AGENTS IN ECHO CHAMBERS: WHEN WILL THEY CHANGE THEIR MIND?

Bachelor's Project Thesis

Annet Onnes, s2485540, A.T.Onnes@student.rug.nl,
Supervisors: Prof. dr. L.C. Verbrugge & Prof. dr. B.P. Kooi

Abstract: Echo chambers are a politically current and interesting information phenomenon. They are epistemic networks which through belief polarization and limited exposure to information have become impregnable to outside information. In this bachelor's project, I have looked at the resolution of these epistemic structures by developing two multi-agent models in NetLogo: DIALx and DIALx2. I based both on DIAL, an existing model created by Dykstra et al. In order to differentiate between epistemic bubbles and echo chambers I added belief entrenchment to the models. Belief entrenchment is an idea by Baumgaertner. Results showed that the DIALx implementation did not allow resolution of echo chambers or bubbles. The limited movement of the agents restricts their exposure to opposing evidence, thereby upholding the belief entrenchment values. DIALx2 is a response to this behavior. In this implementation, the movement of agents depends on their belief entrenchment, allowing resolution. Both models provide telling results about the resolution and constitution of echo chambers and other epistemic structures. A comparison between DIAL, DIALx and DIALx2 contributes to the field by reviewing existing models and ideas. The succession of DIALx and DIALx2 shows the strong influence of agents' movements within topic space, bringing up new research questions.

Keywords: Echo chambers, belief polarization, opinion dynamics, epistemic bubbles

1 Introduction

Echo chambers are social and epistemic phenomena that are becoming more and more apparent in the current state of affairs in the world. In the comfort zone of being surrounded by people with similar opinions, individuals start to feel stronger for their beliefs and become susceptible to discrediting opposing opinions. These closed environments and the discrediting of opposing opinions prevent valuable and needed discussion from happening, as people are unable to thoroughly consider the arguments of others with contradictory opinions. Echo chambers can arise in different kinds of environments, but an environment currently in the spotlight is the online world. The following tweet was posted by the president of the USA on the 30th of December 2017. Words such as 'dishonest', 'unfair', 'phony' and 'fiction' do not pertain to the contents of the information from external sources, but only the external source themselves. It might therefore be an example of the rejection of what might just be opposing opinions, because they are such.

> *I use Social Media not because I like to, but because it is the only way to fight a VERY dishonest and unfair "press," now often referred to as Fake News Media. Phony and non-existent "sources" are being used more often than ever. Many stories & reports a pure fiction!*

Research has already shown that echo chambers are not only a phenomenon perceived by people but that there is also data driven evidence for the phenomenon, even though it is modest (Flaxman, Goel, and Rao, 2016). Thus to have meaningful discussion, echo chambers should be prevented or resolved. In this bachelor's project I will look at the latter.

Segregation is not a new phenomenon and has been researched thoroughly in social psychology, however the concept *echo chamber* is relatively new. Before introducing echo chambers, I first will introduce epistemic bubbles, as they are more moderate epistemic structures. Nguyen (unpublished) defined epistemic bubbles as *a social epistemic structure which has inadequate [knowledge] coverage through a pro-*

cess of mere exclusion. Since the structures are closed off there is a particular information dynamic happing within them, which can lead to belief polarization (Dykstra, Elsenbroich, Verbrugge, and de Lavalette, 2013; Hansen, Hendricks, and Rendsvig, 2013), where everything within the structure points towards one belief.

The definition of *echo chamber* I adhere to is *an epistemic structure which through belief polarization has become impregnable to information*. This belief polarization implies a strong belief in the core belief set of the bubble by its members. This strong belief is described by Baumgaertner (2014) as high belief entrenchment. Looking at the belief entrenchment, epistemic bubbles can converge into echo chambers, as high belief entrenchment prevents the members to change their beliefs and be persuaded by outside knowledge and arguments. The impregnability is thus caused by the high belief entrenchment.

This bachelor's project is an attempt to simulate the resolution of echo chambers and epistemic bubbles, using agent-based modelling as a theoretical tool to study the opinion dynamics. The final goal is to answer the question:

> How can echo chambers and epistemic bubbles be resolved?

A variety of factors influence opinion dynamics. I will look at the initial belief entrenchment, authorities and spatial distribution, which could impact the process of resolution.

The resolution of echo chambers is a social information process that is relevant for the current political landscape and has been tackled from many different angles. Opinion dynamics and social information processes have been thoroughly studied using many different multi-agent models. By combining and comparing existing models and ideas, this project contributes to the evaluation of models in existing research (Flache, Mäs, Feliciani, Chattoe-Brown, Deffuant, Huet, and Lorenz, 2017). Whereas Baumgaertner (2014) and Dykstra et al. (2013) focus on the formation of echo chambers (and prevention thereof), this paper will look at an attempted depolarization of the group of agents after they have reached a stable segregated state. In this paper two different models are presented, both are based on DIAL (also knowns as DIAL1.0), initially introduced by Dykstra et al. (2013). The first is DIALx, which is extended with belief entrenchment. The second, DIALx2, offers a further extension on both previous models which allows agents to spatially break up echo chambers. Section 2 is required reading to understand Section 3.

2 DIALx

2.1 Methods

There are multiple ways to study social phenomena such as echo chambers. One way is to look at people's behaviour in naturally formed echo chambers (Flaxman et al., 2016). Another method to study phenomena, from the field of psychology, is to recreate them in a controlled environment. Although gathering data is beneficial since it can show patterns in behaviour, it requires artificially recreating an echo chamber and controlling the subject's belief states, which is ethically problematic. Even so, knowledge of such patterns can be used to formalize echo chambers, as well as to better understand echo chambers and their development. Therefore, I use a multi-agent simulation to gather data about behaviour in and around echo chambers. A lot of agent-based research has already been done in the field of opinion dynamics. Since there is no use in reinventing the wheel, I am creating an extension of DIAL 1.0 (Dykstra et al., 2013).

2.1.1 Implementation

DIAL is a similarity biased influence model (Flache et al., 2017) in the form of an agent-based simulation, comprising the following main elements: dialogue, argumentation, game structure, reputation status, social embedding and alignment of opinions (Dykstra et al., 2013). The implementation of a dialogue is as a short exchange of opinions, that may take place after an agent makes an announcement. The use of pay and reward in these dialogues activates a game structure in the interactions. The agent's reputation status is based on how well the agent does in this 'game'. Social embedding is implemented by having DIAL agents move around in a topic space, which gives an indication of the held opinions. This is an implicit parallel to social and cultural influences on agents that change based on their held opinions. The main draw towards DIAL in order to look at echo chambers and epistemic bubbles is the alignment of opinions among agents, but all other elements contribute to a more complete simulation.

In Dykstra's simulations, with certain parameter settings, the emergence of a segregation state can be observed. Dykstra et al. (2013) describe this state as a segregation state, but this state strongly resembles epistemic bubbles and echo chambers. Dykstra's model was already able to represent the epistemic structures, which contributed to the choice to extend DIAL.

Adaptations to DIAL Several changes were made to DIAL in order to make the aforementioned distinction between epistemic bubbles and echo chambers and to study the potential resolution of those epistemic structures. Furthermore, some functionalities of DIAL were removed as they served no purpose in this research, including the possibility of dialogues about several different propositions.

The first addition that was made was a change to the initial set-up, which in DIAL is random. In the initial set-up in of DIALx, the agents are already segregated into two bubbles, the proponents and opponents of one proposition, each bubble positioned on one side of topic space. The evidence value for each group is randomly generated on either side of the neutral evidence value (0.5). The following adaptable parameters were added to determine the initial set-up:

- `init-division-ratio`
- `init-belief-entrenchment`
- `init-density`
- `ratio`

A set-up that abides to these parameters is generated semi-randomly, see Figure 2.1. How exactly this is done will follow in the next section.

The second addition is the implementation of belief entrenchment as described by Baumgaertner (2014), closely interacting with the evidence value each agent holds. Belief entrenchment is required to create a distinction between echo chambers and epistemic bubbles; A high belief entrenchment indicates an echo chamber. Emerging from this is also the agents' rejection of evidence from agents from outside their bubble. Table 2.1 shows the interaction between evidence and belief entrenchment as it was implemented. This is the recalculation of the new evidence value of one agent as its previous evidence value (e_1) is influenced by the evidence value of a second agent (e_2). The belief entrenchment of the first agent (b) determines the extent to which this agent updates its evidence value. The belief entrenchment of agents is not a constant number. It (b) is calculated using Equation 2.1, where c is the current belief entrenchment, d is the agreement factor and e is the adaptability of the agents.

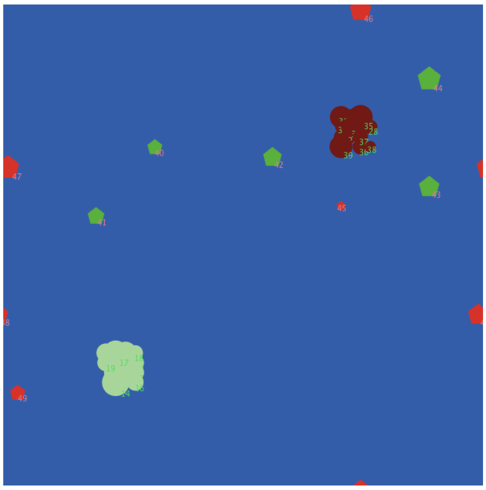


Figure 2.1: Example of initial setup. In the upper right corner there is a bubble of proponents and in the lower left corner there is a bubble of opponents. The pentagon-shaped agents are authorities, randomly dispersed over the topic space.

$e_1 < 0.5$ & $e_2 < 0.5$	$2 * e_1 * e_2$
$e_1 \geq 0.5$ & $e_2 < 0.5$	$e_1 - ((e_1 - e_2) * 0.5 * (1 - b))$
$e_2 \geq 0.5$ & $e_1 < 0.5$	$e_1 - ((e_1 - e_2) * 0.5 * (1 - b))$
$e_2 \geq 0.5$ & $e_1 \geq 0.5$	$2 * e_1 + 2 * e_2 - 2 * e_1 * e_2 - 1$

Table 2.1: Recalculation formulas for evidence, given two evidence values (e_1,e_2) and belief entrenchment (b).

$$b = c + (d * e) \tag{2.1}$$

The adaptability is a new independent variable, which indicates the agent's willingness to change.

Finally, authorized agents ('authorities') were added. These are agents that will not change except for their location and do not experience belief entrenchment. This means their opinion remains unchanged. Their purpose is to test whether authorities can break up a polarized environment. The other, regular agents are now called followers.

2.1.2 Analysis

The described model has been used to gather information about echo chambers and epistemic bubbles.

The factors discussed in the introduction have been implemented in the model in the form of different parameters. To uncover information from the model we will analyse the results produced by different parameter settings. The model described above has a large number of parameters, not all of which are relevant for this data collection process. Here follows a list of relevant parameters and their descriptions. To be able reasonably compare any of the results from these simulations to results from Dykstra et al. (2013), the majority of the parameters are kept the same. Here follow the parameters that did change to produce a smoother running simulation. All settings for parameters as set in Dykstra et al. (2013) can be found in Table 6.1 in the appendix.

Parameter	range	setting
chance-walk	0-100	70
stepsize	0-2	0.32
loudness	0-20	6

Table 2.2: Adapted from Dykstra et al. (2013) parameter settings.

The setting of the following three parameters also require explaining. The first is the `number-of-followers` which are the regular agents as they are in DIAL. This parameter is set to 40, to keep the experiments feasible. Additionally, there are the two parameter that Dykstra et al. (2013) studied: `force-of-argumentation` and `force-of-norms`. Following their results these are set to 0 and 1 respectively, as those are the settings that produce a segregation state in DIAL. These parameters will not change while studying the potential influencing factors.

Several parameters have been added. The first three are fixed during the analysis of DIALx, for the final two I performed parameter sweeps to explore their influence.

- `init-division-ratio` The initial division ratio is the division of followers between the proponent and opponent groups. This is fixed at 0.5, so the groups are the same size.
- `adaptability` This parameter influences the rate at which the belief entrenchment of agents changes. This is set to 0.02.
- `init-belief-entrenchment` This parameter determines with which belief entrenchment the agents are all initialized.
- `init-density` An init-density of 15 means the agents of one group are distributed over a square covering 15% of the topic space.
- `ratio` This variable is the ratio followers:authorities implement as ratio:1. Together with the number of followers it determines the amount of authorities.

For the initial density, I have performed a parameter sweep over its complete range, from 0 to 100. As for the ratio, which does not have a set range, I have chosen to start at a ratio 1:1. This means this project does not cover the situation in which there are no authorities. As will become clear later, a situation with no authorities will not come to resolution, unless other changes are made to the model.

Finally there is one parameter that is kept fixed in Dykstra et al. (2013) for which I will perform a parameter sweep. The parameter `announce-threshold` (announcement threshold) determines which agents are allowed to make announcements, so which agents will have an influence on the others.

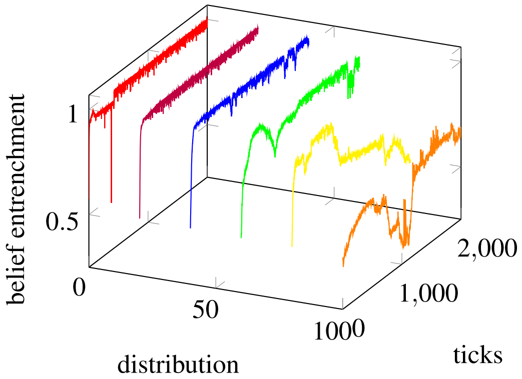
The data has been collected using functions provided by Netlogo, exporting all data points to external files in a format suitable to create graphs. Only the relevant data is collected. This data consists of the average belief entrenchment values of the two opposing and agreeing bubbles. It is important to note that in the implementation the bubbles are not the clustered followers, but the patches corresponding to the held believes. Further collected data is the clustering data and the difference in average evidence value between the two bubbles. All data is collected for each tick of the simulation, as it is the dynamics I am interested in.

The data is presented in graphs below in Section 2.2.

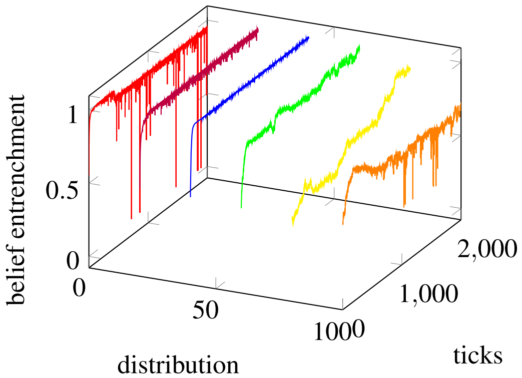
2.2 Results

In this section, I present a selection of the data that can be extracted from the DIALx simulations. Figures 2.2 and 2.3 are included to confirm the expected effects of clustering and initial belief entrenchment. The remainder of the results was selected to show the effects of the influencing factors I set out to study. For all factors I have executed a parameter sweep to study the resulting dynamics for different values of the factor. Each figure has the studied factor on the x-axis, time in ticks on the y-axis and on the z-axis the most indicative dependent parameter to show the effects on the dynamics. For each factor the data of belief entrenchment, clustering and evidence difference was collected. The graph that shows the most important dynamics in the simulations the best is included. All graphs are of data gathered from only the followers, unless noted otherwise.

Figure 2.2 shows the belief entrenchment over time for different values of initial distribution (x-axis). Figures 2.2a and 2.2b show the same parameters, but for the two bubbles in one simulation. The initial distribution parameter is the percentage of space over which the agents are spread out per bubble.



(a) Average belief entrenchment of followers with evidence between 0.5 and 1.



(b) Average belief entrenchment of followers with evidence between 0 and 0.5.

Figure 2.2: Parameter sweep over the initial distribution for DIALx.

The belief entrenchment is the most indicative dependent parameter in this case, because of the notable decrease in entrenchment as the distribution increases. This shows that distribution is a relevant factor in opinion dynamics.

Figure 2.3 shows the clustering of all agents over time with different initial belief entrenchment values. The initial belief entrenchment determines where the bubbles are on the scale between epistemic bubble and echo chamber.

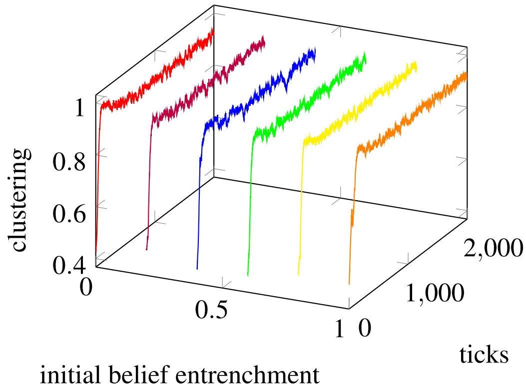


Figure 2.3: Clustering of followers in a parameter sweep over initial belief entrenchment in DIALx.

The values of clustering remain similar, no matter the initial belief entrenchment value. The constant high values on the z-axis suggest a limited influence of initial belief entrenchment.

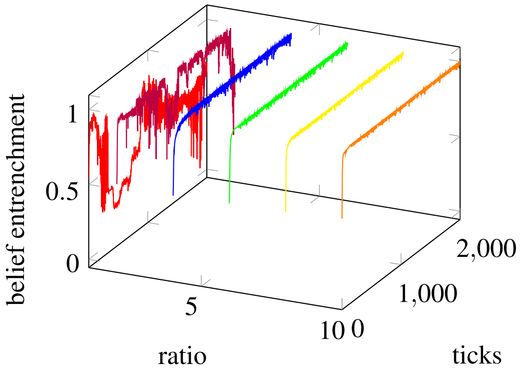
As the simulation did not present unwanted unexpected behaviour, we turn to the results for the influencing factor parameters.

The first parameter is the ratio between the followers and the authorities. Figures 2.4a and 2.4b show the same type of results, but for different bubbles. The x-axis represents the number of followers per authority.

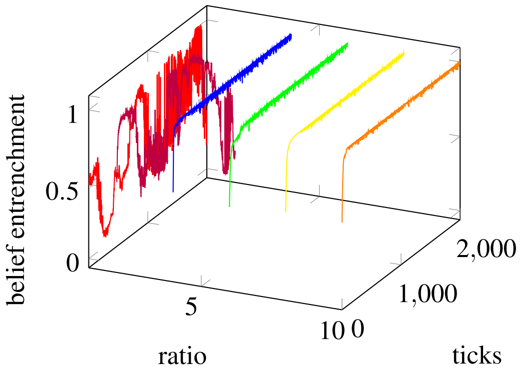
With a ratio of 1:1 or 1:2, the effect of ratio on the belief entrenchment of both bubbles is most visible. This is not remarkable, as the regular agents can not evade the overwhelmingly present authorities. On the other hand, how unaffected the belief entrenchment is with high ratio values is a more interesting observation obtained from these graphs.

The second influencing factor studied is the announcement factor, shown in Figure 2.5. This time the dependent parameter selected as the most indicative is the difference in evidence between the two bubbles, or actually between agents with an evidence value below and above 0.5. Important to note about this parameter is that if all agents agree (are all on the same side of 0.5), there are no values for evidence difference. The x-axis represents the announcement threshold between 0 and 2.5.

The results for the three lowest values of announcement threshold show that the two bubbles hold oppos-



(a) Average belief entrenchment of followers with evidence between 0.5 and 1.



(b) Average belief entrenchment of followers with evidence between 0 and 0.5.

Figure 2.4: Parameter sweep over the followers:authorities ratio for DIALx.

ing opinions, although they are not terribly entrenched. More interesting is the progress for the higher announcement threshold values. At 2, one of the bubbles resolves quickly and at 2.5, the evidence difference almost entirely disappears at certain points.

2.3 Discussion

In this section I will discuss the results presented above, as well as other findings that are not necessarily visible in the graphs.

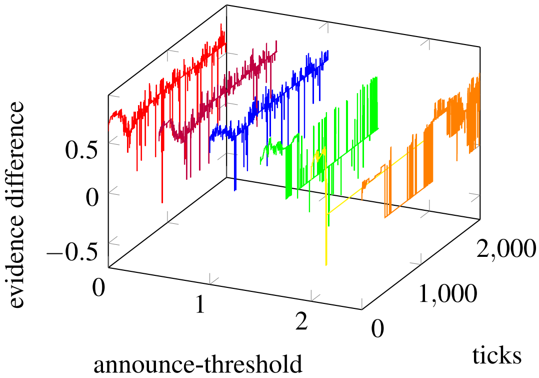


Figure 2.5: Difference between the average evidence values of proposing and opposing bubbles in a parameter sweep over announcement threshold parameter in DIALx.

2.3.1 Bubbles

Figures 2.2a and 2.2b show that initial distribution affects the belief entrenchment value of the followers over time. A low distribution means the followers are positioned in bubbles, which results in high belief entrenchment. As the the followers become more spread out, the belief entrenchment values decrease. Since the topic space is limited in size, a higher distribution implies a lower average distance between agents. As agents can be influenced by more agents and are influenced more strongly, their belief entrenchment is reduced. Figure 2.6 shows highly distributed agents, with varying degrees of belief entrenchment.

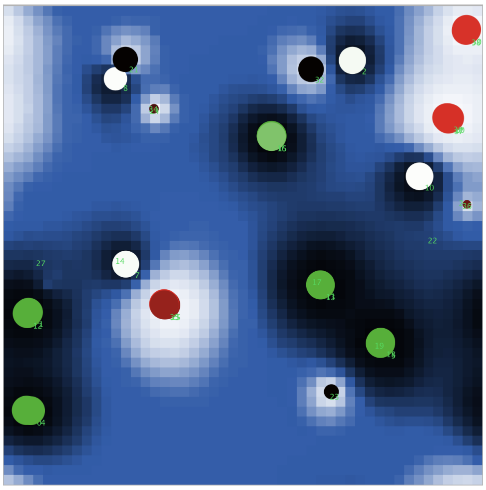


Figure 2.6: Highly distributed agents.

2.3.2 Epistemic bubbles

In order to implement belief entrenchment, the concept had to be reduced to several computations that represent it. In these computations, belief entrenchment affects the agents' evidence values and is affected by the change in their evidence value. One of the features of an epistemic bubble is a strong shared opinion among the members, where the evidence values are all very similar. The minimal changes in evidence values that follow will cause the belief entrenchment to increase. This high belief entrenchment will thus prevent the bursting of the bubble when exposed to new information. This effect can be observed in Figure 2.3. A low initial belief entrenchment eventually gives the same results as a high initial belief entrenchment. From this I conclude that the difference between epistemic bubbles and echo chambers is theoretically shown using belief entrenchment, even though in this model there is no effective difference.

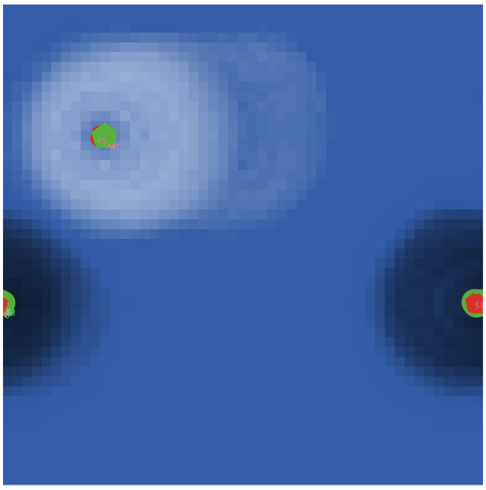


Figure 2.7: Discussion within bubbles.

2.3.3 Authorities

Figure 2.4 shows the results for different follower to authority ratios. Important to take away from these results is that only when the ratio is 1:1 or 1:2, the dynamics are distrupted to such an extent that bubbles break apart, more than two bubbles are formed or agents change opinions. This can be observed by looking at the belief entrenchment graphs, which are less steady for high numbers of authorities.

When there is a lower number of authorities, the final state is a stable state in which the agents remain within their bubbles. However, within the bubbles a fierce discussion is going on due to the presence of the authorities with opposing opinions. This can be seen in Figure 2.7, as the center of the bubble is not entirely white.

Even though these states in which there is discussion within bubbles are on the border of what is an echo chamber and what is not, according to the definition this is still an echo chamber. The authorities might have infiltrated the echo chamber spatially, they are still unable to break it up because the followers stick to their original opinion. What we instead observed is a discussion that takes place on only one side of the neutral position (so with all evidence value below 0.5 or all above 0.5).

2.3.4 Announcement threshold

As the results in Figure 2.5, show the evidence difference at a threshold value of 2.5 is 0, this means that all agents hold the same opinion. This shows a higher announcement threshold can cause an echo chamber to break. An echo chamber can break when at some point followers start to lose their ability to make announcements and authorities become the majority of the announcing agents. See Figure 2.8 for an example of such a break. Even though this factor can indeed allow echo chambers to break, in this model the representation of the difference in opinion in topic space remains flawed.

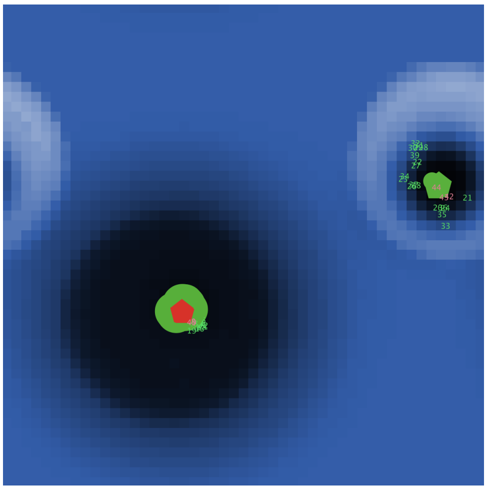


Figure 2.8: Example of a break caused by high announcement threshold.

2.3.5 Movement of the agents

How the agents move has a large impact on the complete opinion dynamics. Therefore, the movement of the agents through topic space should properly represent their opinion. In DIAL and DIALx, the agents move towards patches with similar evidence values as themselves. They will not leave this bubble of their own accord, as long as they do not change their mind. The same patch attracts like-minded agents forming bubbles, see Figure 2.9. Within these bubbles agents only confirm each other. This is the effect of an echo chamber at work. In the situation where the announcement threshold is 0, i.e. each agents can always make an announcement, the following behavior can be observed. For agents in a bubble to move apart, the agents need to be able to decide to move away from others. In the implementations of DIAL and DIALx the direction of the agents is so focussed that even adjusting all parameters to encourage moving apart, the agents have no reason to do so. The only way agents can move away from the core of a bubble is when they move towards another bubble, which happens rarely due to the surrounding like-minded agents. They will also move away from the center of a bubble when the opinion in the center changes, implying when the bubble has already burst. The agents will never spread out into single agents bubbles with unique opinions. When a bubble bursts, bubbles of the opposing opinion will form.

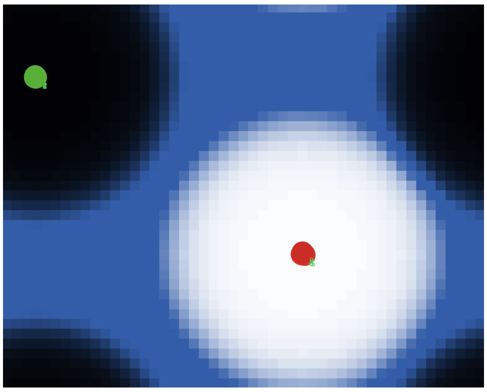


Figure 2.9: Example of clustering.

3 DIALx2

3.1 Methods

The previous model turned out to still have a major restriction: the way the movement of agents is determined does not allow them to move in a direction other than the centre of an echo chamber. Therefore, the model is unable to simulate the resolution of echo chambers in a spatial sense. The adaptation made to create this second model is the way the movement of the agents is determined. I aim for the result that the agents are able to spread out over the topic space in way that it represents their opinions.

3.1.1 Implementation

In DIAL and DIALx the movement of agents is determined by calculating a direction and moving a step in this direction. The agents turn towards a patch with a similar evidence value as their own, where the agent will most likely find like-minded agents. Since the influence that agents have on each other is spatially limited, this positioning indicates a similarity bias influence model. Increasing the complexity of the way in which the agents movement is determined should allow them to move away from other agents as well as moving towards them. This extension means enriching a similarity biased influence model with a repulsive influence (Flache et al., 2017). There are multiple ways to implement repulsed behaviour into this existing model, and the implementation choices will highly influence the resulting behaviour. In this model the agents move away from strong partisan patches when they are more neutrally opinionated themselves. This will result in a spatial representation that matches the agents evidence values. Important to note here is that, theoretically agents were already able to move towards neutral patches outside the bubbles, but the draw was insignificant. Therefore, a more complex implementation of movement is required, that encourages more complex behaviour in the agents. Two separate adaptations were made. The first was to the distance agents move. In DIALx the agents made 1 uncomplicated step forward. In DIALx2 the forward distance is directly proportional to the belief entrenchment and the distance moved backwards is inversely proportional to the belief entrenchment, as well as the naivety of the agent. Naivety is a new independent parameter. In an initial implementation the adaptability parameter and the naivety were the same parameter. This implementation was changed later to allow for more precise fitting. Since the agents can now move backwards, the direction agents move in is also changed. In DIALx they move to a patch with a similar evidence value, whereas in DIALx2 they move to a patch most similar to their evidence value polarized (rounded to 1 or 0).

3.2 Results

This results section is similarly structured to section 2.2 and the results have been attained using the same method as for DIALx. Missing are the results showing the influence of distribution and initial belief entrenchment, because these factors were sufficiently discussed in Section 2. Here follow parameter sweeps for the two influencing factors already previously discussed, namely the ratio and announcement threshold and a new factor, naivety.

First take a look at the ratio. Again belief entrenchment is displayed on the z-axis and the number of followers per authority on the x-axis. Figure 3.1 show that, compared to the results for DIALx, the influence of authorities seems to have decreased. The effects on belief entrenchment with a ratio 1:2 are now similar to those of higher ratios.

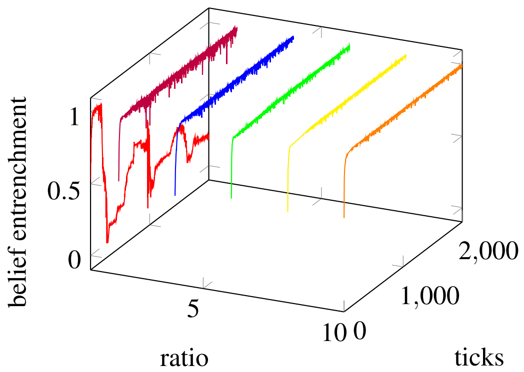
The second factor is the announcement threshold. The parameter indicating the effect on opinion dynamics is, as before, the evidence difference.

In Figure 3.2 we can see the effect of `announcement-thresholds` seems to have decreased compared to 2.5, which shows the effect for DIALx. The same difference between DIALx and DIALx2 as we could see for the ratio parameter. For an announcement threshold value of 1.5, the effect is now similar to the lower values, while in DIALx effects of `announcement-thresholds` were already noticable at that 1.5.

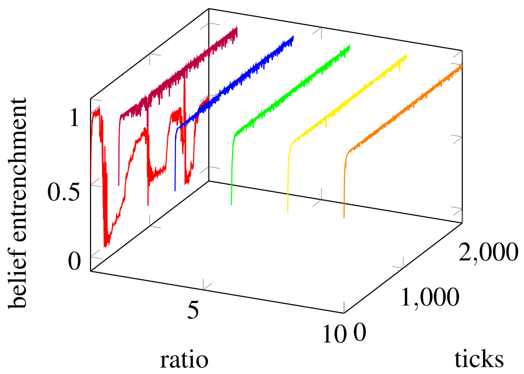
The final factor we will take a look at is naivety, newly introduced to DIALx2. Figure 3.3 shows the influence of this parameter (on the x-axis) on the clustering value (on z-axis). There difference between the different values of naivety is almost unnoticeable, while this is the most indicative parameter.

3.3 Discussion

Even though, theoretically, the implementation has changed, the effect is hard to detect. Since the proportion of moving backward and forward is determined by the belief entrenchment, the effect of echo chambers is still strong. Getting caught in an echo chamber implies high belief entrenchment values, still resulting in no movement away from the center. However, any other way of implementation with stronger results would have had the potential pitfall of being theoretically or empirically nonsensical. Depending on the goal of the simulation it might be an option to chose an implementation that causes agents to be more spread out over the topic space. The question is then 'do individuals move towards more neutral spaces when they hold a more neutral opinion?' An answer to this question could be drawn from other research fields in the form of empirical evidence.



(a) Average belief entrenchment of followers with evidence between 0.5 and 1.



(b) Average belief entrenchment of followers with evidence between 0 and 0.5.

Figure 3.1: Parameter sweep over the ratio for DIALx2.

A side-effect of this implementation is that agents run

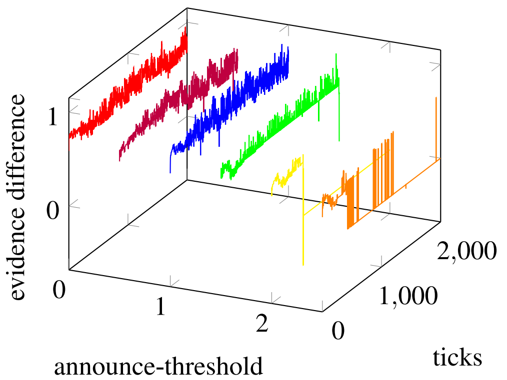


Figure 3.2: Difference between the average evidence values of proposing and opposing bubbles in a parameter sweep over the announcement threshold parameter in DIALx2.

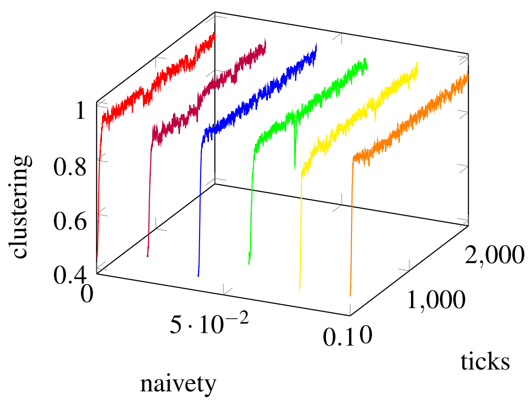


Figure 3.3: Clustering data of all agents in a parameter sweep over the naivety parameter in DIALx2.

away from authorities, as the authorities affect the evidence values of patches. This is caused by the fact that the presence of the authorities changes the evidence values of the patches. Since the regular agents determine their direction by moving towards a patch with a high agreement factor to them, the patches with authorities are less attractive to the regular agents to move to.

As mentioned before, the naivety and adaptability parameters are two seperate parameters in the current implementation, but this could potentially be captured in one single (more complex) parameter representing the agents susceptibility to different influences. Formalizing such a parameter requires a look towards psychological empirical research on group behavior and peer pressure.

4 General discussion

DIAL, DIALx and DIALx2 are successive models. In this section I will compare each of the models to its predecessors and discuss possibilities for further research.

All models are similarity biased influence models (Flache et al., 2017). Agents with similar opinions are closer neighbours and a smaller distance implies a stronger influence of agents on each other. Which means that similarly opinionated agents influence each other more heavily.

As discussed in Section 3.1, the two main changes made between DIAL and DIALx were the addition of belief entrenchment and the possibility to add authorities, agents that do not change their opinion. The former did change the dynamics, as the speed at which agents change their opinion became variable. What remain unchanged was a final state of segregation, starting from a random distribution over topic space.

In DIALx, we could observe that the initial belief entrenchment did not make a difference in the long-term behaviour of the agents. This behaviour discarded the distinction between epistemic bubbles and echo chambers in the simulation. Further research into this is required to find out what *does* define the distinction between these two types of epistemic structures. As it appears that belief entrenchment, as implemented in DIALx, is not doing the concept epistemic bubbles (Nguyen, unpublished) justice. I have shown that not only strong beliefs are a reason to reject opposing opinions, as it is the case that in an epistemic bubble, due to continuous positive confirmation, the belief entrenchment is also high.

When authorities were added, the dynamics changed more drastically, as the opinion of regular agents can no longer stabilize. In the DIALx simulations a form of infiltration of the echo chambers could be noticed as the authorities caused continuous discussion within bubbles. These were exactly the changes I aimed for when implementing DIALx. What remained unchanged was the seemingly unavoidable formation of bubbles in topic space, this I did not intend.

DIALx2 was a response to the clustering that remained in DIALx. I responded by changing the implementation of the agents' movements. The aim was to create repulsive movement and for the position of agents relative to other agents to be a better representation of their opinion relative to the opinion of the other agents. The dialogues between two agents remain unchanged. The agents do explicitly move away from agents with dissimilar opinions, which means DIALx2 is more akin to a repulsive influence model, while actually still being a similarity biased influence model (Flache et al., 2017).

An interesting side-effect of the repulsive movement is that followers move away from authorities in DIALx2. This reduces the effect that authorities have on the other agents, since their relative position to these agents affects their influence on them. This decreases the amount of internal discussion in bubbles.

It is possible for authorities, in both DIALx and DIALx2, to change the opinion in a bubbles by becoming the majority in the pool of agents that are able to make announcements. This is possible because of the announcement threshold. As more regular agents drop below the threshold, the authorities can gain the majority. In DIALx2 this effect decreases as it is harder for authorities to gain the majority. The authorities are less capable at infiltrating bubbles.

The question of this bachelor's project was

> How can echo chambers and epistemic bubbles be resolved?

I have interpreted the 'how' by looking at many different parameters and their influence on the opinion dynamics. Two parameters showed the most potential to contribute to the resolution of echo chambers, namely the threshold at which agents can make announcements and the number of authorities. I studied these parameters using a parameter sweep for each of them individually, while leaving the all other parameters unchanged. In further research the relation between these factors and others can be studied more extensively.

In the methods section I state that empirical research concerning this topic wil be tottering from a moral perspective. Projects like these allow us to distile which, more specific, parts can be tackled by empirical research. I suggest that such research is important to be considered, before any further attempts to extend the current models are made. Multidisciplinary research, as noted by Flache et al. (2017), is critical to support choices concerning the implementation, as well as the parameter settings. Research from the fields of psychology and sociology, especially empirical research can help confirm findings done using modelling as a tool.

Now I will consider several design choices that were made and could reconsidered, with potential back-up from research from other fields.

Both in DIALx and DIALx2 belief entrenchment is implemented with only a direct two-way relation to the evidence value of a particular agent. It has not been ruled out that other factors can also influence the belief entrenchment.

Second, the implementation of the authorities are currently implemented with the idea that they try to convince any agent that needs most convincing. This is done by having them move towards the agent that has an opinion furthest from their own. Other than that, they have similar capabilities as regular agents. Yet, there are lots of parameters that can be separated for regular agents and authorities. Including the range at which authorities are heard or the impact that the evidence of an authority has on the evidence or belief entrenchment of regular agents.

Additionally, for authorities as well as regular agents, the implementation of a form of memory could drastically change the opinion dynamics. This could for example allow authorities to more strategically go on campaigns, instead of chaotically moving towards any agent holding the most opposing opinion. Other information from such a memory that could be taken into account could be for example for how long an agent has held a certain opinion.

Besides these suggestions, there is one final point of improvement for DIAL and its successors, for which empirical back-up seems superfluous. On the one hand, we have the interpretation of the research question, according to which I have studied the models and looked at the results in this project. This is with a focus on what can bring resolution about. On the other hand the research question raises an important implication about the state after resolution. The aim of the resolution of echo chambers and epistemic bubbles is a new situation, contrasting the segretated state or authoritative state (Dykstra et al., 2013). This ideal situation is one where meaningful discussion takes place among the agents and they all hold slightly different opinions in order to have this discussion. Looking at the way the model is abstracted from the real world, removing countless details that are irrelevant, I have distinguished one detail that can stop the agents in DIALx or DIALx2 getting into this ideal situation of meaningful discussion. This is the injection of new knowledge about the proposition on which agents hold an opinion, into the agents' non-existent knowledge base. In a discussion in a real world situation, those in discussion mostly have access to much more than only the opinion others hold on a certain proposition. New information can be added to the discussion from other sources.

5 Conclusion

I asked my research question with the intention to create and research multi-agent models to better understand epistemic bubbles and echo chambers and their resolution. The overal findings from the collection of research being done on echo chambers, may eventually lead to a model that helps to predict the developements of real world echo chambers. I have mostly found what needs to be done following these results, in order to create better understanding of opinion dynamics and better simulations thereof. However, I was able to identify two main influencing factors, the threshold at which agents can make announcements and the number of authorities compared to the number of followers. As could be expected, more authorities increases their influence. A higher announcement threshold allows these authorities to dominate the announcements causing a bubble to burst. DIALx showed that with many different parameter settings, the initial state will converge to unresolvable clustering, but not always with high polarization and belief entrenchment. DIALx2 was still unable to deliver the ideal situation in which agents spread over topic space, representing meaningful discussion. Even though agents can be convinced in both models, new bubbles will form. DIALx2 still showed the same strong echo chamber effects, which is not entirely faulty. Considering this I have concluded that what is missing from these models is the injection of new information over time. A meaningful discussion is the constant formation and resolution of weak epistemic bubbles and what we can see in these models is only a single iteration of such a discussion.

References

Bert Baumgaertner. Yes, no, maybe so: A veritistic approach to echo chambers using a trichotomous belief model. *Synthese*, 191(11):2549–2569, 2014.

Piter Dykstra, Corinna Elsenbroich, Rineke Verbrugge, and Gerard Renardel de Lavalette. Put your money where your mouth is: DIAL, A dialogical model for opinion dynamics. *Journal of Artificial Societies and Social Simulation*, 16(3), 2013.

Andreas Flache, Michael Mäs, Thomas Feliciani, Edmund Chattoe-Brown, Guillaume Deffuant, Sylvie Huet, and Jan Lorenz. Models of social influence: Towards the next frontiers. *Journal of Artificial Societies and Social Simulation*, 20(4), 2017.

Seth Flaxman, Sharad Goel, and Justin M. Rao. Filter bubbles, echo chambers, and online news consumption. *Public Opinion Quarterly*, 80:298–320, 2016.

Jens U Hansen. Pluralistic ignorance: A case for social epistemology and epistemic logic. *Proceedings of the workshop on Epistemic Logic for Individual, Social, and Interactive Epistemology*, pages 1–15, 2014.

Pelle G Hansen, Vincent F Hendricks, and Rasmus K Rendsvig. Infostorms. *Metaphilosophy*, 3(44):301–326, 2013.

Rainer Hegselmann and Ulrich Krause. Opinion dynamics and bounded confidence: Models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5(3), 2002.

Vincent F Hendricks. Knowledge transmissibility and pluralistic ignorance: A first stab. *Metaphilosophy*, 41(3):279–291, 2010.

C. Thi Nguyen. Echo chambers and epistemic bubbles. unpublished.

6 Appendix

6.1 Parameter settings

Parameter	range	Dykstra	DIALx
chance-announce	0-100	38	38
chance-question	0-100	0	0
chance-attack	0-100	12	12
chance-learn-by-neighbour	0-10	0	1
chance-learn-by-environment	0-10	1	0
chance-mutation	0-2	0	0
chance-change-strategy	0-10	0	0
chance-walk	0-100	27	70
stepsize	0-2	0.8	0.32
undirectness	0-45	26	26
visualhorizon	0-20	5	5
loudness	0-20	2.5	6
neutral importance	0-1	0.5	0.5
firmness-of-principle	0-10	2.6	2.6
lack-of-principle-pen	0-1	0.07	0.07
attraction	0-1	0.47	0.47
rejection	0-1	0.47	0.47
winthreshold	0-1	0	0
inconspenalty	0-1	0	0
forgetspeed	0-0.005	0.00106	0.00106

Table 6.1: Dykstra et al. (2013) setting