



Universiteit
Leiden
The Netherlands

Understanding epistemic bubbles and echo chambers: A network epistemological approach

Dana ,

Citation

Dana,. (2025). *Understanding epistemic bubbles and echo chambers: A network epistemological approach*.

Version: Not Applicable (or Unknown)

License: [License to inclusion and publication of a Bachelor or Master Thesis, 2023](#)

Downloaded from: <https://hdl.handle.net/1887/4280912>

Note: To cite this publication please use the final published version (if applicable).

Understanding epistemic bubbles and echo chambers: A network epistemological approach

Student: Dana Bouwknecht

Program: MA philosophy (60 ECTS)

Specialization: Philosophy of Knowledge

Supervisor: dr. Jan Sleutels

Date of submission: 20-07-2025

Abstract

In the past decade, there has been a boom in scholarly literature regarding epistemic bubbles and echo chambers. However, how these concepts have actually been defined and researched has been inconsistent throughout this scholarly literature, making it difficult to compare different findings. In his influential article *Echo chambers and epistemic bubbles* (2020), philosopher C. Thi Nguyen has set out to distinguish epistemic bubbles and echo chambers from each other, arguing that the epistemic bubbles mechanism is based on exposure while the echo chamber mechanism is based on trust. Although Nguyen does argue that we should treat these concepts in a more network-oriented manner, his main approach is arguably still quite individual-oriented and could benefit from a more systemic and network-oriented approach. Analyzing (social) epistemological phenomena on a systemic and network-based level is the domain of *network epistemology*, a philosophical field in which researchers use mathematical models to simulate the spread of beliefs through a network of agents. In this thesis, four questions are identified that are left based on Nguyen's contribution, which all four could benefit from a more network-oriented perspective. These questions relate to the following aspects: 1) Nguyen's discussion of how one gets into an echo chamber, 2) the intentionality behind the emergence of echo chambers, 3) the easiness of shattering an epistemic bubble, and 4) the interplay between epistemic bubbles and echo chambers. Findings from the field of network epistemology are used to answer these questions. Moreover, network epistemological research is used to shine a new light on the much debated potential role of digitalization in the formation of echo chambers and epistemic bubbles. Some main lessons that can be drawn from the endeavors in this thesis are that 1) even rational and epistemically virtuous agents can get trapped in echo chambers based on a simple and justifiable heuristic of trust, 2) it is crucial to correctly distinguish epistemic bubble and echo chamber mechanisms in empirical research regarding these concepts to validly analyze both phenomena, and 3) that digitalization could have a potential effect on echo chamber formation separately from its impact on epistemic bubbles.

Table of Contents

Abstract	2
1. Introduction	4
2. Epistemic bubbles and echo chambers: Nguyen's view.....	7
2.1. Origins.....	7
2.2. Nguyen's contribution	8
2.2.1. Problems and gaps	12
3. Network epistemology	14
3.1. What is network epistemology?	14
3.1.1. The use of network epistemology	17
3.2. Current research.....	18
4. Bubbles and chambers in a network.....	20
4.1. Getting into an echo chamber	21
4.2. The intentionality of echo chambers	26
4.3. Shattering an epistemic bubble	28
4.4. The interplay of bubbles and chambers	30
5. The effect of digitalization	33
5.1. Conflation, again	34
5.2. The downsides of big networks	37
6. Conclusions and interventions	39
6.1. Potential interventions	40
6.2. Limitations and further research.....	42
Bibliography.....	44

1. Introduction

In the past decade, especially since Brexit and the electoral victory of Donald Trump in 2016, interest in the terms ‘epistemic bubbles’, ‘echo chambers’, and ‘filter bubbles’ has risen quite dramatically. Statistics provided by Mahmoudi et al. (2024) even indicate that more papers on echo chambers were published in 2021 and 2022 than in all previous years combined. The increase in attention for these social epistemological phenomena seems to be linked to increasing worry for a ‘post-truth era’ (Lewandowsky et al., 2017; McIntyre, 2018), the spread of fake news and misinformation (Mønsted & Lehmann, 2022), and societal polarization in multiple forms (Arguedas et al., 2022; Munroe, 2023). The growing scholarly concern for polarization even takes the form of concern for factual polarization, which denotes the loss of a shared reality between different groups within western populations (Rekker & Hartevelde, 2024).

However, while interest in epistemic bubbles and echo chambers has peaked in recent years, consensus on what these concepts actually *mean* is missing. Systematic reviews show that epistemic bubbles and echo chambers have been defined and researched in many different ways, sometimes even without a clearly specifying the concept in question, leading to inconsistent and sometimes incomparable findings (Hartmann et al., 2025; Terren & Borge, 2021). This inconsistency is also present in literature analyzing the effect of digitalization on epistemic bubbles and echo chambers (Mahmoudi et al., 2024). One author has even called epistemic bubbles and echo chambers “the dumbest metaphor on the internet” (Bruns, 2019, p. 8), pointing to the lack of a clear definition of these terms.

Next to this inconsistency within the empirical sciences regarding bubbles and chambers, there also seems to be a lack of integration between empirical and philosophical discussions of these concepts. On the one hand, philosophers often discuss epistemic bubbles and echo chambers in a rather hypothetical but sweeping manner, even if the empirical basis for their claims is questionable or lacking. For example, multiple theoretical scholars have warned us for the grave dangers of filter bubbles or echo chambers and have implied that digitalization has made their influence on contemporary society very big and worrisome (e.g. Pariser, 2011; Sunstein, 2018), even though numerous empirical scholars have found evidence against such claims (Arguedas et al., 2022; Masip et al., 2020). On the other hand, social scientists have often neglected to adequately ground their research on these phenomena in a solid philosophical or theoretical basis, leading to doubts about the validity of their research. This lack of a strong theoretical basis could arguably at least partially account for the lack of consensus among social scientists about the exact meaning of concepts like echo chambers and epistemic bubbles and

the consequent uncertainty about how well different findings can be compared (Hartmann et al., 2025; Terren & Borge, 2021).

Responding to this lack of consistency, philosopher C. Thi Nguyen has given a very influential clarification of both epistemic bubbles and echo chambers. In his article *Echo chambers and epistemic bubbles* (2020), cited hundreds of times in the past five years,¹ the author advocates for a clear distinction between the two concepts. In short, Nguyen argues that epistemic bubbles are based on exposure to relevant sources, while echo chambers are based on discrepancies in trust. The author shows that these mechanisms are relatively distinct, and argues that both earlier philosophers and empirical researchers have unjustly conflated the two terms which has led, especially in empirical analyses of these phenomena, to misguided conclusions. According to Nguyen, a systematic distinctive analysis of both concepts is crucial to accurately understand and analyze both. Moreover, Nguyen argues that epistemic bubbles and echo chambers should not only be seen as individual or psychological phenomena, but rather as inherently social and at least partially as network-related. He argues that people can sometimes even end up in echo chambers through no fault of their own, but rather due to the network they are part of and their place in it.

With his conceptualization, Nguyen has made an influential contribution to the clarification of both epistemic bubbles and echo chambers. However, while Nguyen does argue that we should look at these epistemological constructs as network problems, the approach used in his article is still rather focused on the individual and is still rather example-based. This impedes a more generalizable understanding of epistemic bubbles and echo chambers and how they might be analyzed in a more systemic and network-based manner. Moreover, a number of gaps and contradictions can be found in Nguyen's work, which I argue could be at least partially solved by a richer network-based analysis of the bubble and chamber mechanisms. In other words: I argue that to get a complete picture of these epistemological networked structures, it is very important to analyze them as such.

Analyzing (social) epistemological phenomena on a systemic and network-based level is the domain of *network epistemology*. Researchers in this relatively new and exciting philosophical field use mathematical models to simulate the movement of beliefs and knowledge through a network (Weisberg, 2021), analyzing how certain factors or behaviors can influence groups of people at large. The goal of network epistemological approaches is exploratory: they help identify in what could be the *potential* driving factors behind certain

¹ Based on metrics from publisher: see <https://www.cambridge.org/core/journals/episteme/article/abs/echo-chambers-and-epistemic-bubbles/5D4AC3A808C538E17C50A7C09EC706F0#metrics>

(epistemic) phenomena such as epistemic bubbles and echo chambers. I argue that the field of network epistemology can function as the needed bridge between social empirical research and ‘armchair philosophy’ in bettering our understanding of concepts like epistemic bubbles and echo chambers. For the philosopher, network epistemology helps to put hypothetical ideas into models that help analyze the viability of philosophical arguments about potential drivers behind these phenomena. For the social scientist, the theoretical models created this way can help to provide a basis for the development of actual empirical research and can be used as an indicator of the empirical model’s validity.

In this thesis, I will use insights from the field of network epistemology to enrich our understanding of epistemic bubbles and echo chambers. I use Nguyen’s (2020) influential conceptualization of these phenomena as starting point for this endeavor. In the next two chapters, I will further elaborate on Nguyen’s conceptualization of epistemic bubbles and echo chambers and introduce the field of network epistemology. In chapter four, I will use insights from research in the field of network epistemology to enrich and challenge certain aspects of Nguyen’s conceptualization of epistemic bubbles and echo chambers. One topic that has received particular attention in recent literature regarding epistemic bubbles and echo chambers is the potential role of digitalization in their prevalence and impact. In chapter five, I will dive more into this potential role of digitalization in the occurrence of both phenomena using findings from network epistemology. I will conclude with a discussion of possible interventions to alleviate the potential impact from echo chambers.

2. Epistemic bubbles and echo chambers: Nguyen's view

2.1. Origins

Both the concepts of epistemic bubbles and echo chambers already exist for two decades. The term 'epistemic bubble' was introduced by Woods (2005), but under quite a different definition than is used now. To put it shortly, Woods used the term in his paper to describe a state of inability of an agent (or person) to distinguish between their perception of knowing a certain proposition p and *thinking* that they know p , and the consequent idea that the person will therefore always opt for the perception of knowing p . This distinction between actually knowing p and only thinking that one knows p can according to Woods, only become apparent to the person if someone else comes in to correct them with counterevidence of p . According to this definition then, every person lives in their own epistemic bubble, in which actually knowing something and only *thinking* to know something are indistinguishable to that person (Woods, 2005).

The concept of an epistemic bubble has evolved quite a lot since this definition, although a link to the original definition can still be detected. One other important contribution to the current definition of 'epistemic bubble' has been done by Eli Pariser (2011). Pariser introduced the term 'filter bubble' to describe the phenomenon of how, in a digital world, algorithms filter information so that it is hyper-personalized and already in line with an individual receiver's beliefs. According to Pariser, each person that (often unknowingly) makes use of these algorithms ends up in a bubble in which the presented information is biased (filtered) in a very individualized manner. The term 'filter bubble' has become very popular since then. Whereas filter bubbles might thus be a specifically digital concept, an epistemic bubble is a more overarching phenomenon that includes filter bubbles, but which can also exist offline (Nguyen, 2020). In this thesis, I will use the epistemic bubble as an overarching concept of which the filter bubble can be seen as a specifically digital version.

The term 'echo chamber' has been coined by Cass Sunstein (2001), as referring to a social network in which people are selectively exposed to information that is fitting with their own existing attitudes, and less (or not at all) to information sources expressing dissimilar opinions. Sunstein has argued that echo chambers are a danger stemming from personalized information algorithms since the upcoming of the internet (2001) and later of social media (2018).

In the first decade since the introduction of these terms, attention for epistemic bubbles and echo chambers stayed relatively limited. However, since Brexit and the first electoral victory of Donald Trump in 2016, interest in the terms 'epistemic bubbles', 'echo chambers', and 'filter

bubbles' has boomed (Mahmoudi et al., 2024). Since the initial definition of echo chambers by Sunstein (2001, 2018) bears quite some resemblance to the definition of filter bubbles by Pariser (2011), it is not odd that these terms have often been conflated in scientific literature. Even so, the concepts have evolved quite a lot since then. In addition, several authors that have reviewed recent literature on echo chambers and epistemic bubbles have pointed out that there is a lack of consensus on the meaning of these concepts in general, making it difficult to assess how well different findings regarding epistemic bubbles and/or chambers can be compared (Hartmann et al., 2025; Mahmoudi et al., 2024; Terren & Borge, 2021).

2.2. Nguyen's contribution

Wanting to clear up this confusion, Nguyen wrote his paper *Echo chambers and epistemic bubbles* (2020). The author detected that actually not one, but two mechanisms were analyzed under these two concepts that were lumped together. With this paper, Nguyen has left a very influential mark on the discussion of both epistemic bubbles and echo chambers, arguing for a clear distinction of these two terms in scholarly debate and research. He provides clear definitions for both concepts, and sets out why the difference in mechanisms of both phenomena makes it much harder to escape an echo chamber than an epistemic bubble. The reason why these concepts are often conflated is that these phenomena do bear some similarities: they are both social epistemic structures in which an agent can find themselves, and in both structures, relevant knowledge is being excluded in a way. Other than that, epistemic bubbles and echo chambers are actually quite different according to Nguyen. Let us first look at an epistemic bubble.

According to Nguyen, an epistemic bubble is “a social epistemic structure in which other relevant voices have been left out, perhaps accidentally” (p. 141). These structures are ubiquitous and can form through the normal processes of finding friends and a community. According to Nguyen, people often interact with people that are similar to them, for example in ideology. This can lead to a biased or incomplete set of information that gets shared in this network of similar people. For example, if one gets information about nuclear energy from their Facebook friends, but everyone in their community is mostly against nuclear energy, then there is a big chance that these friends share news about the dangers of nuclear power plants, but not about the benefits. According to Nguyen, in epistemic bubbles there is a lack of *coverage reliability*, which is “the completeness of relevant testimony from across one's whole epistemic

community” (2020, p. 143). In our example, the lack of coverage reliability exists in the lack of information about the possible benefits of nuclear power plants.

In epistemic bubbles, the lack of coverage reliability does not necessarily have anything to do with any of the single individuals in it: rather, Nguyen argues, should coverage reliability be seen as a network problem, which arises because of poor connectivity to other relevant sources. In epistemic bubbles, relevant voices are thus “excluded by omission” (p. 143). According to Nguyen, this can either happen through selective exposure, which is the tendency of an individual to seek like-minded sources, or through the external changing of one’s epistemic landscape, for example through censorship or through online algorithmic filtering. The author states that people sometimes speak of a filter bubble in this last situation because of its solely digital existence, just like the term was coined by Pariser (2011).

According to Nguyen, epistemic bubbles are relatively easy to shatter: if someone encounters new relevant information that was first missing in their epistemic bubble, their bubble regarding that topic gets (at least partially) burst. In that sense, Nguyen argues that people also have a certain responsibility in escaping their respective bubbles as well as they can by trying to create a network of information provision that has a high coverage reliability and that is thus relatively complete. Of course there will always be missed information, but the author argues that one is for example responsible to make sure that they once in a while look for news from a number of relevant sources and to not only rely on posts from their Facebook friends.

This is very different in the case of echo chambers, which, Nguyen argues, are “much harder to escape” (p. 153). Just like epistemic bubbles, echo chambers are (problematic) social epistemic structures in which certain voices and knowledge are excluded. But while this happens in an epistemic bubble through (unintentional) exclusion by omission, in an echo chamber this exclusion happens through the active discrediting and undermining of other voices and sources. According to Nguyen, echo chambers lead their members to actively distrust all sources outside of the echo chamber. He argues that while echo chambers could theoretically come to exist unintentionally, they are in practice almost always set up intentionally by an actor to exert control over its public. The author defines an echo chamber in the following way:

I use “echo chamber” to mean an epistemic community which creates a significant disparity in trust between members and non-members. This disparity is created by excluding non-members through epistemic discrediting, while simultaneously amplifying members’ epistemic credentials. Finally, echo chambers are such that general agreement with some core set of beliefs is a prerequisite for membership, where those core beliefs include beliefs that support that disparity in trust. By “epistemic discrediting”, I mean that non-members are not simply omitted or not heard,

but are actively assigned some epistemic demerit, such as unreliability, epistemic maliciousness, or dishonesty. By “amplifying epistemic credentials”, I mean that members are assigned very high levels of trust. Of course, these two processes can feed back into one another. (Nguyen, 2020, p. 146)

In Nguyen’s view, the biggest difference between epistemic bubbles and echo chambers is thus the role of trust. Nguyen argues that while the mechanisms in epistemic bubbles do limit the access of outside information, they do not change the credibility of these sources. In echo chambers, active discrediting of outside sources is used to make members exclude information that is not in line with the beliefs in the echo chamber. With this definition of echo chambers, Nguyen actively departs from Sunstein’s (2001) original use of the term, which he argues is more similar to his own conceptualization of the epistemic bubble.

According to the author, one way the disparity between in-group and out-group trust is created in echo chambers is through *disagreement-reinforcement mechanisms*. In such a mechanism, disagreement with the expressed belief is explained in such a way that, if uttered by outside sources, it will only reinforce the original belief (Nguyen, 2020). An example (that is sadly also an actual conspiracy theory) is if I argue that leftist politicians secretly want to replace the white population with migrants, but they do not want the white population to know so they will deny these accusations and argue that the people that suspect them of this are extreme and crazy. If leftist politicians then actually do deny these accusations and call them bizarre, my original statement is reinforced because I already argued that they would deny it. This example shows that poor connectivity regarding information transfer in the epistemic network is theoretically not necessary for an echo chamber to exist: Nguyen argues that exposure to outside information might even strengthen the echo chamber, or otherwise have no effect. Echo chamber members are thus not necessarily isolated from other information flows, like in epistemic bubbles: rather, they are isolated “credentially” (2020, p. 147), in that they only *trust* and take up information that fit their beliefs. This makes them “overly dependent” (2020, p. 142) on inside sources of information, according to Nguyen. In this conceptualization, echo chambers and epistemic bubbles can thus theoretically exist completely separately from each other, even though they might often overlap in practice.

In addition, the author argues that both the mechanisms of epistemic bubbles and echo chambers are built on what are in moderation healthy epistemic practices. Nguyen states that selection and exclusion of certain sources is necessary in contemporary society since “the world is overstuffed with supposed sources of information, many of them terrible” (2020, p. 143). Moreover, we live in a hyper-specialized world, in which it is impossible to be an expert on more

than a fraction of all the knowledge out there (Goldman, 2001; Nguyen, 2020). This makes it necessary to rely on information provided by others, for example on the expertise of my doctor if I get ill, or on the expertise of a civil engineer for assessing the safety of a bridge. This combination of a hyper-specialized world in which there are also many terrible or unreliable sources of information makes the use of selection criteria and heuristics necessary. Nguyen mentions the *litmus test* as an example of a heuristic that is in practice sometimes rightfully used. In a litmus test a person's standpoint on a certain topic (e.g. vaccination stance) is used to decide whether that person is a reliable source for other kinds of information (e.g. information on human health in general). Excluding or distrusting certain sources of information is thus necessary in today's complex society, and because of the abundance of information this sometimes needs to be done based on heuristics.

Epistemic bubbles and echo chambers even more so thus “function parasitically” (2020, p. 149) on our normal epistemic practices. Nguyen even argues that in the case of an echo chamber, this perversion can result in a situation in which a person that is trapped in an echo chamber can act *epistemically virtuous*, meaning that they use good and well-motivated epistemic practices (e.g. critically assessing new information against their earlier acquired information), and still reach the conclusion that all the information from outside sources is false. This is because the discrediting mechanism of echo chambers has led them to have beliefs that include the (false) idea that all outsiders try to mislead them and are untrustworthy. In this sense, echo chambers “convert individually epistemically virtuous activity into collective epistemic vice” (2020, p. 155). This is what Nguyen calls the “reverse-mandevillian nature” (2020, p. 155) of echo chambers and it is part of what makes it arguably so difficult to leave an echo chamber. Since this reverse-mandevillian nature is not present in epistemic bubbles, it is arguably easier to escape an epistemic bubble than an echo chamber. Therefore, Nguyen argues, people are also more morally blameworthy for staying in a bubble than a chamber: one can act epistemically virtuous and stay trapped in a chamber, but one cannot act virtuous and stay trapped in a bubble.

To sum up, both epistemic bubbles and echo chambers are social epistemic structures that lead to the exclusion of certain (sources of) knowledge. Both structures are the result of a perversion of healthy and virtuous epistemic practices, even though the level of perversion of these practices is much higher in echo chambers. But there are also a lot of differences between the two: Nguyen argues that trust plays no significant role in the mechanism of epistemic bubbles, while it is the motor behind the mechanism of echo chambers. Epistemic bubbles are based on poor connectivity and/or a high homogeneity in one's epistemic network, while poor connectivity or homogeneity of sources is not necessary for a person in an echo chamber: they

have to be credentially isolated. This also means that both structures can exist separately, even though they might occur more often together in practice. In addition, echo chambers are arguably much harder to escape than bubbles, and one is thus also more blameworthy for not escaping their bubble than their echo chamber. Nguyen argues that the difference in mechanism behind the two phenomena calls for a difference in interventions: simple exposure to other sources is theoretically enough to burst one's bubble, but might only reinforce their commitment to an echo chamber. To escape an echo chamber, one needs to radically alter one's beliefs in who to trust. Lastly, Nguyen argues that while epistemic bubbles frequently come into being accidentally, echo chambers are in practice much more likely to be set up intentionally.

All these differences make that Nguyen argues that echo chambers and epistemic bubbles have unjustly been lumped together in (empirical) research on these phenomena. He argues that the distinction of them is crucial for the correct identification and assessment of these phenomena in empirical research, and to be able to use the correct intervention.

2.2.1. Problems and gaps

Nguyen's conceptualization of epistemic bubbles and echo chambers and his arguments for the relevance of distinguishing between the two are arguably very useful and have provided very helpful insights for other (so far mostly theoretical) scholars (e.g. Begby, 2024; Turner, 2023). His conceptualization of both structures has laid a strong foundation for further research on these phenomena. However, I would argue that there are a few ways in which this conceptualization could be enriched or altered. For example, while Nguyen does hint in his paper that we should analyze epistemic bubbles and echo chambers on a more systematic and network-oriented level, his main approach is still quite individual-oriented. Nguyen states that problems behind epistemic bubbles and echo chambers are system-related, but the author still mostly points out how individual epistemic vices can lead someone to enter an echo chamber or an epistemic bubble. Such vices have little to do with the network a person is in. This contradiction can also be detected when Nguyen argues that there are "many cases" (2020, p. 154) in which a person might be blameless for ending up in an echo chamber, but the only case he discusses in his paper is if someone is born and raised in an echo chamber. This case is arguably highly specific, and does not help us to identify factors that are applicable to a bigger part of the population. One question that remains unanswered is then what we could identify as other possible factors that lead people to enter echo chambers, apart from being born into one.

A second aspect where I question Nguyen's argumentation relates to the purported intentionality that Nguyen argues is often present in the formation of an echo chamber. The author argues that echo chambers are almost always set up intentionally by its makers to control

their audience and to increase their power. However, like Munroe (2023), who provides the example of Pizzagate as an unintentional echo chamber, I disagree with Nguyen (2020) that an intentional setup is the most likely cause of echo chambers to arise. There exist more examples of conspiracy theories and fake news stories that seem to come into being unintentionally, such as the fake news story of pet-eating immigrants in Springfield (Lozano, 2024). This raises the question: Can we explain the occurrence and arising of echo chambers that come into being unintentionally?

Another contradiction found in Nguyen's discussion regards to the (lack of) difficulty of shattering one's epistemic bubble. On the one hand, the author argues that it is relatively easy for a person to shatter their own epistemic bubble by including a wider variety of relevant sources of knowledge to their network. On the other hand, the author simultaneously provides some arguments as to why shattering one's bubble could be harder than it seems, for example by discussing the problem of *bootstrapped corroboration* (to be further explained in chapter four) and by arguing that because of our complex and overstuffed information system it is good and even necessary to select and exclude certain sources of information. The question that then remains, is: How does a person know if they have exposed themselves to enough *relevant* sources of knowledge?

Lastly, Nguyen argues that while the epistemic bubble and echo chamber mechanisms can theoretically exist separately, they can also happen simultaneously. However, the author provides no more details regarding how we should understand the interplay of these mechanisms. Still, one could expect that for example someone's decision about which sources to let in their bubble is at least partly based on which sources they trust. This raises a last question, namely: How can we understand the interplay between the epistemic bubble-mechanism and the echo-chamber mechanism?

The four points discussed above are all examples of aspects of the conceptualization of echo chambers and epistemic bubbles that I believe have still remained quite open-ended after Nguyen's contribution and which could arguably strongly benefit from a more network-oriented perspective. Precisely that is what the field of network epistemology can help us with. Further on in this thesis, I will show how network epistemology can be a valuable addition to our methodological tool set for the philosophical analysis of echo chambers and epistemic bubbles, both for addressing the gaps briefly touched upon in this chapter and for addressing the existing controversy surrounding the role of digitalization in the formation of echo chambers and epistemic bubbles discussed in chapter five. First, I will introduce the field of epistemic network modelling in the next chapter.

3. Network epistemology

In our current society, the gathering of knowledge and information is a very social matter: People constantly have to rely on the testimony of others, and have to trust and build on information that they get through other people. This social side of epistemology has individual components, such as the judgements an individual has to make when deciding whether to trust a new source or not, or questions about how one should update their beliefs if they get information from another person. A very big part of the current fields of social epistemology is focused on understanding such two-person testimony relations (Sullivan et al., 2020).

However, some components of knowledge acquisition are not based on the properties of the individual person in question, but on the social environment and the network this person is in. For example, if my friends Angie and Bob both told me the same information (e.g. that there exist blue roses) separately, I would treat it as more likely to be true than if only one of them told me this. But perhaps they heard this information about roses both from the same person, a third person named Charles. In such a case, knowledge of the network relation between Angie, Bob, Charles and myself would tell me that I need to treat Angie and Bob's information as if it came from one person, namely as if it only came from Charles. The analysis of the network can thus unveil certain epistemological properties and vulnerabilities (like the one described above²) that the analysis of only the individual person-to-person relations cannot (Sullivan et al., 2020). Network epistemologists analyze such network structures, and showcase how these structures can for example influence the epistemic vulnerability of agents within these networks (Alfano et al., 2018; Sullivan et al., 2020).

In this chapter, I will explain what network epistemology is and how it is relevant for our current endeavor. I will begin with an explanation of the workings of network epistemology, after which I will discuss its relevance for epistemic bubbles and echo chambers.

3.1. What is network epistemology?

Network epistemologists use models to simulate the spreading of beliefs and knowledge. Singer et al. have identified three guiding ideas behind network epistemology, namely:

² This example is based on the example used by Sullivan et al. (2020)

- “1 Beliefs, information, evidence, and other epistemic elements are held by agents, where ‘agents’ is broadly understood to include individuals, but also families, small groups (like a Parent–Teacher Association), scientific lab groups, companies, journals, and other institutions;
- 2 The agents are connected to other agents in a (not necessarily fixed) network structure; and
- 3 The beliefs, information, evidence, and other epistemic elements are shared via network connections with other agents.” (Singer et al., 2022, p. 134)

Network epistemology is a formal approach used in the field of social epistemology, which means that mathematical and/or logical tools are an important part of its methodology (Weisberg, 2021). The field of network epistemology is quite interdisciplinary: not only scholars with a philosophical background, but also scholars from other disciplines like computer sciences, physics, mathematics and economics contribute to research in this field (e.g. Bala & Goyal, 1998; Pansanella, 2024; Weatherall & O’Connor, 2021).

In network epistemology, mathematical models are created which represent agents and their relationships to each other. Figure 1 shows an example of such a network model. An agent can have multiple *neighbors*, which are the other agents they are directly linked to. A network model can be *static*, meaning that the connections between the agents do not change, but it can also be *dynamic*, meaning that through time it can change which agents are connected to which (Singer et al., 2022). The model in figure 1 is an example of a *random* model: there is no clear structure to be found in how the agents are connected to each other, and some agents have more connections than others. There are several other structures possible: for example, a *complete network* is a network in which all agents are connected to all other agents (Zollman, 2007).

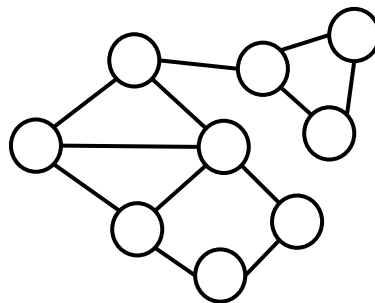


Figure 1: A simple random network model. The circles represent agents, the lines represent connections between agents.

The purpose of epistemological network models is often to explore how certain mechanisms of interest could potentially impact the spread of information (O'Connor et al., 2024), although the visualization of an epistemological network can also mainly serve clarification purposes (such as in Sullivan et al., 2020).

Epistemic network models are almost always run as computer simulations (Weisberg, 2021). In the models, agents are often attributed a certain belief about a proposition p , for example whether the agents believe that eating apples is good for your health. This credence can range between 0 (absolutely certain that *not-p*) and 1 (absolutely certain that p). This thus means that agents with a credence of 0.5 or higher believe (with varying certainty) that p , and agents with credences below 0.5 believe that or *not-p*. In the works of some scholars (e.g. O'Connor & Weatherall, 2018, 2019; Zollman, 2007) the agents' belief is not about p or *not-p* but about whether theory A or B works better. Apart from this, the mechanism is the same.

First, the starting conditions for the model are set out. These include for example the number of agents, the connections between them, or the initial attribution of beliefs about a certain (true) proposition p . Next, the model is ran. A *run* of a model consists out of a certain number of timesteps in which the model's dynamic is executed each in each step (Singer et al., 2022). In these steps, the agents communicate and exchange evidence, arguments or beliefs according to the setup of the model. After communicating with their neighbors, agents then update their credence about p . This communication and belief updating is repeated a certain number of times. A model is often run thousands of times with slightly different starting conditions, after which the results of these runs are statistically analyzed (Singer et al., 2022).

Depending on the research, certain rules are added to the model, such as rules regarding how belief updating is calculated, or how information is shared with neighbors in the model. The way the credence is calculated and updated varies per type of model: In the case of Singer et al. (2019) it is calculated as a weighted sum of the number of reasons an agent has for and against p , with new reasons changing this sum. In a majority of other cases (e.g. Hahn et al., 2024; Olsson, 2013; Weatherall et al., 2020; Zollman, 2007) belief updating is done with the use of a form of statistical belief updating, most often Bayesian. In this form of belief updating, a person's prior credence towards a certain belief is taken into consideration when calculating their new credence after encountering new evidence. Bayes' rule (or a version thereof) is used for this calculation. There is empirical evidence that humans actually use belief updating mechanisms comparable to Bayesian belief updating in real life (Hahn et al., 2024), albeit of course not in the same mechanical and deliberately statistical way.

Sometimes, agents are built to simulate ‘doing research’ themselves, meaning that they acquire certain information about the true state of p in the model (O’Connor & Weatherall, 2018). If we use our apple-example, agents might be thought of to have asked ten people to eat an apple each morning and count the number of people who feel fitter. Such results are then shared with their neighbors in the network (see figure 2 for an example).

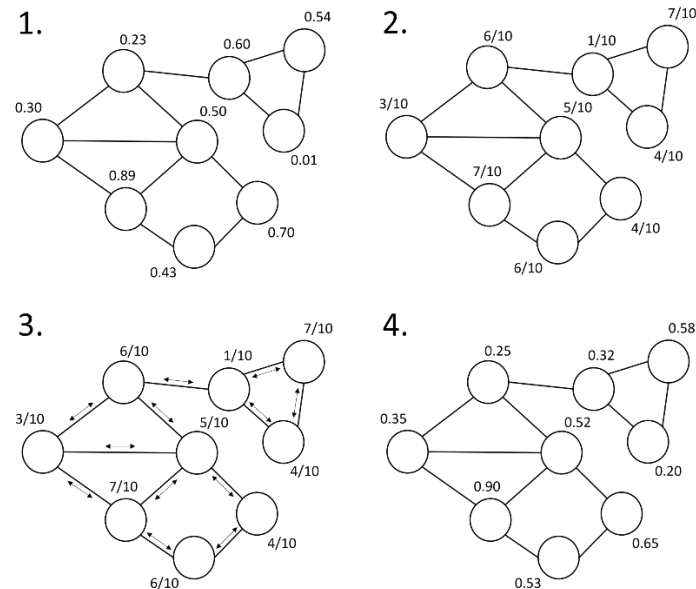


Figure 2: Example of one full cycle in a model run that includes ‘doing research’, in which agents test how many out of ten times a hypothesis yields a positive result (2). These results are shared by the agents with their neighbors (3) and used for belief updating (4)

3.1.1. The use of network epistemology

With the use of epistemic network modeling, two things can be analyzed: 1) how the *structure of a network* influences belief spreading and 2) how certain *behaviors and/or rules* influence belief spreading. The introduction of this chapter already briefly touched upon the first function of network epistemology: the analysis of the influence of *network structure* on belief spreading. One of the researchers that has used network epistemology to this end is Zollman (2007, 2013). He came onto the striking observation that, in some cases, the better the agents in a network are connected, the more likely they are to reach a false consensus and abandon a true belief. This finding has been dubbed the *Zollman effect* (O’Connor & Weatherall, 2019). Sullivan et al. (2020) have also used network epistemology to analyze network structures, namely to gain clarity on how certain epistemic vulnerabilities manifest themselves in networks. They analyzed the network properties of Twitter users to gain insight in these users’ epistemic independence.

Next to the influence of the structure itself, network epistemologists also analyze how certain factors influence how beliefs and knowledge spread through networks. This is done by adding certain rules and dynamics to the model that is used. For example, Singer et. al (2019)

analyzed how limited memory can influence the acquisition of true beliefs by letting agents share reasons for the acceptance or rejection of a certain stance. These agents were modeled in such a way that they were only able to maintain a certain amount of reasons, after which they would throw out one of the reasons in their memory. In certain instances of their model, for example if reasons were kept or thrown out by the agents based on coherence-mindedness (keeping the reasons that cohere best with each other), polarization would occur in the model. Another example of research using network epistemological approaches to analyze the influence of certain rules and behaviors is that of Weatherall et al. (2020), who analyzed the influence of the selective sharing of information on the spreading of true and false beliefs. They did this by adding special agents to the network that represented what they called ‘propagandists’. These propagandists would only share information if it matched with their lobbying goal. Their research showed that such selective sharing of information by propagandists can result in other agents in the model reaching a false conclusion.

3.2. Current research

There is little literature in the field of network epistemology that is directly written about echo chambers and/or epistemic bubbles, with Madsen et al. (2018) and Perfors & Navarro (2019) as notable exceptions. Most literature in the field of network epistemology focuses on *polarization*, which is defined by Singer et al. (2022, p. 137) as “persistent (and sometimes even violent) disagreement between groups, often two”. Two kinds of polarization are of importance for the current discussion, namely *belief polarization* and *affective polarization*. In (inter-group) belief polarization, agents keep holding persistently opposing views that become even more extreme, even after the exchange of information and beliefs with others (Begby, 2024; O’Connor & Weatherall, 2019). In affective polarization, two (or more) subgroups of a society increasingly or persistently harbor more negative feelings to people outside of their group and more positive feelings to the in-group (Munroe, 2023). Most of the literature on epistemic networks focuses predominantly on belief polarization. However, a significant number inadvertently also researches affective polarization in their analysis of trust between agents in their networks.

Many studies in network epistemology try to find answers to why people can come to hold very opposing and stable views in a network even though there is discussion between them: in other words, why polarization occurs in these networks. As discussed above, these researchers often add extra rules or variables to the ‘basic’ model, such as a variable simulating trust in other agents and their testimonies (Hahn et al., 2024; O’Connor & Weatherall, 2018; Olsson, 2013).

Some researchers have added a sense of ‘factuality’ to their model in which there is a ‘fact of the matter’ of what the true value of p is (or whether action A or B is preferable), and analyze what makes that some agents in the network still come to hold radically opposing views even if there is a factually correct answer. Agents in these models are often given the ability to ‘conduct research’ in the way discussed in section 3.1 (Hahn et al., 2024; Madsen et al., 2018; O’Connor & Weatherall, 2018). Other differences in the used models can be found in their structure and size. For example, O’Connor and Weatherall (2018) have used complete networks of 2-20 agents in their research, whereas Madsen et al. (2018) used a scale-free network of 100-1000 agents.

4. Bubbles and chambers in a network

While much of the literature in network epistemology is thus not explicitly linked to the concepts of epistemic bubbles and echo chambers, their findings are still very useful for our discussion. In many cases, this research on polarization can namely also be linked to the mechanisms behind epistemic bubbles and echo chambers as described by Nguyen (2020). Recall from chapter two that the main mechanism behind epistemic bubbles as discussed by Nguyen is exposure to relevant sources. If a person finds themselves in an epistemic bubble, it means that they lack connections to relevant sources. In network epistemology, this is represented by the connections an agent has in the network. If an agent is part of a complete network in which they are connected to all other agents, then they are certainly not trapped in an epistemic bubble. If, however, an agent is only connected to other like-minded agents while there are also disagreeing agents in the network, then this can be argued to represent the occurrence of an epistemic bubble. Research on the effect of network size, structure and/or dynamicity on network polarization can thus also tell us something about epistemic bubbles.

In the case of echo chambers, the most important mechanism is not connectivity, but a high disparity in the level of trust between the in-group and the out-group. If a group of agents in a network only places trust in themselves and radically distrusts all sources with dissimilar views then one could say that these agents form an echo chamber together, especially if their belief does not match empirical evidence or a 'ground truth'. Research that not only includes belief polarization but also affective polarization (i.e. where trust-factors are added to the model) can thus help us understand echo chambers better. Perfors and Navarro (2019) already discuss this link in their work: they name the polarized groups with a high disparity of trust in their work 'echo chambers'.

However, not all research on polarization in network epistemology can be linked to epistemic bubbles or echo chambers. Remember the study by Singer et al. (2019) discussed in section 3.1.1, where a rule simulating limited memory capacity was added to their model. While their study shines a light on a very interesting aspect of how limited agents using rational belief updating can polarize, their analyzed rule does not contain a bubble or chamber mechanism. The studied mechanism is namely not based on either exposure or trust, the two important pillars for bubbles and chambers, but on memory capacities.

In sum, while little research has been conducted in network epistemology that is explicitly linked to epistemic bubbles and echo chambers, a lot of studies can still be used to this end because they are about a phenomenon related to bubbles and chambers, namely polarization.

This does not mean that *all* studies about polarization can be linked to the phenomena discussed in this thesis, but a significant part can, because the polarizing mechanism that is analyzed in these studies is either based on connectivity or on trust. In this chapter, we will have a look at a number of studies in network epistemology that *can* teach us something about epistemic bubbles and echo chambers.

So, let us circle back to epistemic bubbles and echo chambers. As I already stated before, I believe that Nguyen's (2020) conception has helped tremendously in clearing up the definition of both these phenomena. However, some questions still remain. In this chapter, I elaborate on some of my critique on certain aspects of Nguyen's paper and the subsequent questions that were briefly introduced in chapter two. These questions relate to the following aspects: 1) Nguyen's discussion of how one gets into an echo chamber, 2) the intentionality behind the emergence of echo chambers, 3) the easiness of shattering an epistemic bubble, and 4) the interplay between epistemic bubbles and echo chambers.

4.1. Getting into an echo chamber

Let us start with the first one. In his paper, Nguyen argues that people might not always be to blame for ending up in an echo chamber. The example Nguyen gives of when we could speak of possible blamelessness is for a person being raised in an echo chamber. In this specific instance, it is arguably quite easily defensible why such a person is not to blame. Nguyen makes the assumption that it is epistemically reasonable for a child to trust their parents, and I agree with him. In this specific case then, the earliest epistemic contacts of such a person are people that are already in an echo chamber, meaning that the child's first frame of reference to test other information against is the echo chamber's frame of reference in which information of outsiders is deemed severely untrustworthy. Nguyen argues that even if this person would later as a teenager or grownup act in epistemically virtuous ways (such as actively acquiring new information and testing new claims against all evidence and all beliefs they have gathered previously and then deciding whether to accept them or not), they could still remain trapped in the echo chamber due to its problematic mechanisms discussed before.

However, being born and raised in an echo chamber is a very specific example of when one might not be blameworthy of being in an echo chamber. Even though Nguyen argues that there are "many cases" (2020, p. 154) in which a person might be blameless for ending up in an echo chamber, the above example is the only one he gives. In other parts, the author mostly points to the "slow accumulation of minor mistakes" (Nguyen, 2020, p. 151) through epistemic

vices that might lead one to enter an echo chamber, or even to quite strong epistemically vicious acts like purposefully entering one to get a sense of belonging. These are arguably both clear examples of when someone *is* to blame for ending up in an echo chamber.

I would argue that while the example of being raised in an echo chamber is a strong one, it is also very specific and the arguments for the person's blamelessness (trusting one's parents as a child) are not easily transferrable to other cases. Still, it seems like there is quite a significant group of people that enter echo chambers later in life. Take for example the anti-vaxx conspiracy theorists during the covid crisis: these people did not get born into that conspiracy. Another example is that Nguyen himself even argues that we might be able to speak of a "vast partisan echo chamber" (2020, p. 150) on the right-wing side of the current (American) political landscape. While all people might commit small epistemic vices throughout their daily life, it seems to me to be quite a strong statement to argue that all these people on the right-wing side of the political spectrum have acted so epistemically viciously that they are all mostly themselves to blame for having entered an echo chamber. Therefore, one of the questions that still remains after Nguyen's contribution is the following:

Q1: Are there any other factors aside from being raised in an echo chamber that (partly) exempt people from blame for ending up in an echo chamber?

When we look at the literature from network epistemology, the answer to this question is overwhelmingly: Yes. Multiple studies that have used rational and idealized Bayesian agents found that their agents could still end up in what we could argue are echo chambers. Take the work on polarization by O'Connor and Weatherall (2018). These researchers simulated a group of scientists trying to find out whether action A or B yields a better payoff. All agents start off with initial credences about which action is better. These initial credences are randomly distributed around a slight preference for A, of which they already know the success rate ($P_a = 0.5$). The pay-off rate of action B is initially unknown, but in reality it is slightly better than that of A ($0.501 \leq P_b \leq 0.8$). All agents in the model perform the action they themselves prefer n times, and in addition get information about the number of successes found by their neighbors. Since the pay-off rate of A is already known, only results of the performance of action B are shared.

This idea is in line with the models of Zollman (2007, 2013). However, there is one important heuristic added to the model: the evidence gotten from others is treated as uncertain. *How* uncertain this evidence is deemed is based on how distant the beliefs of this other agent are from their own. In other words, the more their neighbor's credence about the superiority of A over

B is different from their own, the less their evidence is trusted. With the addition of this heuristic, O'Connor & Weatherall (2018) thus create (perhaps inadvertently) the possibility to analyze the interplay between potential affective and belief polarization.

The authors tested three variants of their model: In version one, the agents would trust the evidence of agents with a larger belief distance less, but there would always remain a certain level of trust. In the second version an agent would stop listening to a certain neighbor's evidence if the belief distance between them would exceed a certain threshold, and in the third version an agent would actively update their beliefs in the other direction if the belief distance to their neighbor exceeded the threshold. Thus: in version one, if an agent differs a lot in belief from their neighbor, the agent is more cautious but always slightly trusting, in version two they stop listening to the evidence, and in version three the agent distrusts the evidence so much that they think the neighbor is lying and subsequently update their belief according to the disagreement-reinforcement mechanism described by Nguyen (2020) and discussed in chapter two.

O'Connor & Weatherall found that with the addition of this heuristic, polarization is very likely to occur in version two and three of their model, meaning that a subgroup of the population is stably holding the false belief that A is better while not trusting the other agents anymore. In other words, this heuristic leads to the formation of an echo chamber. This is especially true for the models of version three, where a disagreement-reinforcement mechanism is active. In their first version it is not possible to have a completely polarized outcome, due to the fact that an agent never really stops trusting anyone. However, the authors did even in these cases observe what they call "transient polarization" (2018, p. 868) in which two subgroups are highly polarized for a substantial amount of time. Moreover, it takes the agents in this version significantly longer to get to a consensus than in the control version where all agents trust each other equally, "and in the meantime there is a potentially very long period during which some portion of the community is mistrustful of an emerging consensus" (O'Connor & Weatherall, 2018, p. 868).

The results of these three versions together show that the heuristic of trusting a person more if their beliefs are closer to one's own beliefs is possibly enough to create echo chambers, even if the agents in the model are otherwise honest, rational, empirically testing their theories against the real world, and only have as their aim to discover the true best action. Other authors with slight variations in their trust mechanism have found similar results, making clear that O'Connor and Weatherall's findings are not just a product of their specific model settings. Both Madsen et al. (2018) and Olsson (2013) have found that adding a similar simulation of trust to their epistemic model increased polarization (although Madsen et al. call this factor 'open-mindedness' rather than trust). In the model of Perfors & Navarro (2019), who also encounter

similar results, the agents check whether the data provided by the other is likely according to their own data and their idea of how big the variance of data in the population is. In addition, these researchers found the unsurprising result that the less ‘open-minded’ the agents in the initial population are, the more likely they are to polarize. These results found by researchers using a variety of methods thus seem to indicate that such a heuristic of trust is in theory enough to create echo chambers in these networks, even with completely rational and honest agents.

This heuristic actually reminds us of the position Nguyen’s teenager is in: the agents in these models distrust others that do not have the same opinion as them. In this sense, the trust heuristic described here actually looks like a mild version of the precise mechanism behind echo chambers that is discussed by Nguyen (2020), namely the creation of a (high) disparity of trust between in-group and out-group members. Can we then really speak of a valid and rational heuristic of trust?

The answer to this question lies in the word ‘mild’. In his own article, Nguyen (2020) already argues that the mechanisms behind echo chambers “function parasitically” (2020, p. 149) on healthy epistemic practices, like the justifiable use of heuristics in daily life. The author refers for example to the justifiability and the need to resort to heuristics like litmus tests in certain circumstances to be able to make sense of our hyper-specialized world. While O’Connor and Weatherall (2018) themselves do not claim that the heuristic they introduced in their model is individually rational, they do think that the heuristic is justifiable and, in a sense, in practice even essential. Scientists (and other people) need to make judgements about how reliable they think a peer’s evidence is. I would argue that the heuristic as applied by O’Connor and Weatherall (2018) is not directly a parasitic mechanism like in the case of echo chambers, but is rather an example of such of a more mild variant that is actually justified and could be expected from an epistemically well-motivated agent.

One author that provides good insight in why this behavior might actually be considered *rational* is Begby (2024). He argues that people need to have some methods for deciding whether someone can be considered a peer in information gathering or not. He argues that it is quite common sense that one should not update their beliefs based on the beliefs of others if they have reason to think that the other person is misinformed or incompetent, and can thus not be considered a peer. Most people will feel like they have a reason for their own belief in a certain fact. They do not have to be absolute experts in the respective topic for this: people might have done a little research and have drawn certain conclusions based on it that they find logical. If one has such a “reasonable starting confidence” (2024, p. 527) in their own beliefs, then it is only logical to be more inclined to define someone only as a peer to yourself if they share your

judgement. I, for example, have learned at school that eating fruit is good for you and that they contain many essential vitamins. This has led me to believe that apples are good for you. If I meet someone who tells me that apples are actually very poisonous and are only promoted because the government tries to make humans weaker, I have quite good a reason to lower my trust in this person's judgement: it would even be rational to do so.

One way in which the teenager differs from most other people is in their 'starting belief' compared to the bigger epistemic network they are part of: most people have a belief (credence) that is relatively similar to most other agents in the network (the rest of the 'population'). The teenager could be represented as an agent with radically different starting credences than most others in the network, apart from maybe a number of their neighbors (e.g. their parents) that also have radically different credences than most of the population. In most models discussed above, agents were used that, in line with a certain distribution, started out with different initial credences towards p . While I would argue that slight variation in initial beliefs resembles a healthy and realistic society, one could argue under a very crude interpretation of these initializations that in some sense for some agents this resembles the initial state of Nguyen's teenager into the world. What if everyone starts with the same belief in p ?

In the model created by Hahn et al. (2024), all agents start out agnostic about the truth of proposition p , meaning that they all believe that the chances of p or *not- p* are equally likely. These researchers found that polarization can still happen in these cases, especially if the evidence for the true state of p is insufficiently clear. In other words, even if everyone starts out with the same initial agnostic stance towards a certain fact, people can still end up in echo chambers purely based on their coincidental position in the network and the addition of a simple trust heuristic. The conclusions they draw based on their findings is that "those 'conspiracy theorists' in our simulated worlds who strongly end up believing a falsehood, end up in this position through sheer bad luck." (Hahn et al., 2024, p. 13). The authors also found that the bigger the network, the bigger the chance that a subgroup together forms an echo chamber.

To sum up, the literature in network epistemology clearly points to the idea that even rational agents that are motivated to find the truth, that are honest and that are not initially put in echo chamber conditions can come to be trapped in an echo chamber. This does not need to happen through a "slow accumulation of minor mistakes" (Nguyen, 2020, p. 151) or even more radical epistemically vicious behavior: reliance on what seem to be justifiable and rational heuristics of trust, and unfortunately in some cases also a form of 'sheer bad luck', seem to be enough. One softening circumstance seems to be that if agents are more open-minded of their

neighbors' data and are less prone to distrust, they are less likely to polarize and form echo chambers.

4.2. The intentionality of echo chambers

Let us move on to the second aspect: intentionality. In his paper, Nguyen (2020) argues that while echo chambers could in theory come into being unintentionally, they are almost always set up intentionally by its makers to control their audience and to increase their power. Here, Nguyen follows Jamieson & Capella (2008), who conducted an analysis of intentional echo chamber formation around charismatic personalities. Because of Nguyen's dependence on Jamieson and Capella (2008) as his basis, one can understand why he argues that charismatic leaders are important for the setup of an echo chamber. While I agree with Nguyen that some echo chambers might in fact be set up intentionally, I disagree that it is *most likely* that echo chambers are set up intentionally by certain charismatic personalities.

One author that also questions this need of an intentional setup is Munroe (2023). He argues that the Pizzagate conspiracy theory, in which Democrat politicians are said to run a pedophile network from the basement of a certain pizza store, is a prime example of how an echo chamber can arise and flourish without any leadership. This specific echo chamber, built around a conspiracy theory, is said to have originated on the extremist website *4chan* and spread further through mostly anonymous and small social media accounts. It has even been argued that pizzagate started as a sort of online detective game that over time started to actually be believed (Bleakley, 2023). These people were not prepared by certain leaders to distrust anyone claiming pizzagate was false; instead, as Munroe puts it, "the members collectively and dynamically discredit and distrust outside sources as Pizzagate evolves" (2023, p. 7). This sounds to me exactly like an echo chamber as explained by Nguyen (2020): the only thing missing is the intentional setup.

Another example of a bottom-up fake news story is the story of the pet-eating Haitian immigrants of Springfield. During the U.S. election campaign in 2024 in a debate with Harris, Trump made the flagrantly false claim that Haitian immigrants in Springfield capture and kill the cats and dogs of other inhabitants to eat them. This story originated from a Facebook post of an elderly woman in Springfield who misinterpreted a claim made by her neighbor about a missing cat and subsequently posted the story about cat-eating immigrants on Facebook. It was then picked up by numerous other people and reposted, after which it spread rapidly (Lozano, 2024). While this story might have ended up at a leading figure, it actually followed a down-up trajectory

instead of the other way around. It thus did not adhere to the top-down mechanism that would be expected of (false) information created by intentionally set-up echo chambers, although the fake story did get a boost after its mention by Trump (Lozano, 2024).

The last example is an example of a fake news story, not necessarily of an echo chamber. Still, it mostly ended up being believed and spread in a certain part of the American population, one that apparently did not care for or did not trust fact-checkers arguing that the story was fake. The above examples show that the disparity of trust that characterizes echo chambers need not be created intentionally by a charismatic personality, and I question therefore Nguyen's statement that it is most likely for echo chambers to be set up in this way rather than unintentionally. This begs the following question:

Q2: Can we explain the occurrence and arising of echo chambers that come into being unintentionally?

The findings discussed in the previous section also point to the answer of this question. The analyses of polarization in network epistemology show that echo chamber can also form without any malicious intent or setup from above. In all the cases discussed before, the agents in the models were acting as rational and virtuous as possible: the addition of malicious agents was not necessary for the arising of the echo chambers.

This of course does not mean that such malicious agents could not speed up the process. In another study, Weatherall et al. (2020) show that agents that act like lobbyists can have a massive influence on the distribution of knowledge. Still, many of the discussed studies in network epistemology (by e.g. Madsen et al., 2018; O'Connor & Weatherall, 2018; Perfors & Navarro, 2019) show that such malicious agents are not necessary at all. Rather, the unintentional arising of echo chambers can on the system level be explained by the individual use of justifiable heuristics like placing a higher trust in individuals that have beliefs that are more similar to one's own beliefs.

4.3. Shattering an epistemic bubble

Now we turn to the aspect of shattering an epistemic bubble. Recall from chapter two that Nguyen argues that it is relatively easy to shatter an epistemic bubble. According to the author, the only thing one has to do is to expose oneself to a wider variety of relevant sources. However, there are a number of reasons to argue that it might be harder than “mere exposure” for one to get in contact with all relevant voices.

First of all, Pollock (2024) argues that the *shattering* of an epistemic bubble is not the right wording of the concept, since there is always more information to be incorporated and thus always some way in which one’s bubble is not really complete or shattered at all. Rather, Pollock argues that we should talk of the *expansion* of one’s bubble. This is what he calls an “elastic conception” (2024, p. 456) of epistemic bubbles, in which they are “slowly and incrementally expanded to improve our (inevitably partial) perspective on the world” (2024, p. 456).

Still, one could argue that a person does not have to stand in contact with every nuanced and detailed opinion of others to be seen as free from a bubble: as long as they are exposed to a wide enough variety of relevant sources, could that not be considered enough? But even with this in mind, there is another problem with Nguyen’s solution: how does a person know if they have exposed themselves to enough *relevant* sources of knowledge? Nguyen himself already states that “the world is overstuffed with sources of information, many of them terrible” (2020, p. 143). He argues that it is therefore good and even necessary to select and exclude certain sources of information. But how does one decide which sources they can rightfully exclude? Would you not decide this partly based on the information that you already have, and that is thus already part of your bubble? In addition, it might be difficult to adequately value a source of information, for example because of the problem of *bootstrapped corroboration*. In bootstrapped corroboration, a person (perhaps unintentionally) treats dependent sources of certain information as independent, leading to an over-confidence in a certain belief (Nguyen, 2020). However, since this might be unconscious, a person could not know that they should seek out new sources to balance out their unbalanced belief.

It seems that there might be indications that it could be harder to escape an epistemic bubble than is argued by Nguyen (2020), even based on his own description of the problem of bootstrapped corroboration. Based on the scope of this paper, this leads me to the following question:

Q3: Can network modeling help us understand why it might be harder to shatter an epistemic bubble than through ‘mere exposure’?

For the answer to this question, we need to turn the work by Sullivan et al. (2020). The authors argue in their paper that within social epistemology there has been too little focus on the influence of a person's network on their *epistemic vulnerability*, i.e. the threats to the quality and robustness of their knowledge gathering process. To fill this gap, the researchers created a metric that can be used in network settings to calculate the state of an agent's *epistemic position* and their *epistemic security* in that network. With the state of one's epistemic position, Sullivan et al. refer to how epistemically vulnerable the agent is within their current static place in a network. With epistemic security, the authors refer to the changeability of this position, both for better and for worse. Sullivan et al. argue that an agent's epistemic position is based on four components, namely 1) the number of sources the agent is in direct contact with (i.e. the number of neighbors), 2) the independence of these neighbors from each other, 3) the diversity of the viewpoints of their neighbors and 4) their reliability. Sullivan et al. have used the first three components for a calculation of one's epistemic position. The authors argue that the analysis of "these network structures can help us to diagnose when someone is in a problematic filter bubble, along with whether the surrounding network affords them more epistemic threats than opportunities for improving their position" (2020, p. 743). Linking the first two components they identified together can namely help to analyze whether someone has incorporated a wide big enough group of *relevant* sources (i.e. whether one has enough neighbors that are independent from each other), while their diversity metric can help tell whether the *variety* of these relevant sources is big enough (i.e. whether these sources expose the agent to enough different standpoints). Together, one could indeed argue like Sullivan et al. (2020) do that these metrics give us an (albeit simplistic) idea of whether one is part of an epistemic bubble or not.

In addition, Sullivan et al. (2020) used their metric empirically, namely to analyze the epistemic position of individual Twitter users engaging in the topic of vaccine safety. They found that only 15% of the users in their network inhabit a good epistemic position, while the rest inhabit a poor (51%) or even very poor (34%) epistemic position. While the authors concede that their measurement is inclined to be more conservative than generous in the calculation of one's epistemic position, these results arguably still tell us something about how hard it seems to be to escape epistemic vulnerability, and indirectly to escape getting trapped in a bubble. For the users with a very poor epistemic position, the biggest problem for their position was related to the interdependency of their sources. These agents are at a high risk for bootstrapped corroboration, which was also identified by Nguyen (2020) as a threat to how well a person can adequately value the information in their own bubble.

Of course, we do not know whether these Twitter users want to or have ever tried to shatter (or expand) their bubble. Still, one of the goals of this thesis is to show how it is not necessary for people to be epistemically vicious when ending up in a chamber of bubble. The analysis by Sullivan et al. (2020) at least shows us theoretically how one can unwittingly be epistemically vulnerable, and it shows us empirically that a large majority of people have actually ended up in such a position.

4.4. The interplay of bubbles and chambers

The last aspect that will be discussed in this chapter is the interplay between bubbles and chambers. First, I want to state again that I do agree with Nguyen (2020) that there really seem to be two different mechanisms at play here: there is a difference between being exposed to a source and actually taking up and trusting what that source has to say. The idealized distinction between an epistemic bubble as being related to source exposure and an echo chamber as related to the trust in such a source is very useful for (among other things) questioning the validity of current academic research into these phenomena.

However, the underlying mechanisms of epistemic bubbles and echo chambers also seem to be inherently related to each other in some sense and also seem to potentially occur together in practice. For example, one could argue that someone's decision about which sources to add to their bubble is at least partly based on which sources they trust. I do not listen to Fox News, for example, because I do not see them as a trustworthy news outlet. Nguyen (2020) himself argues as well that while these mechanisms can theoretically exist separately, they can also happen simultaneously. This leads to the last question posed in this chapter:

Q4: How can we understand the interplay between the epistemic bubble mechanism and the echo chamber mechanism?

There are multiple articles in the literature of network epistemology in which there is an interplay present between epistemic bubbles and echo chambers. One example of a study where this is the case is that Madsen et al. (2018). The networks that these authors analyze have a relatively large size compared to those of other studies, namely between 50 and 1000 agents. It would never be feasible for one person to listen to the whole network and to take the opinion of almost 1000 others into account when forming their own opinion. Therefore, instead of a complete network, the researchers used *scale-free* and random networks. In scale-free networks, the

number of connections the agents have follows a power law distribution, meaning that most agents have a limited number of connections while a smaller number of agents have a relatively high number of connections (Madsen et al., 2018). The authors argue that real social networks also often have this structure.

Madsen et al. (2018) created a network with a search parameter, in which an agent ‘searches’ across the available network for opinions. The agents only trust others if their opinion falls within their own uncertainty boundaries about their belief, which is based on the variance v of different pieces of evidence they have encountered and taken up times a parameter p for open-mindedness. If a neighbor within their network is no longer trusted, they ‘disconnect’ from this neighbor and go searching for a different agent that does fall within their range of belief. The agent thus ‘stops listening’ to other agents if they are not trusted anymore.

I would argue that the model of Madsen et al. (2018) is a prime example of how the mechanisms behind epistemic bubbles and echo chambers could intertwine. In their model, if an agent does not *trust* a source anymore, they stop *exposing* themselves to that source as well. The authors themselves call this the *pruning* of one’s sources, leading to the formation of relatively disconnected clusters consisting of tightly connected agents that primarily listen to each other. This then leads to agents that are (due to their own pruning) only getting exposed to beliefs that are in line with their own, leading to unwarranted levels of confidence in these beliefs, which then only amplifies the potential group polarization even more (Boyd, 2019; Madsen et al., 2018).

There is a form of bootstrapped corroboration at play in the phenomenon described above, in which the individual agents of an agreeing group illegitimately keep raising their credence in the belief that p because of the testimony of others that are all part of the selective group that all also believe that p . Boyd (2019) has called this phenomenon *groupstrapping*, but other scholars like Olsson (2013) and Begby (2024) regard this as another version of belief polarization.

Next to the (at least semi-)intentional selective exposure mechanism discussed here, the work by Hahn et al. (2024) discussed in section 4.1 gives some insights on how completely *unintentional* selective exposure can amplify echo chamber mechanisms. Remember that all agents in the researchers’ models start out agnostic, meaning that they all think p and *not- p* are equally likely, and they all trust each other equally. These starting conditions mean that no initial echo chamber mechanism is present, since there is no initial disparity of trust (or credence) between the agents. However, once all agents start to gather information by ‘doing research’ (as explained in chapter three) and communicating this research to their neighbors, Hahn et al.

actually did observe the formation of echo chambers. The reason this could happen was because some agents were, by statistical chance, more exposed to findings from themselves and their neighbors that supported the false belief that *not-p*, even though (most) others were more exposed to findings that *p*. After these findings, those specific agents start trusting neighbors that think that *not-p* slightly more, due to their own findings and the early findings of some of their neighbors. This sets off a loop of growing trust disparity as encountered earlier.

In the cases discussed here, echo chamber formation is only possible due to coincidental differences in belief *exposure* between agents, created by their positions in the network they are part of. In other words, in the model of Hahn et al. (2024), epistemic bubble mechanisms and echo chamber mechanisms enhance each other. They show how the interplay of both bubble and chamber mechanisms can, in some cases, be extra effective, even though this interplay might not always be necessary.

Let me stress again that these cases explained here based on models from Madsen et al. (2018) and Hahn et al. (2024) are examples for how epistemic bubbles and echo chambers *could* interact. This does not mean that the mechanisms always *have to* interact or rely on each other. As Nguyen (2020) explained, epistemic bubbles and echo chambers can also function separately from each other.

One thing that the above discussion of these four gaps in Nguyen's (2020) conception arguably has shown is that network epistemology can provide us with a new and more system-oriented perspective on social epistemological phenomena like epistemic bubbles and echo chambers. However, there is one scholarly debate that is not extensively discussed by Nguyen but which is of importance in the current academic discussion surrounding epistemic bubbles and echo chambers: their digital prevalence. In the next chapter, I will further dive into this discussion.

5. The effect of digitalization

Multiple scholars have argued that digitalization plays a significant role in a growing concern for increased polarization (Bennett & Iyengar, 2008) and/or an increase in both echo chambers and epistemic bubbles (Bail et al., 2018; Bakshy et al., 2015; Sunstein, 2018). For example, Sunstein (2018) argues that the vast number of perspectives available through the internet allows people to cherry pick sources that only align with what they already believe, leading them to build their own echo chambers³ around them and increasing polarization. Moreover, multiple empirical researchers have found evidence for echo chambers or epistemic bubbles on the internet (Mønsted & Lehmann, 2022; Terren & Borge, 2021).

However, other scholars have criticized the above perspectives, arguing that new technology has not had this impact. For example, Bruns (2019) has argued that the effect of technology on epistemic bubbles and echo chambers has been grossly overestimated. The author even argues that the prevalence of both bubbles and chambers has been overestimated in general, arguing that much of the disagreement found online can also be attributed to healthy political disagreement, and that even if echo chambers do exist, their prevalence has been overstated due to methodological cherry picking in science. An example of methodological cherry picking, according to Bruns, is the analysis of only topics that are highly polarized in the first place, such as climate change or vaccination stances. Bruns even calls the echo chamber the “dumbest metaphor on the internet” (2019, p. 8). Other, more empirical researchers have also criticized the alleged role of digitalization on epistemic bubbles and echo chambers, with some even arguing that online sources actually decrease the percentage of people in an echo chamber, because online sources actually make for a bigger diversity in ideological exposure⁴ (Arguedas et al., 2022; Masip et al., 2020).

The above examples show the scientific disagreement there is about the role of digitalization for our two phenomena of interest in this thesis. The questions that arise are the following: Firstly, how can we understand where this dispute between scholars about the impact of digitalization comes from, and secondly, can network epistemology help us to understand the possible impact of digitalization on epistemic bubbles and echo chambers better?

³ Notice that the mechanism Sunstein refers to here is actually an epistemic bubble mechanism, not an echo chamber mechanism.

⁴ Here we see the same conflation again.

5.1. Conflation, again

As has already been briefly explained in chapter two, the concepts of epistemic (or filter) bubbles and echo chambers have often been conflated with each other. Many scholars have defined echo chambers similarly to the way epistemic bubbles have been defined in this paper: based on exposure and communication rather than on (radical) trust discrepancies. An example of this is the article of Mønsted & Lehmann (2022), who argue to have found evidence in support of echo chambers. They state that the vaccine discourse on Twitter is highly polarized, and that users on the extreme ends formed “relatively disjoint ‘epistemic echo chambers’ which imply that members of the two groups of users rarely interact, and in which users experience highly dissimilar ‘information landscapes’ depending on their stance” (p. 8). Notice that their use of ‘echo chamber’ in this part is used to refer to the low level of interaction and exposure to similar sources, which were by Nguyen (2020) actually identified as properties assigned to epistemic bubbles instead of echo chambers.

This same mixing of the two concepts is also found in articles arguing that supposed ‘echo chambers’ actually make for a *bigger* diversity in exposure (e.g. Arguedas et al., 2022; Masip et al., 2020). These findings namely again say something about *exposure* to different beliefs, not about the *trust* in these beliefs. The same conflation is also found in most scientific reviews about empirical evidence for echo chambers (e.g. Hartmann et al., 2025; Mahmoudi et al., 2024; Terren & Borge, 2021). In other words, all these studies tell us nothing about how the exposed content is actually received and whether opposing content is actually taken into consideration or just gets dismissed, while we have seen before with the work of for example O’Connor and Weatherall (2018) how much impact the incorporation of these aspects has on the outcome of belief formation models.

So, it seems that most empirical literature that claims to be written about echo chambers is not written about echo chambers at all. Instead, these papers describe mechanisms linked to *epistemic bubbles*. This is also the case for literature discussing the potential effects of personalized algorithms in relation to ‘echo chambers’ (e.g. Fletcher & Nielsen, 2018; Wang & Qian, 2021): Such research has to do with the exposure of individuals to certain content, not directly with their trust in any content they are exposed to. This makes it very hard to say anything empirically about (digital) echo chambers. The fact that much of the research on these two phenomena is conflated is a big problem according to Nguyen (2020). He argues that, since most studies that purportedly analyze echo chambers actually analyze epistemic bubbles, their conclusions about echo chambers are invalid. Nguyen even argues that “in fact, echo chambers should hope that their members are exposed to media from the outside; if the right disagreement

reinforcement mechanisms are in place, that exposure will only reinforce the echo chambers' members' allegiance" (2020, p. 153).

There is another problem with the empirical literature about these concepts: the conclusions different scholars draw based on roughly the same findings are widely dissimilar. This difference is mostly found in what degree of exposure to counter-belief evidence is believed to be enough to no longer talk of an epistemic bubble.⁵ Does an agent need to have completely no exposure to different-minded sources at all before we can say they are in an epistemic bubble, or is it enough if agents are exposed disproportionately more to same-belief sources than different-belief sources?

Bruns (2019) argues that there would have to be (almost) no links at all with different-minded sources before one can really speak of an epistemic bubble. The author argues that it is part of normal and healthy epistemic interaction to be more connected to like-minded than to different-minded individuals, and that this has always been the case, even in the pre-digital era. As an example of counter-evidence for the existence of epistemic bubbles, Bruns references a study on the American blogosphere in which links across the Republican-Democrat spectrum are analyzed (Adamic & Glance, 2005). These researchers found that while blogs were more likely to contain links to blogs from the same side of the political divide, there were also some links across. Bruns (2019) uses this research as one of the examples to show that there is indeed exposure to different ideologies (even if this linking to different-minded blogs was just to criticize them), since there would otherwise be no such links. Other researchers that have drawn similar conclusions based on similar findings are for example Masip et al. (2020).

However, other authors that had similar findings, namely that people were mostly clustered around others with similar beliefs but with still a few links left to persons from different-minded groups, actually saw this as evidence *in favor* of the existence of epistemic bubbles online (Mahmoudi et al., 2024; Mønsted & Lehmann, 2022). These authors believe that the encountered disproportionality of the number of connections to like-minded agents versus different-minded agents was high enough to be able to talk about epistemic bubbles. These differences between authors make it even harder to correctly interpret research on epistemic bubbles in a consistent manner.

Moreover, evidence in favor of the existence of online epistemic bubbles does not mean necessarily that they are *more prevalent* online than offline, which is a crucial aspect of the questions posed in this chapter. Just their online existence could just as well mean that these

⁵ From here on, I will refer to research on exposure and communication to different sources as research on epistemic bubbles, even if the researchers themselves arguably 'mislabelled' their research as being about echo chambers instead of about epistemic bubbles.

epistemic structures exist to the same extent online, which would mean that digitalization does not necessarily impact their existence nor has made it worse than before. Only a few studies have compared online information-sharing to offline-information sharing. These studies found that the use of social media actually had a positive impact on content diversity, meaning that individuals that were more active on social media encountered more different-minded news than those who used it less (Fletcher & Nielsen, 2018; Masip et al., 2020).

It is clear that we can say very little empirically about echo chambers, since most of the literature actually writes about epistemic bubbles. While the conclusions these researchers draw are mixed, there does seem to be an (albeit fragile) scientific consensus that agents are exposed to more similar beliefs than dissimilar beliefs online, and that these agents form clusters. There also seems to be a consensus that agents in these clusters do have at least some exposure to different views, albeit sometimes very little. In addition, while the disagreement about the definitions of echo chambers and epistemic bubbles makes it difficult to draw general conclusions, there does seem to be some evidence that the use of online environments actually leads to a bigger diversity in content exposure than the use of only offline environments (Fletcher & Nielsen, 2018).

In sum, it seems that current empirical research on the existence of epistemic bubbles online yields mixed results, while at the same time there also seem to be some indications that digitalization actually has a positive impact on decreasing the prevalence of epistemic bubbles. Even though the evidence for this last notion is not overwhelmingly clear, let us, for the sake of the argument, follow these authors in the notion that digitalization has not led to more epistemic bubbles. Even then, it is still completely unclear what this means for the link between digitalization and *echo chambers*, which is according to Nguyen (2020) a much bigger threat to our current democratic processes than epistemic bubbles. The question that is thus still left is the following: is it possible for digitalization to influence the prevalence of echo chambers, even if it does not affect (or even negatively affects) the prevalence of epistemic bubbles?

Since there is, to my knowledge, no empirical data that has analyzed echo chambers as defined in this thesis, I will turn to network epistemology again to see if the analyses there could give us some theories about the arising of echo chambers in digital environments. In the upcoming section, I am going to show with the help of research in network epistemology that, even if we would side with researchers that argue that digitalization actually increases diversity in exposure (and thus decreases the prevalence of epistemic bubbles), digitalization could possibly still lead to an increase in the prevalence of echo chambers.

5.2. The downsides of big networks

The internet has radically changed how we look for information and connect with others. First of all, it has made it possible to expand our network much more than before, making it possible to connect to as many people as we would like instead of only to the people in our own neighborhood. We are no longer all dependent on the same encyclopedia or the same newspaper for information as the people around us: Because of the internet, we are able to grow our network much further, and to make use of much more sources of information than before. In addition, people are not only able to have a bigger network, but also to be much more targeted in their search for information (Sunstein, 2018). So, to analyze the impact of digitalization, we must analyze the impact of the properties of epistemic networks that are different in digital societies compared to physical ones.

Let us first start with size. One of the articles in which the impact of network size is discussed is that of Hahn et al. (2024). The authors used a small-world network structure in their model, which they argue is similar to the structure of actual online networks on social media platforms like Facebook. Hahn et al. found that the larger the network is, the higher the percentage of agents that ended up polarized and believing the wrong belief. The authors argue that “being part of a social network fosters polarisation in our simulations and increasing the size of that network increases the rate at which societal polarisation is observed” (2024, p. 16). Madsen et al. (2018) found similar results: the authors studied relatively large networks of up to 1.000 agents, and also found that the formation of echo chambers in their model was more likely to occur with the increase of the number of agents in their model.

What makes the model of Madsen et al. (2018) even more interesting for our current endeavor is that they made use of a *search parameter*. Remember that in their network, an agent ‘searches’ across their network for people with similar beliefs, and only listens to these other agents if their belief is enough in line with their own. The authors ran their model using a range of different search parameters, representing how big the part of the network was that the agents could search. The authors found that the formation of echo chambers in their model was not only more likely to occur with the increase of the number of agents in the model, but also with the increase of the search parameter. They argue that this is probably the case due to a higher chance for more closed-minded agents with initial extreme beliefs to be able to find other agents with similarly extreme beliefs, since their pool of reachable agents is simply bigger. In smaller networks, Madsen et al. argue, these more extreme agents are more often “‘starved’ of sufficient numbers of like-minded agents, and therefore their belief confidence is somewhat stymied” (Madsen et al., 2018, p. 2). In other words, extreme agents that are able to search across a bigger

number of sources are more likely to be able to find another like-minded extremist, leading to an (unjustly) higher confidence in their own beliefs.

Some believed that the internet would have led to an “epistemic democracy” (Waisbord, 2018, p. 1870), in which the internet would flatten existing hierarchical structures of knowledge production and would bring more equality for all agents wanting to participate. But what the results discussed above show is that the democratizing forces of the internet have had different consequences on the systemic level than one might have expected. On a personal level, the possibility to have a bigger and further-reaching network indeed mostly seems to grant people access to much more information than before, giving them the possibility to make more informed inferences about their belief. However, once a trust-variable is added, this reasoning does not seem to hold on the systemic level. In fact, the overall performance of the epistemic network only worsens with an increase in size, meaning that relatively more agents end up polarized and in an echo chamber. Hahn et al. even state that

“the clear effects of network size in our simulations underscore the importance of supra-individual, systems level variables. Specifically, those results suggest that the prevalence of conspiracy theories and misinformation can rise simply because of changes to the effective size of our everyday information networks in ways that the early enthusiastic reception of the internet and social media could not have anticipated.” (2024, p. 19)

These findings seem to support Nguyen’s (2020) argument that the biggest problem for online communication might not be epistemic bubbles, but echo chambers. These models also clearly show what has been lacking from empirical research on these epistemic phenomena: the factor of trust. All empirical studies discussed above have focused on the exposure to different content and different sources, but not on how these different sources actually have been perceived and to what extent these sources are trusted differently. The models created in network epistemology show that this empirical measuring of trust could potentially lead to very interesting and new insights into why people are still polarizing.

6. Conclusions and interventions

In this thesis, I have tried to show how network epistemology can help us to understand the social epistemological concepts of epistemic bubbles and echo chambers better. The processes of information gathering and belief formation have an individual, but also a social and system-level component. This system-level component has long been relatively under-studied. I hope to have shown why filling up this gap is a worthwhile and relevant endeavor. Analyzing social epistemological concepts on a system-based level can significantly increase our understanding of them, and help us answer questions that cannot be answered by considering psychological or individual-level perspectives alone.

Aside from this note, there are several main takeaways that I hope readers have understood from this discussion.

Firstly, it is very important to distinguish between echo chambers and epistemic bubbles, to research them separately, and to do so in a scientifically valid manner. All empirical studies discussed in this paper that allegedly analyzed the prevalence of echo chambers focused on the exposure to different content and different sources, not on how these different sources actually have been perceived and to what extent these sources are trusted differently. In other words, these studies researched the prevalence of epistemic bubbles, not of echo chambers. This takeaway was already one of the main messages of Nguyen (2020) in his own article, but the findings discussed here only strengthen the author's point that the mechanisms behind epistemic bubbles and echo chambers are different and can drive polarization separately from each other. For example, the results from the network epistemological models discussed in chapter five seem to support Nguyen's argument that the biggest problem for online communication might not be epistemic bubbles, but echo chambers. With the discussion of current empirical research on the role of digitalization on the prevalence of echo chambers and epistemic bubbles, I hope to have made clear that the conflation of these phenomena has significant implications for the validity of this research.

Moreover, The results found with the use of models created in the field of network epistemology show that the empirical measuring of trust in epistemological sources and presented information could potentially lead to very interesting and new insights in why people are polarizing. For this reason, one recommendation for future research is to find ways to incorporate such a factor of trust, so that the echo chamber phenomenon is adequately captured in empirical research.

One last, and arguably most important takeaway, is that people do not have to be stupid, lazy, or otherwise epistemically vicious to end up in an echo chamber. The findings from the field of network epistemology discussed in this thesis show that the addition of a simple and arguably justifiable heuristic for trust is enough to even let idealized and rational agents form echo chambers. They also show that it is not even necessary to have powerful malicious agents enforcing such an echo chamber from above: unfortunately, the rational idealized agents from these models are able to form an echo chamber all on their own, even when they are able to ‘test’ their hypothesis against a ground truth.

6.1. Potential interventions

This conclusion paints a rather gloomy picture of how likely we are to succeed in preventing polarization and misinformation to take (an even bigger) hold on society. If we could say that people in echo chambers are namely just epistemically vicious or at least poorly informed, there are clear and potentially very effective possible interventions, namely to educate people better on epistemic virtues and to expose them one way or another to correct information.⁶ Of course, such interventions might still be relevant and valid from an individual and psychological perspective, but the current discussion has shown that they do not solve the whole problem. As O’Connor and Weatherall put it:

“we do think that recognizing the importance of social effects on the spread and persistence of false beliefs, even if we assume that all individuals in a society are perfectly rational (which, alas, they are not), shows that whatever else we do, we also need to think about interventions that take networks into account.” (2019, p. 176)

So, while the interventions discussed above (especially the latter one) would still help against epistemic bubbles, from the system-level perspective, they do not help against echo chambers. As a possible intervention against echo chambers, Nguyen himself proposes what he calls a “*social epistemic reboot*” (2020, p. 157) for any person that has become trapped in one. The echo chamber mechanism is dependent on the order in which someone gets their information: just like in the trust mechanisms described in this thesis, a person tests new information they get (and its trustworthiness) at moment t against the beliefs and information they have gathered up until

⁶ Of course it is debatable what would be understood as correct information, but diving into these philosophical discussions lies outside of the scope of this thesis. For now, I will define ‘correct’ information as information endorsed by current scientific practice.

moment $t-1$. According to Nguyen, a way to escape one's echo chamber is to "undo the influence of historical ordering of their encounters with the evidence" (p. 157). Nguyen argues that this social epistemic reboot is done by letting go of one's beliefs about what sources can be trusted, and look at all gathered information with equal trust. However, expecting this task from a person in an echo chamber seems highly unrealistic, something Nguyen agrees with.

Still, this does not mean that there are no other (more realistic) interventions possible. There is still a variety of possible interventions and advices, some of which have been proposed by researchers in the field of network epistemology.

The first one is to try to teach people to stay as open-minded and as cautiously trusting as possible. O'Connor and Weatherall (2018) show us with the first version of their model, in which agents never completely lost trust in another agents, how important it can be to remain open-minded and at least cautiously trustful of others. Because agents in this model version always kept listening to others and stayed at least cautiously trusting of them, they never definitively polarized, even if they did polarize for a significant amount of time. An important side note to this proposal is that their model represents an ideal situation in which no one is deliberately lying and everybody *can* be trusted. In real life, this is of course not the case, and other research of these same authors has shown how open-mindedness in the presence of deceitful propagandists and lobbyists can even be harmful (Weatherall et al., 2020). Still, I think that one lesson that could be drawn from their findings is that one should be cautious to write another person's opinions completely off, unless one has good evidence that the other agent has malicious intents.

A second potential intervention, proposed by Madsen et al. (2018), is to begin young with creating a common educational ground between people, for example through a strong and trusted public news broadcast. In the multiple network models that have been discussed in this thesis, problems regarding trust arose once agents differed too much from each other in opinion to still trust the other person's opinion and take it into consideration. One way to prevent this problem is to stay ahead of such levels of distrust and corresponding deviant beliefs by creating a trusted educational source that is regarded as a common ground by the majority of people. Such a public broadcast will not persuade agents that already have extreme beliefs, but it does help to prevent agents that still trust such broadcasting platforms from radicalizing beyond the scope of influence. This way, veristic public broadcasting channels do help to slow down polarization and the formation of echo chambers, according to Madsen et al. (2018).

Another possibly effective intervention, proposed by O'Connor and Weatherall, is to keep issues as local as possible in order to "give the world more chances to push back." (2019, p. 176).

The authors argue that such an intervention works because keeping an issue as local as possible “generates a situation better matched by models in which there is a notable difference between the success rates of actions A and B” (2019, pp. 176–177). In other words, if people have more possibilities to empirically ‘test’ their beliefs about a certain issue, they have to rely less on testimony alone, consequently leaving less room for unintended polarizing mechanisms.

Lastly, I would like to add one personal point of advice to the reader: Do not write off conspiracy theorists (or people that simply have a very different opinion than you) as stupid and biased. If there is one thing that the findings discussed in this thesis have shown, it is that people from other backgrounds can just from a simple different chain of events end up on the other side of the belief spectrum from you. It is important to keep this in mind when interacting with different people.

6.2. Limitations and further research

The arguments discussed in this thesis have some limitations. One of these limitations is that while the number of social epistemological studies using networked approaches is rising, the still relatively small number of relevant studies in this field means that there remain still a lot of questions about how certain other mechanisms or specific model properties influence the echo chamber and epistemic bubble mechanisms discussed in this thesis. For example, in the models discussed in this paper where agents were able to ‘do research’ themselves, all agents were equally able to gather evidence. But would the findings in these studies be the same if some agents would have to rely on testimony alone? And how would the outcomes of these models be influenced if some agents would have one-way communication pathways? To be able to answer such questions, more research is needed in which the influence of individual parameters and specific scenarios is tested more thoroughly.

Another limitation is that the models of network epistemology discussed here are only that; *models*. Models can only provide *how-possibly* explanations, meaning that they help explain how a phenomenon can *possibly* be explained. This is different from how a phenomenon can *actually* be explained (Sullivan, 2022). Network epistemology has an exploratory goal, and the goal of this thesis has also been to explore how we could understand echo chambers and epistemic bubbles differently. Moreover, the exploration of the mechanisms behind epistemic bubbles and echo chambers discussed in this thesis have also helped to show echo chambers and epistemic bubble mechanisms can influence a network of agents separately, pointing out flaws in the validity of current empirical research. Theoretical and empirical research can thus

definitely strengthen each other. Still, models can be wrong. Empirical evidence is necessary to actually link the working of a model to reality and to validate the explanatory power of a model (Sullivan, 2022).

One way in which the limit of the actual explanatory power is evident in the case of the network models used in this thesis is in the differences in operationalization of certain concepts, for example the concept of trust. These different operationalizations have not led to radically different outcomes between studies, meaning that we can relatively safely say that a form of the trust heuristic analyzed in this paper is robustly influencing the formation of echo chambers in the models. Still, these differences in exact operationalization show that it is unclear which one is most representative of a real-world mechanism of trust. Moreover, different scholars have different interpretations of what similar rules actually represent: Madsen et al. (2018) interpret their mechanism as a measure of open-mindedness, while other researchers have interpreted similar mechanisms as mechanisms of trust.

While models like the network models explained in this thesis can thus certainly help us to understand the possible mechanisms behind epistemic bubbles and echo chambers and to give direction to empirical research, actual empirical research is still an integral part to really understand the real-world phenomena of epistemic bubbles and echo chambers in question. This limitation only strengthens the recommendation to conduct more empirical research on these phenomena in which factors of trust and connection are incorporated.

Bibliography

- Adamic, L. A., & Glance, N. (2005). The political blogosphere and the 2004 U.S. election: Divided they blog. *Proceedings of the 3rd International Workshop on Link Discovery (LinkKDD '05)*, 36–43. <https://doi.org/10.1145/1134271.1134277>
- Alfano, M., Cunningham, S., Meulemans, W., Rutter, I., Sondag, M., Speckmann, B., & Sullivan, E. (2018). Social network-epistemology. *2018 IEEE 14th International Conference on E-Science*, 320–321. <https://doi.org/10.1109/eScience.2018.00073>
- Arguedas, A. R., Robertson, C. T., Fletcher, R., & Nielsen, R. K. (2022). *Echo chambers, filter bubbles, and polarisation: A literature review*. <https://doi.org/10.60625/risj-etxj-7k60>
- Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. B. F., Lee, J., Mann, M., Merhout, F., & Volfovsky, A. (2018). Exposure to opposing views on social media can increase political polarization. *Proceedings of the National Academy of Sciences of the United States of America*, 115(37), 9216–9221. <https://doi.org/10.1073/pnas.1804840115>
- Bakshy, E., Messing, S., & Adamic, L. A. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, 348(6239), 1130–1132.
- Bala, V., & Goyal, S. (1998). Learning from neighbours. *The Review of Economic Studies*, 65(3), 595–621. <https://doi.org/10.1111/1467-937X.00059>
- Begby, E. (2024). From belief polarization to echo chambers: A rationalizing account. *Episteme*, 21(2), 519–539. <https://doi.org/10.1017/epi.2022.14>
- Bennett, L., & Iyengar, S. (2008). A new era of minimal effects? The changing foundations of political communication. *Journal of Communication*, 58(4), 707–731. <https://doi.org/10.1111/j.1460-2466.2008.00410.x>
- Bleakley, P. (2023). Panic, pizza and mainstreaming the alt-right: A social media analysis of Pizzagate and the rise of the QAnon conspiracy. *Current Sociology*, 71(3), 509–525. <https://doi.org/10.1177/00113921211034896>
- Boyd, K. (2019). Epistemically pernicious groups and the groupstrapping problem. *Social Epistemology*, 33(1), 61–73. <https://doi.org/10.1080/02691728.2018.1551436>
- Bruns, A. (2019). It's not the technology, stupid: How the 'echo chamber' and 'filter bubble' metaphors have failed us. *International Association for Media and Communication Research Conference*, 1–12. [https://eprints.qut.edu.au/131675/1/It's Not the Technology%2C Stupid %28paper 19771%29.pdf](https://eprints.qut.edu.au/131675/1/It's%20Not%20the%20Technology%2C%20Stupid%20paper%2019771.pdf)
- Fletcher, R., & Nielsen, R. K. (2018). Are people incidentally exposed to news on social media? A comparative analysis. *New Media and Society*, 20(7), 2450–2468.

- <https://doi.org/10.1177/1461444817724170>
- Goldman, A. I. (2001). Experts: Which ones should you trust? *Philosophy and Phenomenological Research*, 63(1), 85. <https://doi.org/10.2307/3071090>
- Hahn, U., Merdes, C., & von Sydow, M. (2024). Knowledge through social networks: Accuracy, error, and polarisation. *PLoS ONE*, 19(1), 1–30. <https://doi.org/10.1371/journal.pone.0294815>
- Hartmann, D., Pohlmann, L., Wang, S. M., & Berendt, B. (2025). A systematic review of echo chamber research: Comparative analysis of conceptualizations, operationalizations, and varying outcomes. *Journal of Computational Social Science*, 8(52), 1–59. <https://doi.org/10.1007/s42001-025-00381-z>
- Jamieson, K. H., & Capella, J. N. (2008). *Echo chamber: Rush Limbaugh and the conservative media establishment*. Open University Press.
- Lewandowsky, S., Ecker, U. K. H., & Cook, J. (2017). Beyond misinformation: Understanding and coping with the “post-truth” era. *Journal of Applied Research in Memory and Cognition*, 6(4), 353–369. <https://doi.org/10.1016/j.jarmac.2017.07.008>
- Lozano, A. V. (2024, September 12). “It just exploded”: Springfield woman claims she never meant to spark false rumors about Haitians. *NBC News*. <https://www.nbcnews.com/news/us-news/-just-exploded-springfield-woman-says-never-meant-spark-rumors-haitian-rcna171099>
- Madsen, J. K., Bailey, R. M., & Pilditch, T. D. (2018). Large networks of rational agents form persistent echo chambers. *Scientific Reports*, 8(1), 1–8. <https://doi.org/10.1038/s41598-018-25558-7>
- Mahmoudi, A., Jemielniak, D., & Ciechanowski, L. (2024). Echo chambers in online social networks: A systematic literature review. *IEEE Access*, 12, 9594–9620. <https://doi.org/10.1109/ACCESS.2024.3353054>
- Masip, P., Suau, J., & Ruiz-Caballero, C. (2020). Incidental exposure to non-like-minded news through social media: Opposing voices in echo-chambers’ news feeds. *Media and Communication*, 8(4), 53–62. <https://doi.org/10.17645/mac.v8i4.3146>
- McIntyre, L. (2018). *Post-truth* (1st ed.). MIT Press.
- Mønsted, B., & Lehmann, S. (2022). Characterizing polarization in online vaccine discourse—A large-scale study. *PLoS ONE*, 17(2 February), 1–19. <https://doi.org/10.1371/journal.pone.0263746>
- Munroe, W. (2023). Echo chambers, polarization, and “post-truth”: In search of a connection. *Philosophical Psychology*, 37(8), 2647–2678.

- <https://doi.org/10.1080/09515089.2023.2174426>
- Nguyen, C. T. (2020). Echo chambers and epistemic bubbles. *Episteme*, 17(2), 141–161.
<https://doi.org/10.1017/epi.2018.32>
- O'Connor, C., Goldberg, S., & Goldman, A. I. (2024). *Social epistemology*. Stanford Encyclopedia of Philosophy.
<https://plato.stanford.edu/archives/sum2024/entries/epistemology-social/>
- O'Connor, C., & Weatherall, J. O. (2018). Scientific polarization. *European Journal for Philosophy of Science*, 8, 855–875. <https://doi.org/10.1007/s13194-018-0213-9>
- O'Connor, C., & Weatherall, J. O. (2019). *The misinformation age* (1st ed.). Yale University Press.
- Olsson, E. (2013). A Bayesian simulation model of group deliberation and polarization. In F. Zenker (Ed.), *Bayesian Argumentation* (pp. 113–133). Synthese Library.
https://doi.org/10.1007/978-94-007-5357-0_6
- Pansanella, V. (2024). *Biased echoes: Unraveling mechanisms of opinion dynamics and their impacts in online social networks* [Scuola Normale Superiore].
https://doi.org/10.25429/pansanella-valentina_phd2024-01-30
- Pariser, E. (2011). *The filter bubble: What the internet is hiding from you* (1st ed.). The Penguin Press.
- Perfors, A., & Navarro, D. J. (2019). Why do echo chambers form? The role of trust, population heterogeneity, and objective truth. *Proceedings of the Annual Meeting of the Cognitive Science Society*, 41, 918–923. <https://escholarship.org/uc/item/3zx0w51m>
- Pollock, J. (2024). Epistemic bubbles and contextual discordance. *Philosophy*, 99(3), 437–459.
<https://doi.org/10.1017/S0031819124000093>
- Rekker, R., & Hartevelde, E. (2024). Understanding factual belief polarization: The role of trust, political sophistication, and affective polarization. *Acta Politica*, 59(3), 643–670.
<https://doi.org/10.1057/s41269-022-00265-4>
- Singer, D. J., Bramson, A., Grim, P., Holman, B., Jung, J., Kovaka, K., Ranginani, A., & Berger, W. J. (2019). Rational social and political polarization. *Philosophical Studies*, 176(9), 2243–2269. <https://doi.org/10.1007/s11098-018-1124-5>
- Singer, D. J., Grim, P., Bramson, A., Holman, B., Jung, J., & Berger, W. J. (2022). Epistemic networks and polarization. In M. Hannon & J. De Ridder (Eds.), *The Routledge handbook of political epistemology* (pp. 133–144). Routledge. <https://doi.org/10.4324/9780429326769>
- Sullivan, E. (2022). Understanding from machine learning models. *British Journal for the Philosophy of Science*, 73(1), 109–133. <https://doi.org/10.1093/bjps/axz035>

- Sullivan, E., Sondag, M., Rutter, I., Meulemans, W., Cunningham, S., Speckmann, B., & Alfano, M. (2020). Vulnerability in social epistemic networks. *International Journal of Philosophical Studies*, 28(5), 731–753. <https://doi.org/10.1080/09672559.2020.1782562>
- Sunstein, C. R. (2001). *Republic.com*. Princeton University Press.
- Sunstein, C. R. (2018). *#Republic: Divided democracy in the age of social media*. Princeton University Press.
- Terren, L., & Borge, R. (2021). Echo chambers on social media: A systematic review of the literature. *Review of Communication Research*, 9, 99–118. <https://doi.org/10.12840/ISSN.2255-4165.028>
- Turner, C. (2023). Online echo chambers, online epistemic bubbles, and open-mindedness. *Episteme*, 22(1), 1–26. <https://doi.org/10.1017/epi.2023.52>
- Waisbord, S. (2018). Truth is what happens to news: On journalism, fake news, and post-truth. *Journalism Studies*, 19(13), 1866–1878. <https://doi.org/10.1080/1461670X.2018.1492881>
- Wang, D., & Qian, Y. (2021). Echo chamber effect in rumor rebuttal discussions about COVID-19 in China: Social media content and network analysis study. *Journal of Medical Internet Research*, 23(3), 1–19. <https://doi.org/10.2196/27009>
- Weatherall, J. O., & O'Connor, C. (2021). Endogenous epistemic factionalization. *Synthese*, 198(s25), S6179–S6200. <https://doi.org/10.1007/s11229-020-02675-3>
- Weatherall, J. O., O'Connor, C., & Bruner, J. P. (2020). How to beat science and influence people: Policymakers and propaganda in epistemic networks. *British Journal for the History of Philosophy*, 71, 1157–1186. <https://doi.org/10.1093/bjps/axy062>
- Weisberg, J. (2021). *Formal epistemology* (E. N. Zalta (ed.)). The Stanford Encyclopedia of Philosophy. <https://plato.stanford.edu/entries/formal-epistemology/#FifCasStuSocEpi>
- Woods, J. (2005). *Epistemic bubbles*. 1–39.
- Zollman, K. J. S. (2007). The communication structure of epistemic communities. *Philosophy of Science*, 74(5), 574–587. <https://doi.org/10.1086/525605>
- Zollman, K. J. S. (2013). *Network epistemology: Communication in epistemic communities*. 1, 15–27. <https://doi.org/10.1111/j.1747-9991.2012.00534.x>