1080/1369118X.2025.2456561.

[48] Angelina Nedić, Alex Olshevsky, and César Uribe. "Graph-theoretic analysis of belief system dynamics under logic constraints". In: *Scientific reports* 9 (2019), p. 8843. DOI: 10.1038/s41598-019-45076-4.

[49] M. E. J. Newman. "Modularity and community structure in networks". In: *Proceedings of the National Academy of Sciences* 103.23 (2006), pp. 8577–8582. DOI: 10.1073/pnas.0601602103.

[50] Masaya Ohagi. "Polarization of Autonomous Generative AI Agents Under Echo Chambers". In: *Proceedings of the 14th Workshop on Computational Approaches to Subjectivity, Sentiment, & Social Media Analysis*. Aug. 2024, pp. 112–124. URL: https://aclanthology.org/2024.wassa-1.10.

[51] Kerem Oktar and Tania Lombrozo. "How aggregated opinions shape beliefs". In: *Nature Reviews Psychology* 4.2 (Jan. 2025), pp. 81–95. DOI: 10.1038/s44159-024-00398-7.

[52] Sergey E. Parsegov et al. "Novel Multidimensional Models of Opinion Dynamics in Social Networks". In: *IEEE Transactions on Automatic Control* 62.5 (2017), pp. 2270–2285. DOI: 10.1109/TAC.2016.2613905.

[53] Jinghua Piao et al. *Emergence of human-like polarization among large language model agents*. 2025. arXiv: 2501.05171 [cs.SI]. URL: https://arxiv.org/abs/2501.05171.

[54] Yimeng Qi, Dong Xue, and Fangzhou Liu. "Opinion diversity shaped by higher-order interactions on multi-dimensional issues". In: *Chaos, Solitons & Fractals* 194 (2025), pp. 116188–116195. DOI: 10.1016/j.chaos.2025.116188.

[55] Kazutoshi Sasahara et al. "Social influence and unfollowing accelerate the emergence of echo chambers". In: *Journal of Computational Social Science* 4.1 (2020), pp. 381–402. DOI: 10.1007/s42001-020-00084-7.

[56] Shalom H. Schwartz. "A Theory of Cultural Values and Some Implications for Work". In: *Applied Psychology* 48.1 (Jan. 1999), pp. 23–47. DOI: 10.1111/j.1464-0597.1999.tb00047.x.

[57] Nicolas Schwind et al. "Belief Revision Games". In: *Proceedings of the 29th AAAI Conference, Austin (TX US)*. 2015, pp. 1590–1596. DOI: 10.1609/aaai.v29i1.9415.

[58] Nicolas Schwind et al. "On Opinion Control in Belief Revision Games". In: *Proc. 30th Annual Conference of the Japanese Society for Artificial Intelligence (JSAI), Kita-Kyushu (JP)*. 2016, 1F24–1F24. DOI: 10.11517/pjsai.JSAI2016.0_1F24.

[59] Nicolas Schwind et al. "BeliefFlow: A framework for logic-based belief diffusion via iterated belief change". In: *Proceedings of the 38th AAAI Conference, Vancouver (CA)*. 2024, pp. 10696–10704. DOI: 10.1609/aaai.v38i9.28941.

[60] Jean Springsteen, William Yeoh, and Dino Christenson. "Algorithmic Filtering, Out-Group Stereotype, and Polarization on Social Media". In: *Proceedings of the 23rd International Conference on Autonomous Agents and Multiagent Systems*. Auckland, New Zealand, 2024, pp. 1782–1790. ISBN: 9798400704864.

[61] Michele Starnini et al. *Opinion dynamics: Statistical physics and beyond*. 2025. arXiv: 2507.11521 [physics.soc-ph]. URL: https://arxiv.org/abs/2507.11521.

[62] Yanni Tang, Jiamou Liu, and Wu Chen. "Exchange, adopt, evolve: Modeling the spreading of opinions through cognition and interaction in a social network". In: *Information sciences* 551 (2021), pp. 1–22. DOI: 10.1016/j.ins.2020.11.043.

[63] Hans van Ditmarsch, Wiebe van der Hoek, and Barteld Kooi. *Dynamic Epistemic Logic*. Springer, Heidelberg (DE), 2018.

[64] Chenxi Wang et al. *Decoding Echo Chambers: LLM-Powered Simulations Revealing Polarization in Social Networks*. 2024. arXiv: 2409.19338 [cs.SI]. URL: https://arxiv.org/abs/2409.19338.

[65] Tao Wen et al. "Formulating opinion dynamics from belief formation, diffusion and updating in social network group decision-making: Towards developing a holistic framework". In: *European Journal of Operational Research* 325.3 (2025). DOI: 10.1016/j.ejor.2024.12.015.

[66] Junchi Yao et al. "Social opinions prediction utilizes fusing dynamics equation with LLM-based agents". In: *Scientific Reports* 15.1 (May 2025). DOI: 10.1038/s41598-025-99704-3.