The Logical Dynamics of Social Networks: From Homophily to Polarization

MSc Thesis *(Afstudeerscriptie)*

written by

Jonathan Thul
(born November 9th, 1996 in Duisburg, Germany)

under the supervision of **Prof. Dr. Sonja Smets**, and submitted to the Examinations Board in partial fulfillment of the requirements for the degree of

MSc in Logic

at the *Universiteit van Amsterdam*.

Date of the public defense:
September 28th, 2023

Members of the Thesis Committee:
Dr. Balder ten Cate (chair)
Prof. Dr. Sonja Smets
Dr. Malvin Gattinger
Dr. Soroush Rafiee Rad

INSTITUTE FOR LOGIC, LANGUAGE AND COMPUTATION

Abstract

This thesis studies the effect of homophily on the development of social networks. Homophily refers to the tendency of individuals to be socially connected to others that are similar to themselves. We contend that this phenomenon can be traced back to two tendencies, social influence and social selection. Using techniques from dynamic epistemic logic, we provide a formal setting to represent social networks and define model transforming upgrades that correspond to social influence and social selection.

In the first part of the thesis, we introduce the notion of cluster-split models and argue that they represent socially fragmented networks. We show how social selection gives rise to such models and argue that this suggests a connection between homophily and polarization.

In the second part we introduce epistemic social network models. This allows us to define different epistemic update versions of social influence and social selection. We compare these updates, showing that the different ways agents deal with epistemic uncertainty can lead to different network developments. Finally, we survey phenomena of learning that arise in our framework.

Contents

Introduction

The last three decades have seen the introduction of wholly new channels of communication. In the internet age, influential new social structures have emerged through the rise of social media and societies across the globe have undergone unexpected social and political shifts. Explaining and controlling such changes in social dynamics requires us to understand how groups and opinions form and how they influence each other. There are many ways of gaining such an understanding. This thesis attempts to do so using the formal tools of *social network analysis* – methods from graph theory, logic and game theory.

A central aim of social network analysis is to propose and study principles that can explain the structural features and development of social networks[1] – i.e. who is socially connected to whom, and who acts or thinks in certain ways. One such organizational principle often proposed is *homophily* – the tendency of people to be socially connected to others that are similar to themselves.

In this thesis, we will introduce a formal framework to describe social networks and the way they are affected by homophily-driven changes. Social influence and similarity-driven group formation have been studied independently of each other in previous works. This thesis contributes to the existing research by

1. providing a formal framework that can model both network- and opinion-changing processes happening to the same social network,

2. giving a number of stabilization results for homophily-driven network change, some of which suggest a connection between tendencies of homophily and social polarization and

3. introducing epistemic social network models which explicitly model the knowledge and uncertainty of agents, allowing us to investigate the influence of epistemic factors on the development of our social network models.

[1] It is worth noting that in the field and within this thesis, "social network" is not referring to social media sites. Instead, we mean to refer to any social structure made up of individuals, e.g. school classes, clubs and towns. Social networks found on social media sites are a special case of this more general concept.

The thesis is structured as follows. In Section 1, we give a brief introduction to the concept of homophily. We then sketch the way threshold models are used to represent processes of social influence. In Section 2, we formally define social network models and introduce two threshold-style model transforming updates that correspond to social influence and social selection. We then show how both updates interact, proving a number of stabilization results for single-issue models. We introduce the notion of a cluster-split, show how it connects to our updates and argue that our results have implications for social fragmentation and polarization. Section 2 ends with a first attempt at characterizing the phenomenon of oscillation. Section 3 introduces the reader to the basic ideas and techniques of epistemic logic. In Section 4, we employ these techniques to extend our models with an epistemic dimension. We go on to define different updates, corresponding to different ways of dealing with uncertainty about the network state. We then compare these updates, proving that they lead to different development dynamics. Finally, we briefly survey the phenomenon of learning in our epistemic social network models.

1 Preliminaries: Homophily and Threshold Models

As mentioned above, homophily refers to the tendency of people to be socially connected to others that are similar to themselves. Observations of this principle were contained in the first works on social network analysis, and suggestions of it can be found in works as early as Plato's *Phaedrus*.[2] But one does not need to consult academic literature to observe homophily – examples are easy to spot "in the wild": In school, children are more likely to make friends with other children that share their interests; family members are more similar to each other in habits and behaviors than to people outside of the family; and conventional wisdom has it that in the age of social media, people online seek out contact to like-minded others, leading to the creation of "echo chambers" that lead everyone involved to converge on ever more extreme views.

Reflecting on such phenomena, it seems that there are two distinct tendencies that contribute to homophily. Given a group of people and some dimension of similarity, e.g. musical taste or political leaning, we find that

1. people form and maintain social connections to others that are similar to them, and
2. people are influenced in their views by their social contacts.

The first tendency is sometimes called *social selection*, the second is *social influence*.

There is a lot of work in social network analysis and social network logic on social influence [4, 19, 14, 22, 32]. The basic framework employed in many investigations of social influence are so-called *threshold models*. Such models work with the following four "ingredients": A behavior that agents can adopt or not adopt; a set of agents, some of which show the behavior as initial adopters; a social relation that holds between some of the agents and a threshold that expresses what fraction of an agent's neighbors need to adopt the behavior for the agent herself to do the same. From these initial conditions, threshold models describe how the behavior spreads from agent to agent in steps, assuming that this happens through a mechanism similar to conformity pressure: In each time-step t, agents that showed the behavior at $t-1$ continue to do so; all other agents adopt the behavior if and only

[2] See [21] for a brief overview on the history of the concept.

if the fraction of their social contacts that show the behavior matched or exceeded the threshold at $t-1$.[3]

Formally, many instances of threshold models make use of techniques from the field of dynamic epistemic logic (DEL).[4] Dynamic epistemic logics are formal frameworks to represent changes in the knowledge of agents.[5] To this end, the epistemic state of each relevant agent, i.e. what they know or do not know, is represented through a model. Events that change the agents' knowledge – a classic example is the announcement of some fact to every agent – then correspond to a *model transforming update*: Given a DEL model $\mathbb{M}$, an application of such an update leads to a new model $\mathbb{M}'$ that differs from $\mathbb{M}$ in what the agents know and do not know. The inclusion of model transforming updates gives DEL its *dynamic* character, enabling it to be a tool to reason about changes in its models.

Classic threshold models as described above are not concerned with knowledge, but the dynamic nature of the behavior spread they describe is a good fit for the use of DEL-style methods. Instead of representing knowledge, a singular model represents the state of a social network; and instead of knowledge change, model transforming updates are used to illustrate how a behavior spreads within the network.

Often the spreading behavior in a threshold model takes the form of a chain reaction: Starting with the initial adopters, a large number of agents is eventually made to adopt the behavior step by step. This kind of spreading dynamic is called a *cascade*, and is said to be *complete* when it leads to every agent adopting the behavior in the long run.[6] Threshold models allow for describing cascades precisely and to characterize, for example, all models that lead into a complete cascade.

But with respect to homophily, classical threshold models only tell half the story – they take the social relation between agents to be unchanging, and so they cannot represent the tendency of social selection. Recent work in social network logic [30, 29, 28, 2] has taken steps to change this and proposed mechanisms through which agents in networks begin new social

[3]See [12, Chapter 19] for a detailed introduction to threshold models.

[4]See for instance [10, 4, 28, 2].

[5]A short introduction to dynamic epistemic logic can be found in [6], a more extensive one is provided by [11].

[6]Note that cascades in threshold models are different from so-called "informational cascades" in which sequentially acting Bayesian-rational agents are found to ignore their private information to act in accordance with group behavior. This phenomenon is described in e.g. [15] and [3].

relations and end existing ones.

The aim of this thesis is to further explore this by modelling both tendencies of homophily in one formal framework. This will allow us to better understand how social influence and social selection interact, and how their interaction gives rise to certain structural properties of social networks. We begin by introducing the basic formal model in the next section.

2 Social Network Models

In this thesis, we represent a social network as a set of agents supplemented with a binary relation, interpreted as a social *neighborhood relation*.

Definition 1 (Social Network Model)**.** *A social network model (SNM)* M *is a tuple* $(A, N, F, V, \omega, \tau)$*, where:*

- A *is a non-empty finite set of agents;*
- $N : A \to \wp(A)$ *is the neighborhood function, assigning a set of neighbors* $N(a)$ *to each agent* a *such that for all* $a, b \in A$*:*
  - $a \notin N(a)$ *(irreflexivity) and*
  - $a \in N(b)$ *if and only if* $b \in N(a)$ *(symmetry);*
- F *is a non-empty finite set of issues;*
- $V : A \to \wp(F)$ *is a valuation function, assigning to each agent a set of issues they accept or support;*
- $\omega, \tau \in \mathbb{Q}$ *are two rational numbers s.t.* $0 \leqslant \omega \leqslant 1$ *and* $0 \leqslant \tau \leqslant 1$*, interpreted, respectively, as similarity threshold and influencability threshold.*

This model is similar to models defined in [2] and [4]. In contrast to the former paper, we require the neighborhood relation between agents to be irreflexive and symmetric. We thus focus on social relations that are mutual between agents, friendship and neighborhood being obvious examples. In contrast to the latter paper, we do not require the relation to be serial: We will go on to define update operations that change the relational structure of a given model, and these changes can lead to isolated agents making new connections and formerly connected agents being isolated. Requiring the neighborhood relation to be serial would make it impossible to implement such updates.

The valuation function V takes an agent a and returns a set of issues supported by the agent $V(a)$; all other issues, i.e. $F \setminus V(a)$ are not supported by a. In this thesis, we will refer to the position an agent takes with respect to some issue $f \in F$ as her *opinion*. When $f \in V(a)$, we will say that a *supports* f. Otherwise, we will say that a supports $\neg f$. This way of speaking is slightly misleading, as we make no attempt to explicitly model the structure of agents' opinions or beliefs. Instead, we take opinions to be primitive in the model. Every agent has an opinion on every issue in F; for every issue, the two possible opinions to have are *support* or *not support*; no two issues are logically interdependent. There is an extensive literature on doxastic logic and belief revision that dispenses with these simplifying assumptions.[7] But the focus of this thesis is on the interplay between network-driven behavior changes and behavior-driven network changes. To keep our formalism as simple and as general as possible, we will not go into the topic of belief structures – and since opinions are an intuitive example of a behavior that is influenced through social networks and exerts influence on the development of social networks, we will stick to the conventions just described.

Following the same ideas as the classic threshold models introduced in Section 1, we include two different thresholds in our model definition: The similarity threshold ω is used to formally represent what fraction of their total opinions two agents have to agree on in order to count as "similar enough" to become neighbors. The influencability threshold τ represents what fraction of an agents' neighbors need to hold an opinion for this agent to be influenced to adopt the same opinion. For any given model, both thresholds are universal and fixed: They apply to all agents equally, and they do not change, even when other aspects of the model do.[8]

Figure 1 below shows how social network models can be represented graphically. Here, there are three agents a, b and c. a and b are neighbors, and so are a and c, while b and c are not neighbors. Further, b is the only agent supporting the issue f. Whenever we represent *single-issue SNMs* – i.e. social network models with $|F| = 1$ – in this thesis, we will color supporting nodes black and non-supporting nodes white.

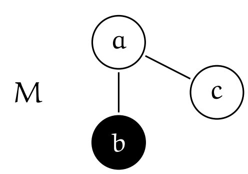

Figure 1: A single-issue social network model with $F = \{f\}$, $\tau = \frac{1}{2}$ and $\omega = 1$. Black nodes represent agents that support the issue f, white nodes represent agents that do not support f.

[7] See [20], [8] and [16] as well as [6] for an overview.

[8] In [19], the authors consider the case of diffusion through social network models with agent-specific thresholds.

2.1 Opinion Updates

Given a social network model, we want to be able to update it to represent some change in the model's structure. The first type of update we will consider represents a change in the agents' opinions. There are many mechanisms by which such a change in opinions could conceivably take place. We focus on a simple type of *social influence*: Given some issue $f \in F$, an agent begins to support it if there is a sufficiently large portion of f-supporters in the neighborhood.

Definition 2 (Monotonic Opinion Update)**.** *Given a social network model* $M = (A, N, F, V, \omega, \tau)$*, the updated model* $M_{\triangle} = (A, N, F, V_{\triangle}, \omega, \tau)$ *is such that for any* $a \in A$ *and* $f \in F$*:*

$$f \in V_{\triangle}(a) \quad \textit{iff} \quad \left\{ \begin{array}{ll} f \in V(a), & \textit{if } N(a) = \emptyset \\ f \in V(a) \vee \frac{|N^f(a)|}{|N(a)|} \geqslant \tau, & \textit{otherwise} \end{array} \right\}$$

Where $N^f(a) = \{b \in N(a) : f \in V(b)\}$ *is the set of* a*'s* f-neighbors*.*

For this update, the threshold τ acts as a formal representation of what a "sufficiently large portion" of neighbors is. Note that the only thing changing through the update are the agents' opinions; the set of agents, the set of issues, the neighborhood function and both thresholds are kept fixed. Under repeated application of the update $\triangle$, agents can only start supporting issues, but they can never stop their support. This means that the set of f-supporters in a model can never decrease in size when $\triangle$ is applied, which is why we call it a *monotonic* opinion update. This mechanism corresponds to opinion spread updates in classical threshold models, used e.g. in [10] and [4]. As a mechanism of how opinions diffuse through a social network, a monotonic update like this is not entirely plausible: Most people are not so stubborn that they can never change their opinion once they have started to support a given issue. Still, we include monotonic opinion updates in this thesis. They serve as a comparison to other dynamics of opinion change introduced below, and they provide a link to other work in social network analysis. Overstretching our terminology of issues and opinions, monotonic opinion updates can also be understood as modelling e.g. technology adoption: Suppose there is a new version of a popular messaging app. Updating the app is a bit of a hassle, but after the update, agents can access useful features when communicating with other up-to-date app users. In such a scenario, agents will be inclined to update their app as soon as enough of their social contacts have done so.[9] But once they do this, they cannot "un-update" and return to the older version of the app. The adoption of the update within a social network could be modeled using the monotonic opinion update from Definition 2.

The next definition describes an opinion update mechanism that does away with monotonicity and allows agents to also stop supporting an issue:

Definition 3 (Non-Monotonic Opinion Update)**.** *Given a social network model* $M = (A, N, F, V, \omega, \tau)$*, the updated model* $M_{\triangle} = (A, N, F, V_{\triangle}, \omega, \tau)$ *is such that for any* $a \in A$ *and* $f \in F$*:*

$$f \in V_{\triangle}(a) \quad \textit{iff} \quad \left\{ \begin{array}{ll} f \in V(a), & \textit{if } N(a) = \emptyset \\ f \in F, & \textit{if } \tau = 0 \\ \frac{|N^f(a)|}{|N(a)|} \geqslant \tau, & \textit{otherwise} \end{array} \right\}$$

On the right side of the definition, the first condition pertains to neighborless agents: Through an update with $\triangle$, they do not change their position on any of the issues in F. The second condition deals with the special case of the threshold τ equaling zero: In this case, all agents support all issues in F after the update. The third condition covers the remaining cases: In models with $\tau > 0$, agents that have at least one neighbor will support an issue after the update with $\triangle$ if a sufficient amount of their neighbors already supported it before the update. Otherwise, such agents will stop supporting the issue after the update. As such, non-monotonic opinion updates can be used to model how e.g. political opinions, hobbies or habits diffuse from agent to agent.

The difference between the two types of opinion update is illustrated in Figure 2. In the initial model M, agents a and b support the issue, but both of their neighbor sets contain less than $\tau = \frac{2}{3}$ supporters. When updated with $\blacktriangle$, agent c starts supporting the issue, and since the update is monotonic, agents a and b keep supporting it, too. In contrast, the update with $\triangle$ leads to the model $M_{\triangle}$, where c has likewise started supporting the issue, but a and b have stopped their support.

Using opinion updates on social network models, we can simulate how agents influence each other and how opinions "travel" through the network.

[9]Clearly the specific threshold τ – representing what fraction of an agents' contacts have to update before the agent is inclined to do the same – depends on how useful the update is, and how much work it is to update. [12, Chapter 19.2] shows how this idea can be made formally precise, relating the influencability threshold to payoffs in a coordination game between agents.

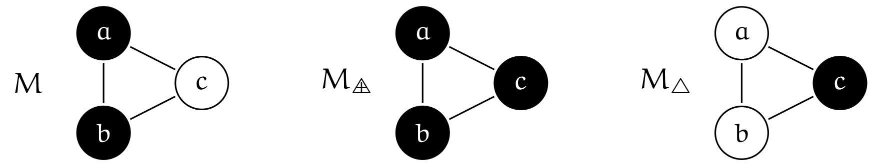

Figure 2: Three single-issue SNMs with $\tau = \frac{2}{3}$. $M_{\triangle\!\!\!\!\dagger}$ has been obtained by applying a monotonic opinion update to M; $M_{\triangle}$ has been obtained by updating M with a non-monotonic opinion update.

In general, there are two polar states that a model can be in with respect to a given issue $f \in F$: Everyone supports it – i.e. for all $a \in A$, $f \in V(a)$ – or no one does, i.e. for all $a \in A$, $f \notin V(a)$. Throughout this thesis, we will refer to these states as f-*consensus* and $\neg f$-*consensus*, respectively.

Now, given repeated application of an opinion update, which models will eventually become an f- or $\neg f$-consensus? Or, put more technically: Which models will *cascade* into an f- or $\neg f$-consensus under application of $\triangle\!\!\!\!\dagger$ or $\triangle$? For $\triangle\!\!\!\!\dagger$, the answer is given by a classic result, first proved for a related, but somewhat different type of model.[10] This result can be adapted for our present setting – but to do so, we need to first introduce *update streams* and the concept of a *cluster*.

Definition 4 (Update Stream)**.** *Let* **U** *be the set of all updates on social network models. An update stream* $\vec{\dagger}$ *is an infinite sequence of updates* $(\dagger_n)_{n \in \mathbb{N}}$ *(with* $\dagger \in \mathbf{U}$*). A* repeated update *is an update stream of the form* $(\dagger_1, \dagger_2, \dagger_3, \dots)$*, i.e. an update stream that consists of only one type of update.*

An update stream $\vec{\dagger}$ *induces a function mapping every model* M *into an infinite sequence* $\vec{\dagger}(M) = (M_n)_{n \in \mathbb{N}}$ *of models, defined inductively by:*

$$M_0 = M \quad \textit{and} \quad M_{n+1} = \dagger_n(M_n)$$

The notion of update streams, adapted from [8], gives us a convenient way to talk about the "long term" development of a social network model under sequences of updates. By introducing the idea of a cluster, we gain a way to describe models that have a certain structural property, which will likewise help us in precisely describing which SNMs become an f-consensus under which repeated update.

[10] Originally found in [22]. For a simpler version of the proof that is closer to our present setting, see [12, Chapter 19.3].

Definition 5 (Cluster of density d). *Let A be a set of agents and $N : A \to \wp(A)$ be a neighborhood function, as defined in Definition 1. For any nonempty set of agents $C \subseteq A$, C is a* cluster of density d, *where $d \in [0,1]$ is the greatest value such that for all $c \in C$:*

$$\frac{|N(c) \cap C|}{|N(c)|} \geqslant d \quad \textit{or} \quad N(c) = \emptyset$$

It is worth noting that this is an unorthodox definition of a cluster, different from the ones given in [10, p. 141] and [12, p. 574]. This is motivated by the fact that, in the context of this thesis, we understand clusters with a high density as antagonists of opinion spread. Since we do not presuppose the neighborhood relation N to be serial, we have to adapt the definition of clusters to account for agents without neighbors.[11] Keeping in mind how the opinion updates above are defined, we see that neighborless agents are an insurmountable obstacle to the spread of an opinion: They will not change their mind on an issue f after an update with $\mathbin{\triangle\mkern-13mu+}$ or $\triangle$. To reflect this, we change the definition of clusters so that (singleton sets of) neighborless agents count as clusters of density 1. Figure 3 below is used to illustrate how the definition works: There, the sets $\{a\}$, $\{b\}$ and $\{c\}$ are all clusters of density 0. $\{a,b\}$, $\{a,c\}$ and $\{b,c\}$ all have density $\frac{1}{2}$, and $\{a,b,c\}$ and $\{d\}$ have density 1. A "mixed" set like $\{a,b,d\}$ has density $\frac{1}{2}$.

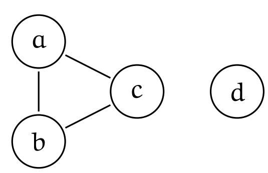

Figure 3: A single-issue social network model.

Now we can make explicit how clusters are related to reaching an f-consensus under the repeated monotonic opinion update $\vec{\mathbin{\triangle\mkern-13mu+}}$.

Theorem 2.1 (Cluster-Cascade Theorem). *Let $M = (A, N, F, V, \omega, \tau)$ be a social network model with $\tau > 0$ and a set $B \subset A$ of f-supporting agents.*

[11] With the more classic definition, a cluster of density d is any subset $C \subseteq A$ such that for all $c \in C$, $\frac{|N(c) \cap C|}{|N(c)|} \geqslant d$. In a serial social network, this clause applies to all agents and ensures that $0 \leqslant d \leqslant 1$. With neighborless agents, this is not the case, so we modify the definition by adding a disjunct covering neighborless agents and explicitly restricting the value of d to the interval $[0,1]$.

Then there exists a model $M_n \in \vec{\mathbb{A}}(M)$ such that M_n is an f*-consensus if and only if there are no clusters of* $\neg$f*-adopters of density* $d > (1 - \tau)$ *in* M.

Proof. $(\Rightarrow)$ By contraposition. Suppose there is a $C \subset A$ such that for all $c \in C$, $f \notin V_0(c)$. Further, let $N_0(c) = \emptyset$ or $\frac{|N_0(c) \cap C|}{|N_0(c)|} \geqslant d > (1 - \tau)$ for each $c \in C$. Then none of the agents in C will ever support f. To see this, suppose otherwise: Then there must be an update stage $M_m \in \vec{\mathbb{A}}(M)$ at which the first agents in C first support f. Let c' be one of these agents. By assumption, $f \notin V_{m-1}(c')$, so we know that $N_{m-1}(c') \neq \emptyset$. Then it must hold that $\frac{|N^f_{m-1}(c')|}{|N_{m-1}(c')|} \geqslant \tau$. We know that no agent in C supported f in M_{m-1}, so $(N_{m-1}(c') \cap C) \cap N^f_{m-1}(c') = \emptyset$, so $|N_{m-1}(c') \cap C| \leqslant (1 - \tau)$. But by Definition 2, we know that $N_{m-1}(c') = N_0(c')$, so this contradicts our original assumptions about C.

$(\Leftarrow)$ By contraposition. Suppose that an f-consensus is not reachable from M via $\mathbb{A}$-updates, meaning that M does not stabilize in an f-consensus. Let $M_n \in \vec{\mathbb{A}}(M)$ be the first stable state of M. Since it is not an f-consensus, we know that the set of $\neg$f-supporters $C = \{b \in A : f \notin V_n(b)\}$ is nonempty. Further, we know that for each agent $c \in C$, either (i) $N(c) = \emptyset$ or (ii) $\frac{|N^f_n(c)|}{|N(c)|} < \tau$. The latter of these two conditions is equivalent to $\frac{|N(c) \cap C|}{|N(c)|} > (1 - \tau)$. So C is a cluster of density $d > (1 - \tau)$. □

Under the repeated monotonic opinion update $\vec{\mathbb{A}}$, and given that there are already f-supporters in the network (and that $\tau > 0$), the only obstacle to a spread of support for f are clusters of $\neg$f-supporters, i.e. groups of agents that do not support the issue and are so interconnected that they keep each other from changing their opinion. For the remainder of this thesis, we will refer to such clusters of $\neg$f-supporters of density $d > (1 - \tau)$ as *very dense clusters of* $\neg$f*-supporters*. On a technical level, Theorem 2.1 shows that, equipped with only the monotonic opinion update $\mathbb{A}$, the social network models from Definition 1 work like the more classical threshold models described in Section 1.

In contrast, the direct analogue of the theorem for the non-monotonic opinion update does not hold. To see this, consider Figure 4: M is a social network model with a non-zero threshold τ, a set $\{a\}$ of f-supporters and no very dense clusters of $\neg$f-supporters – the only such cluster is $\{b\}$, which has a density of $0 \not> \tau$. Yet M does not cascade into an f-consensus under application of $\triangle$-updates.

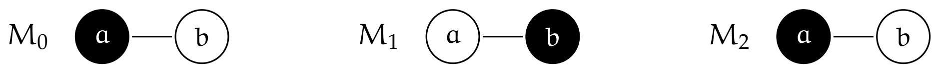

Figure 4: A succession of models, starting from a single-issue M with $\tau = \frac{1}{2}$ and $\omega = 1$ and continuing with $M_1, M_2 \in \vec{\triangle}(M)$. Note that $M_2 = M$, $M_3 = M_1$ *et cetera*.

The counterexample gives us the opportunity for another observation: Under the repeated update $\vec{\triangle}$, models are not guaranteed to *stabilize* at all, instead they can *oscillate* back and forth. These concepts will be precisely defined and explored in Section 2.3.

An important thing to note is that the model in Figure 4 only acts as a counterexample to the right-to-left direction of Theorem 2.1. The other direction of the theorem still holds for non-monotonic opinion updates.

Proposition 2.1. *Let* $M = (A, N, F, V, \omega, \tau)$ *be a social network model, with* $\tau > 0$ *and a set of agents* C *such that for all* $c \in C$, $f \notin V(c)$ *and* C *has density* $d > (1-\tau)$. *Then there is no model* $M_n \in \vec{\triangle}(M)$ *such that* M_n *is an* f*-consensus.*

Proof sketch. This proof works very similar to the one given for the left-to-right direction of Theorem 2.1: Assume that the initial model M contains a cluster of $\neg f$-adopters C of density greater than $(1 - \tau)$. In addition, assume that there is an f-consensus $M_m \in \vec{\triangle}(M)$. Then there must be a set of agents in C that first adopt f. For each of them, it cannot be the case that their set of neighbors is empty – as otherwise they would not switch to supporting f – so the ratio of f-supporting neighbors to all neighbors must be sufficiently large. But then it cannot be true that C is a cluster of the required density, which is a contradiction. □

Thus, we see that dense clusters of $\neg f$-supporters continue to be an obstacle to f-cascades under $\vec{\triangle}$. And since under this type of opinion update, models can also cascade into a $\neg f$-consensus, we can prove a mirror image of Proposition 2.1 for $\neg f$-cascades.

Proposition 2.2. *Let* $M = (A, N, F, V, \omega, \tau)$ *be a social network model with a set of agents* C *such that (i) for all* $c \in C$, $f \in V(c)$ *and (ii)* C *has density* $d \geqslant \tau$. *Then there is no model* $M_n \in \vec{\triangle}(M)$ *such that* M_N *is a* $\neg f$*-consensus.*

Proof. For contradiction, assume that there is a model $M_n \in \vec{\triangle}(M)$ such that M_n is a $\neg f$-consensus. Then there must be a set $C' \subseteq C$ of agents that first drop support for f in a model M_m (with $m \leqslant n$). Consider some such agent $c' \in C'$. We know that $f \in V_{m-1}(c')$ and $f \notin V_m(c')$. So it must hold that $N_{m-1}(c') \neq \emptyset$ and $\frac{|N^f_{m-1}(c')|}{|N_{m-1}(c')|} < \tau$. But $(C \cap N_{m-1}(c')) \subseteq N^f_{m-1}(c')$ (since c' is assumed to be among the first agents of C that drop support for f in M_m). So then $\frac{|C \cap N_{m-1}(c')|}{|N_{m-1}(c')|} < \tau$, which contradicts the assumption that C is a cluster of density $d \geqslant \tau$. □

For the remainder of this thesis, we will call clusters of f-supporters of density $d \geqslant \tau$ *very dense clusters of* f*-supporters*. While Proposition 2.2 shows us that the non-existence of these clusters is necessary for a model to become a $\neg f$-consensus, the model in Figure 4 once again proves that it is not sufficient: M has no very dense clusters of f-supporters, and yet it will never become a $\neg f$-consensus under $\vec{\triangle}$.

What do these results mean? The counterexample given in Figure 4 and Proposition 2.1 illustrate that social network models have to fulfill stricter conditions in order to become an f-consensus under the repeated non-monotonic opinion update $\vec{\triangle}$ than they would have to under its monotonic counterpart $\vec{\mathbb{A}}$. In addition, Proposition 2.2 shows us that the requirements for becoming a $\neg f$-consensus are similarly strict. This is not unexpected: Under the repeated monotonic opinion update $\vec{\mathbb{A}}$, agents that were "convinced" to adopt an issue $f \in F$ at some point will stay convinced indefinitely. But in a process like $\vec{\triangle}$ which allows agents to change their mind and start or stop supporting any issue in F any number of times, it is harder to get everybody to agree on one stance.

2.2 Social Relation Updates

The opinion updates defined in the last section reflect the tendency that in a social setting, agents tend to become more similar to their peers or neighbors. For this section, we look at the converse effect: Determined by similarities or differences in opinions, agents might end existing social relations or begin new ones. Threshold-based mechanisms of social network creation like the one proposed here have also been used in [28], [30], [27] and [2].

Definition 6 (Social Relation Update)**.** *Given a SNM* $M = (A, N, F, V, \omega, \tau)$, *the updated model* $M_\square = (A, N_\square, F, V, \omega, \tau)$ *is such that for any* $a, b \in A$,

$$a \in N_\square(b) \quad \textit{iff} \quad \left(a \neq b \textit{ and } \frac{|(V(a) \cap V(b)) \cup (\overline{V(a)} \cap \overline{V(b)})|}{|F|} \geqslant \omega \right)$$

where "$\overline{V(a)}$" and "$\overline{V(b)}$" are used to refer to the set-theoretic complement of $V(a)$ *or* $V(b)$.

Through a social relation update, all agents in the model simultaneously re-evaluate their social ties: They end relations with agents whose opinions on the issues on F do not overlap to a sufficient degree with their own opinions, and they connect to agents with whom they do have a sufficient overlap. The similarity threshold ω acts as a formal reflection of what a "sufficient degree" is.

In Definition 1, we defined the neighborhood relation N to be irreflexive and symmetric. For any updated model $M_{\Box}$, the relation $N_{\Box}$ is also irreflexive and symmetric, so our definition of the relation update does indeed yield another proper social network model.

Corollary 2.2.1. *Given any social network model* M, $N_{\Box} \in M_{\Box}$ *will be such that for all* $a, b \in A$, *(i)* $a \notin N_{\Box}(a)$ *and (ii)* $a \in N_{\Box}(b) \Leftrightarrow b \in N_{\Box}(a)$.

Proof. Follows immediately from Def. 6. □

We also now see how the social relation update of Definition 6 can produce a non-serial neighborhood relation: Agents that do not share opinions to a sufficient degree with anybody else will have no neighbors after the update.

2.3 Stabilization, Fragmentation and Polarization

Having introduced updates on agents' opinions and relations, we now turn to questions of *stabilization*. Intuitively, a social network model is stable under a series of updates if the updates do not change anything about the structure of the model.

We can make this idea more precise by using the concept of update stream introduced in Definition 4, essentially adapting the idea that a model that is stable under an update stream $\vec{\dagger}$ is a fixed point of the model change induced by $\vec{\dagger}$ from [8].

Definition 7 (Stabilization and Stability)**.** *A model* M stabilizes under an update stream $\vec{\dagger}$ *if there exists some* $n \in N$ *such that* $M_n = M_m$ *for all* $m > n$ *and* $M_n, M_m \in \vec{\dagger}(M)$. *A model* M is stable under an update stream $\vec{\dagger}$ *if* $M_0 = M_n$ *for all* $n \geqslant 0$, *i.e. if the point of stabilization has already been reached.*

It turns out that under repeated updates, there is an easy way to tell if any given model stabilizes: As soon as an update step returns the same model, the model is stable. This will turn out to be handy in the following proofs, and it also shows that our notion of stabilization coincides with the notion expressed in [2].

Proposition 2.3. *For any repeated update* $\vec{\dagger}$ *and any social network model* M, M *stabilizes under* $\vec{\dagger}$ *if and only if there are two models* $M_n, M_{n+1} \in \vec{\dagger}(M)$ *such that* $M_n = M_{n+1}$.

Proof. ($\Leftarrow$) For any update $\dagger \in \mathbf{U} = \{\triangle\!\!\!\!\!\text{+}, \triangle, \square\}$, we know that the outcome of any application of $\dagger$ to a given SNM M is uniquely determined by the model and Definition 2, 3 or 6. So if for $M_n, M_{n+1} \in \vec{\dagger}(M)$, $M_n = M_{n+1}$, then $M_n = M_{n+m}$ for any $m \in \mathbb{N}$. Specifically:

- For $M_n, M_{n+1} \in \vec{\triangle\!\!\!\!\!\text{+}}(M)$: Assume that $M_n = M_{n+1}$. Then, by Definition 2, we know that for any agent $a \in A$ and any issue $f \in F$, $f \in V_n(a)$ if and only of one of the two conditions in the definition holds in M_n. But by assumption, the same then holds for M_{n+1}, so then $M_{n+1} = M_{n+2}$, *et cetera*.

- For $M_n, M_{n+1} \in \vec{\triangle}(M)$: Suppose that $M_n = M_{n+1}$. By Definition 3, we know that for any agent a and any issue f, $f \in V_n(a)$ if and only if at least one of the three conditions given in the definition is satisfied. Again by assumption, the same holds for M_{n+1}, meaning that $M_{n+1} = M_{n+2}$ and so on.

- For $M_n, M_{n+1} \in \vec{\square}(M)$: Suppose $M_n = M_{n+1}$, then by Definition 6 we know that for any two agents a, b, $a \in N_n(b)$ and $b \in N_n(a)$ if and only if $a \neq b$ and $\frac{|(V(a) \cap V(b)) \cup (\overline{V(a)} \cap \overline{V(b)})|}{|F|} \geqslant \omega$. But then, by assumption, the same holds in M_{n+1}, meaning that $M_{n+1} = M_{n+2}$ and so forth.

So then M_n is stable under the repeated update $\vec{\dagger}$, so M stabilizes.

($\Rightarrow$) If M stabilizes under the repeated update $\vec{\dagger}$, then $M_n = M_{n+1}$ for some $n \in \mathbb{N}$ by definition. $\square$

Corollary 2.3.1. *For any repeated update* $\vec{\dagger}$ *and any social network model* M, M *is stable under* $\vec{\dagger}$ *iff for* $M_0, M_1 \in \vec{\dagger}(M)$, $M_0 = M_1$.

Armed with this knowledge, we can now begin investigating which social network models stabilize under which update streams. For the remainder of this section, we will mostly focus on single-issue models. It turns out that under repeated monotonic opinion updates and repeated relation updates, all such models stabilize:

Proposition 2.4. *Any single-issue SNM* M *stabilizes under a repeated update* $\vec{\mathbb{A}}$ *in at most* $|A-1|$ *steps, i.e.* $M_{|A-1|} = M_m$ *for* $m \geqslant |A|$ *and* $M_{|A-1|}, M_m \in \vec{\mathbb{A}}(M)$.

Proof. First, note that if M_0 is an f-consensus or a $\neg f$-consensus, it stabilizes under $\vec{\mathbb{A}}$: To see this, consider that either $\tau > 0$ or $\tau = 0$. If M_0 is an f-consensus, then $M_0 = M_1$ by Definition 2 in either of those cases, since no agent will ever stop supporting f. If M_0 is a $\neg f$-consensus, then we know that for any agent $a \in A$, it holds that $f \notin V_0(a)$ and $N_0^f(a) = \emptyset$. By Definition 2, we then know that if $\tau > 0$, $f \notin V_1(a)$ for any a, meaning that $M_0 = M_1$. In case τ equals zero, M_1 will be an f-consensus, which stabilizes by the argument just given above.

Now, suppose M_0 is neither type of consensus. Let A_0^f be the set of f-supporters in M_0, and let $A_0^{\neg f}$ be the set of $\neg f$-supporters in the same model. Then $A_0^f \neq \emptyset$ and $A_0^{\neg f} \neq \emptyset$. Since we are working with the repeated monotonic opinion update $\vec{\mathbb{A}}$, we know that for each update step from M_n to M_{n+1}, either $A_n^f = A_{n+1}^f$ or $A_n^f \subset A_{n+1}^f$. In the first case, we know by Corollary 2.3.1 that M_n is stable. In the second case, we can proceed to the next update step and see if $A_{n+1}^f = A_{n+2}^f$ or $A_{n+1}^f \subset A_{n+2}^f$. This process can only go on for as long as $A_{n+m}^f \neq \emptyset$. So it can take at most $|A_0^f|$ steps (or less, if at some update step more than one agent adopts f). But $|A_0^f| \leqslant (|A|-1)$. This completes the proof. $\square$

This means that under the repeated monotonic opinion update $\vec{\mathbb{A}}$, any social network will stop changing after a finite number of steps: Once, for any issue $f \in F$, all agents that are not in a very dense cluster of $\neg f$-supporters have adopted f, the model enters a kind of deadlock and does not change under any further application of $\mathbb{A}$. Returning to our example of a messaging app update (given on page 9), this means that some agents might never install the update. If they are part of a close-knit group of agents that use the old version of the app – or if they are not connected to anybody else – they will never have incentive enough to go through the hassle of updating.

Proposition 2.5. *Any SNM* M *stabilizes under a repeated update* $\vec{\Box}$ *in one step, i.e.* $M_n = M_{n+1}$ *for* $M_n, M_{n+1} \in \vec{\Box}(M)$ *and* $n > 0$*.*

Proof. After one update with $\Box$ on M_0, i.e. in the model $M_1 \in \vec{\Box}(M)$, we know by Definition 6 that for all $a, b \in A$, $b \in N(a)$ if and only if $\frac{|(V(a)\cap V(b))\cup(\overline{V(a)}\cap\overline{V(b)})|}{|F|} \geqslant \omega$. This condition involves V, F and ω. But these elements do not change from M_0 to M_1, so the same condition still applies w.r.t. M_1, so then $M_1 = M_2 \in \vec{\Box}(M)$. So M_1 is stable under the repeated update $\vec{\Box}$, so M stabilizes under it. $\Box$

Proposition 2.5 shows that the relations that form between agents with the social relation update $\Box$ stay unchanged as long as the opinion distribution of the social network stays the same. Since we do not include other social relation update mechanisms in our framework, this means that the social structures that form in our models are quite rigid. After an update with $\Box$, the space of agents A is partitioned into groups of relatively like-minded agents, and the only way for a given agent to leave or enter a group is to change her opinions. Since the only way for that to happen is via a monotonic or non-monotonic opinion update – both of which work by making agents more (or at least not less) similar to their direct neighbors – the prospects of agents switching from one $\Box$-created group to another are limited. Figure 5 shows an example model that demonstrates that such switches are still possible in multi-issue models: After one social relation update, the model contains two groups: $\{a, b, c, d\}$ and $\{e\}$. Through a subsequent opinion update, d changes her opinion, adopting all three issues in F. When another social relation update is applied, this leads to her leaving her previous social group and forming a new group with agent e.

Such group switches are not possible in single-issue models. Here, an application of the social relation update $\Box$ leads to the formation of groups that are almost completely immutable. We can make this observation precise using the concept of a *cluster split*, introduced through Definition 8.

Definition 8 (Cluster-Split Social Network Model)**.** *Given a single-issue social network model* $M = (A, N, F, V, \omega, \tau)$*, let* $A^f = \{a \in A : f \in V(a)\}$ *and* $A^{\neg f} = \{a \in A : f \notin V(a)\}$ *be the set of* f*-supporters and* $\neg f$*-supporters, respectively. We say that* M *is cluster-split if* A^f *and* $A^{\neg f}$ *are clusters of density* 1*.*

Proposition 2.6. *Let* $M = (A, N, F, V, \omega, \tau)$ *be a single-issue SNM with* $\omega > 0$*. Further, suppose that* M *is neither an* f*- nor a* $\neg f$*-consensus. Then* $M_\Box$ *is cluster-*

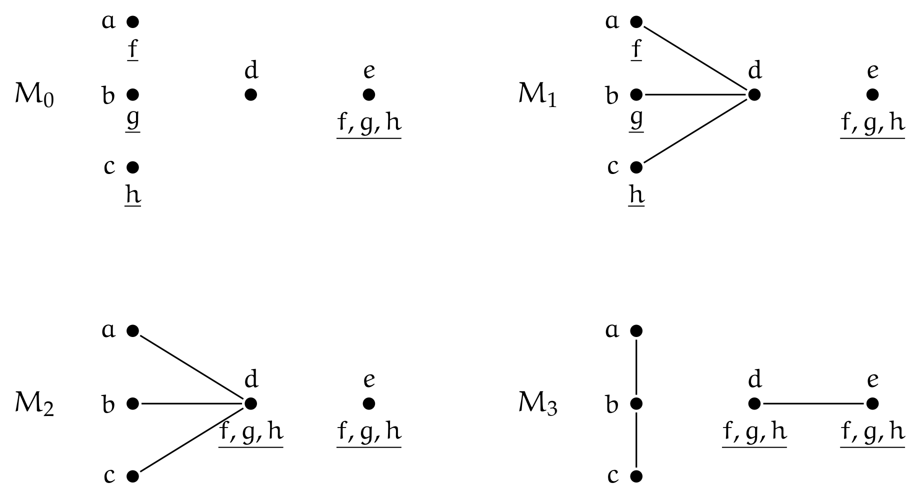

Figure 5: Four models. The initial model M_0 contains three issues $\{f, g, h\}$ and the thresholds $\tau = \frac{1}{3}$ and $\omega = \frac{2}{3}$. M_1, M_2 and M_3 are obtained by applying an update stream $\vec{\dagger} = (\square, \triangle, \square, \dots)$ to M_0. This means that M_1 is the result of applying a social relation update to M_0; M_2 is the result of applying a non-monotonic social influence update to M_1; M_2 is the result of applying a social relation update to M_1. In each model, every node is labelled with the corresponding agent's name and the issues the agent supports (the latter is underlined). Agents without underlined text next to them do not support any of the issues in F.

split, and $A^f_\square = \{a \in A : f \in V_\square(a)\}$ *and* $A^{\neg f}_\square = \{a \in A : f \notin V_\square(a)\}$ *are completely connected.*

Proof. Since V does not change with when applying the update $\square$, we know that $A^f = A^f_\square \neq \emptyset$ and $A^{\neg f} = A^{\neg f}_\square \neq \emptyset$. Take some $a \in A^f$ and some $b \in A^{\neg f}$. We know that $V(a) = F = \{f\}$ and $V(b) = \emptyset$, so $a \notin N_\square(b)$ and $b \notin N_\square(a)$. As this holds for any such pairs of agents and $A^f \sqcup A^{\neg f} = A$ (where "$\sqcup$" denotes the disjoint union), we know that A^f and $A^{\neg f}$ are clusters of density 1, so $M_\square$ is cluster-split.

To see that both sets are completely connected, consider that either $|A^f| > 1$ or $|A^f| = 1$. In the first case, take any $a, a' \in A^f$: We know that $V(a) = V(a') = F = \{f\}$, so $a \in N_\square(a')$ and $a' \in N_\square(a)$. In the second case, complete connectedness holds vacuously. An analogous argument holds for

$A^{\neg f}_{\Box}$. □

A cluster-split model is a model where the set of agents has fragmented into camps with opposing opinions that are not in contact with each other. Any single-issue SNM that is not already a consensus will become cluster-split through a single social relation update.

It is worth taking a few paragraphs to reflect on this result. We have defined the update $\Box$ to be a simple formal version of the homophilic tendency of social selection – *seeking connections with people that are similar to oneself*. It seems intuitively clear that in isolation, such a tendency can contribute to social fragmentation. To a degree – arriving at the result requires many simplifying assumptions – Proposition 2.6 seems to be the formal confirmation of this intuition. The result also links similarity-based social relation updates like $\Box$ to the topic of *polarization*:

In much of the literature on social network analysis and formal epistemology, polarization is understood as an effect of group deliberation.[12] The concept was neatly summarized by Isenberg, who states that "polarization is said to occur when an initial tendency of individual group members toward a given direction is enhanced following group discussion" [18]. Hansen, Hendricks and Rendsvig identify four factors that are required for polarization to take place [15]:

1. A set of agents,
2. an issue on which agents' degree of agreement can vary on a scale with neutral midpoint and two extreme poles,
3. a division of agents into subgroups, which are homogenous with respect to their degree of agreement relative to the midpoint, and
4. a group deliberation process in which agents are free to discuss their opinions and arguments.

Point 2 and 4 are not to be found in our formal framework: We model the attitude that agents take with respect to issues in F as a binary "agree or disagree", not as a scale with a midpoint (recall the brief discussion of this on page 7); and we do not provide a way to model deliberation between agents, which is surely a more complex process than the mechanism for social influence we attempt to capture with opinion updates. What *is* found

[12] See for instance [31] and [15].

in our framework, though, is a potential explanation of how a set of agents comes to fulfill point 3. If similarity in opinions is used as a principal factor to manage social relations, it can easily lead to the formation of fragmented, echo-chamber-like social structures – a prerequisite to polarization.

In what follows, we further spell out the consequences that a cluster-split has for the development and stabilization of a model under update streams. It turns out that almost all cluster-split models stabilize under *any* update stream – and that models with non-trivial, i.e. nonzero thresholds even stabilize in a cluster-split. We present these results below, and then return to another reflection on their significance.

Proposition 2.7. *Let* $M = (A, N, F, V, \omega, \tau)$ *be a cluster-split social network model. Then* M *stabilizes under the repeated update* $\vec{\triangle}$.

Proof. If $\tau > 0$, then $V_0 = V_1$, so $M_0 = M_1$. To see this, take any agent $a \in A$. Suppose that $a \in A_0^f$, then since this is a cluster of density 1 by assumption, we know that either $N(a) = \emptyset$ or $\frac{|N(a) \cap A^f|}{|N(a)|} \geqslant 1$. So $f \in V_1(a)$. Now, suppose that $a \in A_0^{\neg f}$. By the same argument, we know that $N(a) = \emptyset$ or $\frac{|N(a) \cap A_0^{\neg f}|}{|N(a)|} \geqslant 1$. So then $\frac{|N_0^f(a)|}{|N(a)|} = 0 < \tau$, so $f \notin V_1(a)$.

If $\tau = 0$, then M_1 is an f-consensus, so it is stable under $\vec{\triangle}$: Then it holds for all $a \in A$ that $f \in V_1(a)$, so it is likewise the case that $N_1^f(a) = N_1(a)$, so by Definition 3 we get that $f \in V_2(a)$ for any agent a, meaning that $M_1 = M_2$. □

The fact that cluster-split models stabilize under the repeated update $\vec{\triangle}$ is notable since this is not the case for all social network models, as shown in Figure 4. Having covered all repeated updates through Propositions 2.4, 2.5 and 2.7, one might wonder if cluster-split models are also stable under *any* update stream. The answer is "yes", but only for cluster-split models with certain thresholds. We prove this in Propositions 2.9, 2.10 and 2.11, but first we present an auxiliary result on stabilization of models under update streams:

Proposition 2.8. *Let* M *be a single-issue social network model. If* M *is stable under the repeated updates* $\vec{\mathbin{\triangle\!\!\!\!\!+}}$, $\vec{\triangle}$ *and* $\vec{\square}$, *then* M *is stable under any update stream* $\vec{\dagger}$ *(with* $\dagger \in \{\mathbin{\triangle\!\!\!\!\!+}, \triangle, \square\}$*).*

Proof. By contradiction. Take some such model M, some update stream $\vec{\dagger}$ and suppose that for some $M_n \in \vec{\dagger}(M)$, $M \neq M_n$. Then there must be some first update stage M_m (with $0 \leqslant m \leqslant n$) and some update $\dagger \in \{\triangle\!\!\!\triangle, \triangle, \square\}$ where the model first changed through application of $\dagger$. But then the original model M is not stable under the update $\dagger$, which contradicts our assumption. $\square$

It is important to emphasize that Proposition 2.8 works with models that are stable under repeated updates, not models that stabilize under repeated updates. Using the proposition, we can now characterize which cluster-split models stabilize under any update stream.

Proposition 2.9. *Let $M = (A, N, F, V, \omega, \tau)$ be a cluster-split social network model with $\tau > 0$ and $\omega > 0$. Then M stabilizes in a cluster-split under any update stream $\vec{\dagger}$.*

Proof. Let $\vec{\dagger}$ be some update stream. We show that no matter what the first two operations $\dagger_1, \dagger_2 \in \vec{\dagger}$ are, $M_2 \in \vec{\dagger}(M)$ is cluster-split and stable under any repeated update. Then, by Proposition 2.8, we know that M stabilizes under any update stream.

Suppose that $\dagger_1 = \triangle$. By Proposition 2.7 and since $\tau > 0$ by assumption, we then know that $M_0 = M_1$.

Suppose that $\dagger_1 = \triangle\!\!\!\triangle$. Then $M_0 = M_1$ by the same argument given in Proposition 2.7: Agents in A_0^f and $A_0^{\neg f}$ are in clusters of density 1 by assumption, so neither type of agent will be convinced to change their opinion about f through the update.

Suppose that $\dagger_1 = \square$. Then by Proposition 2.6, we know that M_1 is cluster-split with completely connected components A_1^f and $A_1^{\neg f}$. Now by Proposition 2.5 we know that M_1 is stable under further application of $\square$. And by the same arguments as given in the previous two paragraphs, we know that if $\dagger_2 = \triangle$ or $\dagger_2 = \triangle\!\!\!\triangle$, then $M_1 = M_2$. So M_1 is stable under the repeated updates $\vec{\triangle\!\!\!\triangle}$, $\vec{\triangle}$ and $\vec{\square}$. Since $M_0 = M_1$ if $\dagger_1 = \triangle\!\!\!\triangle$ or $\dagger_1 = \triangle$, this paragraph also shows that in these cases, the model will become stable under any of the repeated updates. $\square$

Proposition 2.10. *Let $M = (A, N, F, V, \omega, \tau)$ be a cluster-split social network model with $\tau = 0$ and $\omega = 0$. Then M stabilizes under any update stream $\vec{\dagger}$.*

Proof. Let $\vec{\dagger}$ be any update stream, and consider in particular $\dagger_1, \dagger_2 \in \vec{\dagger}$.

Suppose that $\dagger_1 = \square$. Then M_1 is a fully connected model. If $\dagger_2 = \square$, then $M_1 = M_2$ by Proposition 2.5. If $\dagger_2 = \triangle\!\!\!\triangle$ or $\dagger_2 = \triangle$, then M_2 is a fully connected f-consensus, which is stable under any repeated update.

Suppose that $\dagger_1 = \triangle\!\!\!\!+$ or $\dagger_1 = \triangle$. Then M_1 is an f-consensus. Thus, if $\dagger_2 = \triangle\!\!\!\!+$ or $\dagger_2 = \triangle$, $M_1 = M_2$. So M_1 is stable under $\vec{\triangle\!\!\!\!+}$ and $\vec{\triangle}$. Now, if $\dagger_2 = \square$, we know that M_2 is a fully connected f-consensus. So since $\tau = 0$ and $\omega = 0$, M_2 is stable under any repeated update.

So we know that no matter what $\dagger_1$ and $\dagger_2$ are, $M_2 \in \vec{\dagger}(M)$ is a stable under any repeated update. So by Proposition 2.8, we know that M_2 is stable under any update stream $\vec{\dagger}$, so M stabilizes under any update stream. $\square$

Proposition 2.11. *Let* $M = (A, N, F, V, \omega, \tau)$ *be a cluster-split social network model with* $\tau = 0$ *and* $\omega > 0$. *Then* M *stabilizes under any update stream* $\vec{\dagger}$.

Proof. Let $\vec{\dagger}$ be any update stream, and consider in particular $\dagger_1, \dagger_2 \in \vec{\dagger}$.

First, suppose that $\dagger_1 = \triangle\!\!\!\!+$ or $\dagger_1 = \triangle$. Then M_1 is an f-consensus. Then M_1 is stable under $\vec{\triangle\!\!\!\!+}$ and $\vec{\triangle}$. Suppose that $\dagger_2 = \square$. Then M_2 is a fully connected f-consensus, stable under $\vec{\triangle\!\!\!\!+}$ and $\vec{\triangle}$ and $\vec{\square}$.

Now, suppose that $\dagger_1 = \square$. Then M_1 is a cluster-split model with completely connected sets A_1^f and $A_1^{\neg f}$ by Proposition 2.6. M_1 is stable under $\vec{\square}$, then. Suppose that $\dagger_2 = \triangle\!\!\!\!+$ or $\dagger_2 = \triangle$: Then M_2 is an f-consensus. Then by the same argument as in the previous paragraph, M_2 stabilizes under all repeated updates. So no matter what the first updates in $\vec{\dagger}$ are, M develops into some model that is stable under all repeated updates. So M develops into a model that is stable under any update stream $\vec{\dagger}$. Thus, M stabilizes under any update stream $\vec{\dagger}$. $\square$

From a purely formal perspective, the three previous propositions show us that cluster-split models are well-behaved: Most of them will stabilize, no matter what sequence of updates is applied to them. And we can obtain such models by simply making agents update their social relations once, provided that (i) there isn't already a consensus in the model – in which case the model would stabilize under any update stream, anyway – and (ii) ω is not 0 – which incidentally is also a sufficient condition for the resulting cluster-split model to stabilize. In single-issue models, the social relation update $\square$ can thus be seen as a push towards stabilization.

However, assuming the more "social" perspective from above again, Propositions 2.9 to 2.11 are not a positive result. We find that within our framework, the possibilities of resolving the social fragmentation represented by a cluster-split are quite limited. In particular, Proposition 2.9 shows that a cluster-split model with nonzero thresholds will remain cluster-split indefinitely. We had argued that the emergence of cluster-split models through our social relation update hints at the fact that in isolation, similarity-driven social selection leads to the development of fragmented social structures. With Propositions 2.9 to 2.11, we now also know that none of the homophily-based mechanisms we defined is likely to resolve the issue of fragmentation.

2.4 Oscillation

Attentive readers will have noticed that we did not provide a general stabilization result for cluster-split models with $\omega = 0$ and $\tau > 0$, yet. This is due to the fact that not all such models stabilize under any update stream:

Proposition 2.12. *Not all cluster-split SNMs stabilize under any update stream. Consider as counterexample the model* $M = (A, N, F, V, \omega, \tau)$, *with* $A = \{a, b\}$, $N(a) = N(b) = \emptyset$, $F = \{f\}$, $V(a) = F$ *and* $V(b) = \emptyset$, $\omega = 0$ *and* $\tau = 1$*; further, define the update stream* $\vec{\dagger}^{\wr} = (\Box, \triangle, \triangle, \dots)$ *with* $\dagger_n = \triangle$ *for all* $\dagger_n \in \vec{\dagger}^{\wr}$ *with* $n \geqslant 2$. M *never stabilizes under* $\vec{\dagger}^{\wr}$.

See Figure 6 for a visual representation of the first models in $\vec{\dagger}^{\wr}(M)$.

Figure 6: The first four models in the sequence $\vec{\dagger}^{\wr}(M)$. Note that M_0 is a cluster-split model. Since $\omega = 0$, the cluster-split can be broken by application of the opinion update $\Box$. The resulting model never becomes stable.

This observation serves to shift our attention away from the special case of cluster-split models, and towards a more general phenomenon: Some social network models never stabilize. We had already seen this behavior earlier (see Figure 4), and in this subsection, we make a first step towards a characterization of *oscillation*. To do so, we need a precise formal description:

Definition 9 (Oscillation). *Given an update stream* $\vec{\dagger}$ *and a single-issue SNM* M, *we say that* M oscillates *under* $\vec{\dagger}$ *if* $\exists a \in A$ *and* $\exists M_n \in \vec{\dagger}(M)$ *such that for all* $m \geqslant n$, $f \in V_m(a) \Leftrightarrow f \notin V_{m+1}(a)$.

It is clear that a model that oscillates under $\vec{\dagger}$ never stabilizes under $\vec{\dagger}$.[13]

[13]This shows that our notion of oscillation is different from the concept of oscillation used in [2]. There, any stable model is a special type of oscillating model.

By Propositions 2.4 and 2.5, we thus know that single-issue social network models cannot oscillate under repeated updates $\vec{\triangle\!\!\!\!\triangle}$ or $\vec{\square}$. But Figure 6 clearly shows that oscillation can happen under the repeated update $\vec{\triangle}$. Proposition 2.13 outlines conditions sufficient for a social network model to oscillate under $\vec{\triangle}$.

Proposition 2.13. *Let M be a single-issue SNM with $\tau > 0$. M will oscillate under $\vec{\triangle}$ if $\exists G, H \subseteq A$ such that:*

1. G *and* H *are nonempty.*
2. *Either G is a set of f-supporters and H is a set of not-f-supporters, or the other way around*
3. *For all* $g \in G$*:* $\frac{|H \cap N(g)|}{|N(g)|} \geqslant \tau$ *and* $\frac{|H \cap N(g)|}{|N(g)|} > (1 - \tau)$.
4. *For all* $h \in H$*:* $\frac{|G \cap N(h)|}{|N(h)|} \geqslant \tau$ *and* $\frac{|G \cap N(h)|}{|N(h)|} > (1 - \tau)$.

Proof. We use induction to prove that M will behave in a certain way under the repeated update $\vec{\triangle}$. It then immediately follows that there is a "witness" to the model's oscillation.

Base case: Consider $M_0 \in \vec{\triangle}(M)$. Suppose that $\forall g \in G : f \in V_0(g)$. Then, $\forall h \in H : f \notin V_0(h)$. Then, by point 3 above, it follows that $\frac{|N_0^f(g)|}{|N(g)|} < \tau$ for any $g \in G$, and by point 4 it holds $\frac{|N_0^f(h)|}{|N(h)|} \geqslant \tau$ for any $h \in H$. So in M_1, no agent in G will support f, and every agent in H will support f. An analogous argument applies when we start by assuming that $\forall g \in G : f \notin V_(g)$.

Induction hypothesis: Consider M_{n-1} and suppose that the following three things hold:

- Either $(\forall g \in G : f \in V_{n-1}(g) \wedge \forall h \in H : f \notin V_{n-1}(h))$ or $(\forall g \in G : f \notin V_{n-1}(g) \wedge \forall h \in H : f \in V_{n-1}(h))$
- $\forall i \in (G \cup H) : f \in V_{n-1}(i) \Rightarrow \frac{|N_{n-1}^f(i)|}{|N(i)|} < \tau$
- $\forall i \in (G \cup H) : f \notin V_{n-1}(i) \Rightarrow \frac{|N_{n-1}^f(i)|}{|N(i)|} \geqslant \tau$

Induction step: Now consider M_n. By the induction hypothesis, we know that either G contains only f-supporters and H contains only $\neg f$-supporters, or the other way around. Suppose the first case holds: Then, since N did not change from M_0, we can again use point 3 and 4 from above to deduce that

$\frac{|N_n^f(g)|}{|N(g)|} < \tau$ for any $g \in G$, and $\frac{|N_n^f(h)|}{|N(h)|} \geqslant \tau$! And once again, an analogous argument applies for the other possible case.

Now any agent from $H \cup G$ can be used to prove that M oscillates under the repeated update $\vec{\triangle}$.

$\square$

A social network model that satisfies the conditions given above contains two groups of agents with opposing views which are heavily connected to each other – so much so that either group has enough influence on each member of the other group to make them switch their opinion. In Figure 6, for instance, agents a and b have opposing views with respect to the issue f, and from M_1 on, both are each other's only neighbor.

There are parallels between these conditions and the conjecture put forth by Liu, Seligman and Girard in [20] about characterizing communities "in flux". There, the authors focus on mechanisms and dynamics of belief change in social networks, modelling three possible belief states per proposition for each agent and explicitly defining different types of influences that neighbors have on each other. The formal setting is thus quite different from the one in this thesis. Still, they suggest that communities that are in flux, i.e. never become stable with respect to the agents' belief changes are characterized by the following condition: *For every agent in the community, either (i) all of her neighbors believe* p *and all of her neighbors' neighbors believe* $\neg p$*, or (ii) all of her neighbors believe* $\neg p$ *and all of her neighbors' neighbors believe* p.[14] For any agent, all of that agent's neighbors agreeing on belief or disbelief of a proposition is sufficient to induce a belief change in the agent. It is easy to see that this expresses a similar idea as Proposition 2.13: A social network model will not stabilize if it contains groups of agents with differing attitudes which are also in a position to influence each other's attitudes to a sufficient degree.

However, the conditions we give above are not necessary. In other words, there are oscillating social network models that do not satisfy them. One such model is shown in Figure 7: In M_0, there is a set of $\neg f$-supporters, namely $\{a\}$, and there are a number of available sets of f-supporters. But $\{a\}$ makes up less than a τ fraction of the neighbors of any available f-supporter. The model thus violates condition 3 or 4 of Proposition 2.13, and yet it oscillates. This shows that a precise characterization of oscillating models

[14] Paraphrased from the formally expressed "$(FBp \wedge FFB\neg p) \vee (FB\neg p \wedge FFBp)$", see [20, p. 2411].

must make use of weaker conditions than the ones we gave above. We do not pursue this line further here, but we will return to it in the conclusion.

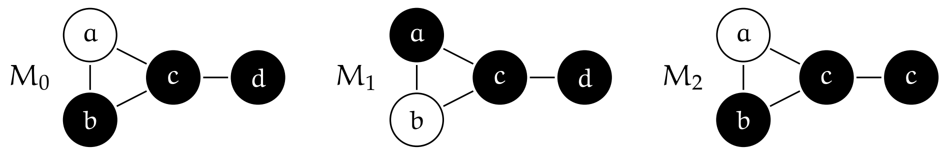

Figure 7: A sequence of single-issue SNMs under the repeated non-monotonic opinion update $\vec{\triangle}$ with $\tau = \frac{2}{3}$. The model oscillates, but it does not fulfill the conditions given in Proposition 2.13.

3 Preliminaries: Knowledge and Epistemic Logic

The framework presented in Section 2 revolves around modelling the two homophilic tendencies of social selection and social influence. While the opinion updates and social relation updates given in Definition 3 and 6 are plausible representations of these tendencies, they are also quite simple. In particular, they sidestep the topic of knowledge: In our model, we tacitly assumed that for a given group of agents, everyone knows exactly what opinions anybody else has, and everyone knows the exact structure of the network, i.e. who is friends with whom. This seems to be an unrealistically strong assumption.

In the following section, we set ourselves the task to amend this and explicitly introduce an epistemic dimension to the model; this will allow us to define more fine-grained versions of the updates introduced so far, and study their characteristics and interactions. To this end we will make use of techniques from epistemic logic and dynamic epistemic logic. We will use this section to introduce these techniques and the ideas behind them.

Representing Knowledge with Modal Logic

Epistemic logics are frameworks to reason about attitudes such as knowledge and belief.[15] Further developments of the framework have made it possible to extend the modelling to different epistemic concepts, including justification, evidence *et cetera*. Both in "classical" epistemic logic and dynamic epistemic logic, this is usually achieved using formal techniques from modal logic, specifically Kripke frames.[16]

A Kripke frame is a pair (W, R), where W is a nonempty set of states and R is a binary relation on the set W. Given a set of propositional atoms Prop, we can define a modal language by specifying the set of well-formed formulas φ of the language, using the Backus-Naur form:

$$\varphi ::= p \in \mathsf{Prop} \mid \bot \mid \neg\varphi \mid \varphi \vee \psi \mid \Box\varphi$$

Kripke frames can then be used to interpret the well-formed formulas of our modal language. To do so, we can extend any Kripke frame $\mathbb{F} = (W, R)$ to a Kripke model $\mathbb{M} = (W, R, V)$ by adding a valuation function $V : \mathsf{Prop} \to$

[15][17] is considered to be a foundational text for epistemic logic; an overview is given in [13].

[16]For an introduction to dynamic epistemic logic, see [6] or [11].

$\wp(W)$ that maps each propositional atom to a set of states. Given a Kripke model $\mathbb{M}$, we can introduce a notion of *satisfaction*, formally denoted by "$\models$": For any formula φ of our language and any state $w \in W$,

$$
\begin{array}{lll}
\mathbb{M}, w \models p & \text{iff} & w \in V(p) \\
\mathbb{M}, w \models \bot & & \text{never} \\
\mathbb{M}, w \models \neg\varphi & \text{iff} & \text{not } \mathbb{M}, w \models \varphi \\
\mathbb{M}, w \models \varphi \vee \psi & \text{iff} & \mathbb{M}, w \models \varphi \text{ or } \mathbb{M}, w \models \psi \\
\mathbb{M}, w \models \Box\varphi & \text{iff} & \text{for all } v \text{ such that } Rwv \text{ we have } \mathbb{M}, v \models \varphi
\end{array}
$$

The conditions given for the propositional connectives are standard; it is the satisfaction condition for the modality $\Box$ that makes the semantics interesting: $\Box\varphi$ holds at a state w if and only if φ holds at all states that can be reached *from w via* the relation R.[17]

Mathematically, Kripke frames and models are interesting structures by themselves, and nothing forces us to give a further interpretation of what formulas like $\Box\varphi$ or $\neg\Box(\varphi \vee \psi)$ *mean*. But from a philosophical perspective, an epistemic reading is quite natural: We can take the set of states W as a set of *possible worlds*, and understand R as some relation of *epistemic accessibility*. The modality $\Box$ can then be interpreted as a formal characterization of knowledge; to distinguish this interpretation from the philosophically neutral one described so far, we will write "$K\varphi$" instead of "$\Box\varphi$". Recapitulating the satisfaction conditions given just above, we get: A proposition φ is known at a possible world w in a model $\mathbb{M}$ – formally, $\mathbb{M}, w \models K\varphi$ – if and only if φ is true in all worlds that are epistemically accessible from w.

We use an example to illustrate how this works.[18] Picture the following scenario: Anna is sitting in her office in Amsterdam, and she is wondering about the weather. Specifically, she is wondering if it is windy or raining. There are four possible situations she could find herself in: It could be both raining and windy, it could only be raining, it could only be windy, or it could be neither raining nor windy. Figure 8 shows how we can represent these situations as possible worlds, using p to represent the proposition "It is raining", and q to represent "It is windy".

Anna's office has a window, and it is overlooking the canal. So if it is raining, she can see this by looking at the water surface. But there are

[17]Using the propositional connectives and modality contained in our language, we can define other connectives and modalities in the usual way, i.e. $\varphi \to \psi ::= \neg\varphi \vee \psi$ or $\Diamond\varphi ::= \neg\Box\neg\varphi$.

[18]In our explanation, we follow the overall structure of [25].

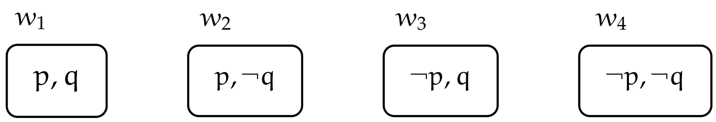

Figure 8: A set of states $W = \{w_1, w_2, w_3, w_4\}$ with a valuation V on them. The states are represented as boxes with round edges, labeled above with their name; V is represented in the figure by labelling states in W with the atomic propositions that hold or do not hold at them.

no trees outside her office, so she can not assess if it is windy by looking outside. This means that Anna can distinguish situations with rainy weather from those where it is dry – but situations that only differ in wind level are indistinguishable for her. This too can be represented in our model, by adding a relation between possible worlds. As the relation represents Anna's situation, we name it R_a. In Figure 9, R_a is depicted through labeled arrows connecting possible worlds. We can see that, for Anna, w_3 and w_4 are indistinguishable; w_2 and w_3 are not indistinguishable; every world is indistinguishable from itself *et cetera*.

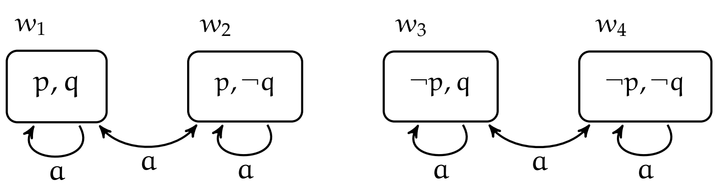

Figure 9: A Kripke model $\mathbb{M} = (W, R_a, V)$. W and V remain unchanged from Figure 8; the relation R_a is represented as double-headed arrows between states, labeled with the letter a.

Figure 9 depicts a Kripke model, and we can now ask what Anna does or does not know in this model. The answer depends on the state we consider. For instance, in state w_2, we get:

1. $\mathbb{M}, w_2 \vDash K_a p$. Anna knows that it is raining.[19]

[19] Note that we write "$K_a\varphi$" in this example to make explicit that we are quantifying over R_a-accessible states and that we are thus modelling *Anna's* knowledge.

2. $\mathbb{M}, w_2 \vDash \neg K_a q \wedge \neg K_a \neg q$. Anna does not know that it is windy, and she does not know that it is not windy.

3. $\mathbb{M}, w_2 \vDash K_a K_a p$. Anna knows that she knows that it is raining.

4. $\mathbb{M}, w_2 \vDash \neg K_a \neg q$. Anna considers it epistemically possible that it is windy.

By spelling out the available scenarios from our example as possible worlds, and adding an epistemic accessibility relation, we receive a precise characterization of Anna's knowledge.

In this example, we have interpreted the relation R_a as indistinguishability, and we have correspondingly modeled the relation R_a to be reflexive, symmetric and transitive – an equivalence relation. It is one of the characteristic features of modal logic that any choice of properties for R_a will commit us to accept a number of principles involving K_a, our notion of knowledge.[20] Specifically, the following principles are valid at any state of any Kripke model with a relation R_a that is reflexive, symmetric, and transitive.

$$\begin{array}{lc} \textbf{K} & K_a(\varphi \to \psi) \to (K_a\varphi \to K_a\psi) \\ \textbf{T} & K_a\varphi \to \varphi \\ \textbf{4} & K_a\varphi \to K_aK_a\varphi \\ \textbf{5} & \neg K_a\varphi \to K_a\neg K_a\varphi \end{array}$$

These principles spell out properties of knowledge as modelled by the operator K_a: **K** commits us to closure of knowledge under *modus ponens*, **T** commits us to knowledge being factual, and **4** and **5** commit us to knowledge being positively and negatively introspective: If an agent knows something, she knows that she knows it; and if an agent does not know something, she knows that she does not know it.

The notion of knowledge that is modelled by frames with equivalence relations is thus quite strong, and many arguments have been given for modifying it by requiring the relation R_a to satisfy other properties.[21] Nevertheless, we will accept all principles listed above and model epistemic

[20] *Modal correspondence theory* is the study of the connection between frame properties, i.e. properties of a modal relation, and frame validities, i.e. formulas or formula schemes that hold at any state of any model based on a frame with the relation that satisfies these properties. An introduction to this can be found in [9].

[21] [7] contains a formal setting that accommodates both this strong notion of knowledge and other epistemic and doxastic attitudes.

accessibility as an equivalence relation in this thesis. This is convenient from a modelling standpoint, as an equivalence relation partitions the state space W and is thus easy to work with. Further, the choice will make our results more compatible with other fields that work with these assumptions. For instance, foundational work on epistemic game theory done by Aumann[22] makes use of this notion of "hard" knowledge.

We now return to our example of Anna and the weather to illustrate a final feature of epistemic logic. So far we have worked with Kripke frames that contain one relation. By including multiple relations on the state space instead, we can model the knowledge of multiple agents in a situation. For context, picture Anna's colleague, Bob. Bob has an office in the same building, but it is not overlooking a canal. Instead, his window is right in front of a big tree. This means that Bob can not see if it is raining or not (he can't quite make out raindrops with all the leaves as the backdrop), but he can tell if it is windy or not by seeing if the tree's branches are bending. We can add this information into our model from Figure 9 by giving a second relation R_b. The resulting model is shown in Figure 10.

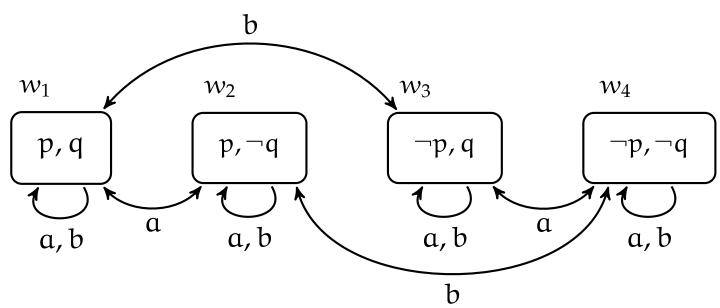

Figure 10: Another Kripke model $\mathbb{M}' = (W, R_a, R_b, V)$. W and V remain unchanged from Figures 8 and 9. The relations R_a and R_b are both depicted as labelled double-headed arrows between states.

Once again, we can now list things Anna and Bob know and don't know at a state, using two modalities K_a and K_b that correspond to the two indistinguishability relations of the model:

1. $w_2 \vDash K_b \neg q \wedge \neg K_a \neg q$. In w_2, Bob knows that it is not windy and Anna does not know that it is not windy.

[22] Originally introduced through [1]. For a contemporary introduction, see [24].

2. $w_2 \vDash K_a(K_b q \lor K_b \neg q)$. In w_2, Anna knows that Bob knows whether it is windy or not.

In the following section, we will use the ideas presented here to represent agents information about the social network they are in. This means that agents can be certain or uncertain, both regarding the opinions other agents have and the relational structures between them. We focus on a model theoretic account, defining a state space and assigning opinion distributions and neighborhood structures to each state. Our method of combining social network models with epistemic logic draws from the work presented in [4]. In this thesis, we use it to explore

1. what ways there are to define homophily-based updates of social networks when epistemic uncertainty is present,
2. how these different ways of updating compare to each other and
3. how agents can learn about the network they are in.

4 Epistemic Social Network Models

We begin by introducing a new type of model, supplementing the non-epistemic social network models of Definition 1 with a state-structure and an accessibility relation between states.

Definition 10 (Epistemic Single-Relation SNM)**.** *An epistemic single-relation social network model* M *is a tuple* $(W, A, N, F, V, \omega, \tau, \{\sim_a\}_{a \in A})$ *where*

- W *is a finite, non-empty set of states,*
- A *is a finite, non-empty set of agents,*
- $N : W \times A \to \wp(A)$ *is the neighborhood function, assigning a set of neighbors* $N(w, a)$ *to each pair of* $w \in W$ *and* $a \in A$ *such that for any* $w \in A$ *and all* $a, b \in A$*:*
  - $a \notin N(w, a)$ *(irreflexivity) and*
  - $a \in N(w, b)$ *if and only if* $b \in N(w, a)$ *(symmetry);*
- F *is a finite, non-empty set of issues;*
- $V : W \times A \to \wp(F)$ *is a valuation function, assigning to each pair of a state* w *and an agent* a *a set of issues the agent accepts or supports in* w*;*
- $\omega, \tau \in \mathbb{Q}$ *are two rational numbers s.t.* $0 \leqslant \omega \leqslant 1$ *and* $0 \leqslant \tau \leqslant 1$*, interpreted as similarity threshold and influencability threshold;*
- $\sim_a \subseteq W \times W$ *is an equivalence relation between states in* W *for each agent* $a \in A$.

For each agent a, the equivalence relation $\sim_a$ represents indistinguishability: If $w \sim_a v$, then agent a considers the state of affairs represented by state v as an epistemic possibility in state w. Through this, we introduce the possibility of agents being uncertain about the state of the social network they are in. To keep the framework as general as possible, we have not included any constraints on the equivalence relation $\{\sim_a\}_{a \in A}$ in the definition above. But for our purposes, it will be plausible to require all agents to have a minimal level of knowledge about the network. To make this precise, we introduce the notions of n*-reachability* and *sight*, following the way they are introduced for opinion models in [4].

Definition 11 (n-reachable)**.** *Let* M *be an epistemic SNM and let* $n \in \mathbb{N}$*. Define* $\underline{N}^n : W \to A \to \wp(A)$ *as follows, for any* $w \in W$ *and any* $a, b, c \in A$*:*

- $\underline{N}^0(w)(a) = \{a\}$
- $\underline{N}^{n+1}(w)(a) = \underline{N}^n(w)(a) \cup \{b \in A : \exists c \in \underline{N}^n(w)(a) \text{ and } b \in N(w, c)\}$

If $b \in \underline{N}^n(w)(a)$*, we say that* b *is* n-reachable *from* a *in* w*.*

Given a model, a state w in W and an agent a, another agent is considered to be n-reachable from a in w if one can get to it in n steps, starting at a and tracing the social relation N.

Definition 12 (Sight (n, m) model)**.** *An epistemic SNM of sight* (n, m) *is a model such that, for* $n, m \in \mathbb{N}$ *and for any* $a, b \in A$ *and any* $w, v \in W$*:*

- $n \leqslant m + 1$
- *If* $w \sim_a v$ *and* $b \in \underline{N}^n(w)(a)$*, then* $V(w, b) = V(v, b)$*.*
- *If* $w \sim_a v$ *and* $b \in \underline{N}^m(w)(a)$*, then* $N(w, b) = N(v, b)$*.*

The sight of a model represents how much any agent in the model knows about the network surrounding them: The higher n is, the further all agents can "see" with respect to others' opinions on the issues in F; the higher m is, the further all agents can see with respect to the network's structure, i.e. who has which agents in their set of neighbors.

Four details of the definition are worth commenting on. First, since $n, m \geqslant 0$, agents are guaranteed to know their own opinions and who their neighbors are for any sight (n, m). Second, in an epistemic SNM of sight (n, m), all agents know the distribution of opinions and the network structure *at least* up to n or m, respectively – but some agents might know even more, i.e. see further than that. Third, the requirement that $n \leqslant m + 1$ makes it so that agents will not know other agents' opinions without also knowing how they are socially related to them. This reflects the idea that within our models, the social relation N is the primary channel of information about other agents.

Fourth, it should be noted that even in models with sight $(0, m)$, social influence can still take place. Having $n = 0$ should be understood as agents not (generally) being *certain* about their neighbors' opinions, not as them having no information about their opinions whatsoever. Depending on the behavior one wants to model, it can be sensible to require that an agent is absolutely certain about the state of the network before she acts on it, but it can also be sensible to set the bar for social influence lower than this, requiring e.g. belief in the possibility that a set of conditions is fulfilled. In the following sections, we will spell out two different ways of including knowledge in update mechanisms for relations and opinions. It will turn out that these mechanisms lead to different adoption behaviors in exactly those epistemic SNMs that have a sight of $(0, m)$, i.e. those models where agents are not certain about their neighbors opinions.

4.1 Cautious Epistemic Updates

In this section and the following section, our definitions of epistemic relation updates will require us to express that the "overlap" in two agents' opinions is larger than the threshold ω. We will use a shorthand to express this similarity condition to make the coming definitions less unwieldy:

$$\mathrm{Smlty}(w)(a, b) := (V(w, a) \cap V(w, b)) \cup (\overline{V(w, a)} \cap \overline{V(w, b)})$$

With this, we now have everything we need to define and evaluate our first pair of updates, starting with an epistemic version of our non-monotonic opinion update from Definition 3.

Definition 13 (Cautious Epistemic Opinion Update)**.** *Given an epistemic social network model* M*, applying the* (n, m) *sight epistemic opinion update* $\triangle C$ *results in a model* $M_{\triangle C} = (W, A, N, F, V_{\triangle C}, \omega, \tau, \{\sim_a^{\triangle C}\}_{a \in A})$*, where for any* $w, w' \in W$*, any* $a \in A$ *and all* $f \in F$*:*

- $f \in V_{\triangle C}(w, a)$ *iff* $\forall v \sim_a w : \left\{ \begin{array}{ll} f \in V(v, a), & \textit{if } N(v, a) = \emptyset, \\ f \in F, & \textit{if } \tau = 0, \\ \frac{|N_f(v,a)|}{|N(v,a)|} \geqslant \tau, & \textit{otherwise.} \end{array} \right\}$,
- $w \sim_a^{\triangle C} w'$ *if and only if*
  - *(i)* $w \sim_a w'$ *and*
  - *(ii)* $\forall b \in \underline{N}^n(w)(a) : V_{\triangle C}(w, b) = V_{\triangle C}(w', b)$.

Where $N_f(w, a) = \{b \in A : f \in V(w, b)\}$

The definition above consists of two points. The first point spells out the conditions under which agents will support an issue in F after the update:

If they know they have no neighbors, they will stick to the opinion they had before the update – unless $\tau = 0$, in which case all agents are so easily influenced that they adopt all issues in F after one update step. Otherwise, they will support an issue only if they are *certain* that at least a τ-fraction of their neighbors supports it, too. With this type of update, agents are modeled as reluctant to support issues in F: They only start doing so if they are certain (enough of) their neighbors are doing it, and they stop as soon as they lose this certainty. Figure 11 illustrates this, showing how an agent that is unsure of the position of her neighbor errs on the side of caution and stops supporting all issues in F.

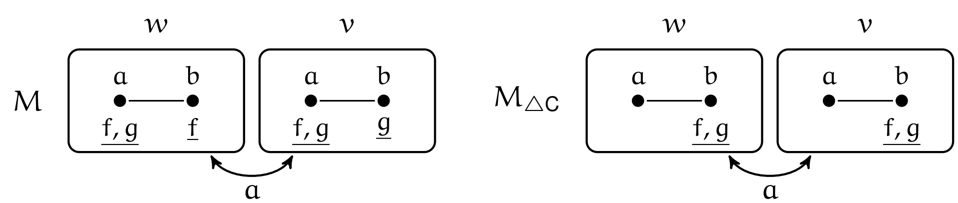

Figure 11: A succession of models with $\tau = 1$, $F = \{f, g\}$ and initial sight $(0, 1)$. Both models have two states, w and v. These states are represented as rounded rectangles enclosing the agents a and b. The equivalence relations are shown as two-headed arrows between states, labeled with the corresponding agent's name; reflexive equivalence arrows for all agents are omitted for visual clarity. Each agent is represented as a black node, labeled with her name and the issues she supports in that state; the latter are underlined. If an agent is not labelled with some issue in F, she does not support it. On all occasions that follow, we will represent multi-issue epistemic SNMs in this manner.

Cautious epistemic opinion updates can be understood to model an opinion adoption behavior that arises when adopting the issues in F comes with a cost or a risk. Perhaps the issues in F represent technologies, and supporting (read: using) some technology $f \in F$ incurs costs that only "pay off" if a significant portion of neighbors does so, too. Or perhaps – taking inspiration from [14] and [4, p. 524] – we are in a sociopolitical sphere, and supporting an issue $g \in F$ is equivalent to supporting a revolutionary movement or a riot. In such a case, agents will want to be certain that a sufficient portion of their neighbors also supports before they "join the movement", lest they find themselves part of a group that was too small to be successful and get punished.

Returning to the formal details, the second point of Definition 13 ensures that the update mechanism is *sight-preserving*: If a (n, m) sight epistemic opinion update is applied to a model M with sight (n, m), the result is a model $M_{\triangle C}$ with the same sight.

Proposition 4.1 ($\triangle C$ is sight-preserving). *Let M be a model of sight (n, m). Then, through applying $\triangle C$ with sight (n, m), the resulting model $M_{\triangle C}$ has (n, m), too.*

Proof. Suppose not: Then one of two things is true.

1. There are two states $w, v \in W$, two agents b, c and an issue f such that: $w \sim_b^{\triangle C} v$ and $c \in \underline{N}^n(w)(b)$ but $V_{\triangle C}(w, c) \neq V_{\triangle C}(v, c)$. This directly contradicts the second point of Definition 13.
2. There are two states w, v and two agents b, c such that $w \sim_b^{\triangle C} v$ and $c \in \underline{N}^m(w)(b)$ and $N(w, c) \neq N(v, c)$. By Definition 13 we then know that $w \sim_b v$, and since N did not change with the update, we know that $c \in \underline{N}^m(w)(b)$ and $N(w, c) \neq N(v, c)$. So M was not a model of sight (n, m), which contradicts our initial assumption.

□

Note that this result does not mean that models cannot *increase* in sight with updates – in fact, we will give examples of update sequences on models where this does happen in Subsection 4.4. In this thesis, whenever there is a sequence of epistemic social network models, we take the sight of the first model to determine the *initial sight* of the model sequence; each update within this sequence is then also performed with the initial sight.

Next, we deal with the counterpart of $\triangle C$, an epistemic version of a relation update.

Definition 14 (Cautious Epistemic Relation Update). *Given an epistemic SNM M, applying the (n, m) sight epistemic relation update $\square C$ results in a model $M_{\square C} = (W, A, N_{\square C}, F, V, \omega, \tau, \{\sim_a^{\square C}\}_{a \in A})$ where, for any $a, b \in A$ and any $w, w' \in W$:*

- $b \in N_{\square C}(w, a)$ *iff* $a \neq b$ *and* $\forall v \sim_a w : \frac{|Smlty(v)(a,b)|}{|F|} \geqslant \omega$ *and* $\forall u \sim_b w : \frac{|Smlty(u)(a,b)|}{|F|} \geqslant \omega$,

- $w \sim_a^{\Box C} w'$ *if and only if*

  (i) $w \sim_a w'$ *and*

  (ii) $\forall c \in \underline{N}^n_{\Box C}(w)(a) : V(w, c) = V(w', c)$ *and*

  (iii) $\forall d \in \underline{N}^m_{\Box C}(w)(a) : N_{\Box C}(w, d) = N_{\Box C}(w', d)$.

Where $\underline{N}^n_{\Box C}(w)(a)$ *is the set of n-reachable agents from agent* a *in state* w *known from Definition 11, but constructed using* $N_{\Box C}$ *from the updated model* $M_{\Box C}$.

The first point of Definition 14 makes it so that after an update with $\Box C$, agents will only be connected as neighbors to other agents if they are certain to agree with each other on at least an ω fraction of all issues. Similar to $\triangle C$-updates, this renders agents reluctant to engage in or uphold neighborhood: If they don't know for sure that the similarities to another agent are big enough, they will not be neighbors with them.

It is worth noting that the definition requires *de dicto* knowledge of this similarity, not *de re* knowledge.[23] This means that it suffices for the agents to know that there is a sufficiently large number of issues they agree upon, possibly without knowing *which* issues they agree upon exactly. Figure 12 illustrates how this distinction works in the present setting: In the depicted model M, agent a knows *de dicto* that her and agent b share half of their opinions – she knows that b agrees with her on one of two issues; but a does not know *de re* that her and b share half their opinions – she does not know that her and b agree on issue f, and she also does not know that her and b agree on issue g. For the cautious epistemic relation update the former type of knowledge is sufficient for both agents to connect. In this, the update mechanism we use is similar to the ones proposed in [30, p. 384] and [4, p. 505].

Like in the latter, it would be formally possible to define a *de re* version of our cautious epistemic relation update. This would be an "extra cautious" relation update, requiring that both agents are certain they agree on a specific and sufficiently large subset of the set of issues F before connecting. We choose not to explore this in this thesis, and instead continue with the less strict version as given by Definition 14.

We now return to that very definition. Taking another look at its first point, we can see that within our framework both agents are needed to enter

[23] For an introduction to the distinction, see the supplement "The De Re/De Dicto Distinction" to [23].

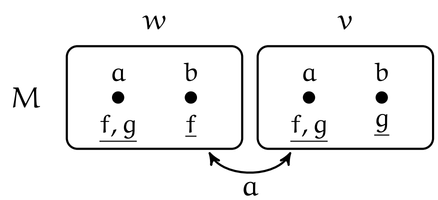

Figure 12: A social network model with $F = \{f, g\}$.

or maintain a neighborhood relation, as they both need to be certain of their similarity in opinions. On a formal level, the symmetry in the first condition ensures that after an update with $\Box C$, the relation $N_{\Box C}$ is again symmetric.

The second point of Definition 14 ensures that, like $\triangle C$, the update $\Box C$ preserves sight.[24]

Proposition 4.2 ($\Box C$ is sight-preserving)**.** *Let* M *be a model of sight* (n, m)*. Then, by updating with* $\Box C$ *with sight* (n, m)*, we obtain the model* $M_{\Box C}$ *which also has sight* (n, m)*.*

Proof. Take any $a, b \in A$ and any $w, v \in W$. Assume that $w \sim_a^{\Box C} v$ and $b \in \underline{N}^n_{\Box C}(w)(a)$. Then, by Definition 14, we know that $V(w, b) = V(v, b)$.

Now, take any $a, b \in A$ and any w, v and assume that $w \sim_a^{\Box C} v$ and $b \in \underline{N}^m_{\Box C}(w)(a)$. Then $N_{\Box C}(w)(b) = N_{\Box C}(v, b)$, also by Definition 14.

So following from Definition 12, $M_{\Box C}$ is a model of sight (n, m). $\Box$

4.2 Eager Epistemic Updates

The second type of epistemic updates we introduce, so-called *eager* updates, differs in the previous pair only in a change of quantifiers. We start by defining the eager epistemic opinion update $\triangle E$.

Definition 15 (Eager Epistemic Opinion Update)**.** *Given an epistemic social network model* M*, applying the* (n, m) *sight epistemic opinion update* $\triangle E$ *results in a model* $M_{\triangle E} = (W, A, N, F, V_{\triangle E}, \omega, \tau, \{\sim_a^{\triangle E}\}_{a \in A})$*, where for any* $w, w' \in W$*, any* $a \in A$ *and all* $f \in F$*:*

[24] In contrast to Definition 14, Definition 13 contains the additional condition (ii) for states to be indistinguishable for an agent after the update. This is necessary because through a cautious epistemic relation update, agents might newly connect to other agents whose opinions they were not sure about before.

- $f \in V_{\triangle E}(w, a)$ *iff* $\exists v \sim_a w : \left\{ \begin{array}{ll} f \in V(v, a), & \textit{if } N(v, a) = \emptyset, \\ f \in F, & \textit{if } \tau = 0, \\ \frac{|N_f(v,a)|}{|N(v,a)|} \geqslant \tau, & \textit{otherwise.} \end{array} \right\}$,

- $w \sim_a^{\triangle E} w'$ *if and only if*

  (i) $w \sim_a w'$ *and*

  (ii) $\forall b \in \underline{N}^n(w)(a) : V_{\triangle E}(w, b) = V_{\triangle E}(w', b)$ *and*

Where $N_f(w, a) = \{b \in A : f \in V(w, b)\}$

As before, the second point of the definition ensures that the update operation is sight-preserving. The first point spells out the conditions for an agent to adopt an opinion. At any state in W, agents will keep their opinions if they have no neighbors. If they have neighbors, they will support any given issue unless they are certain that less than a τ fraction of their neighbors supports it, too. In contrast to the cautious epistemic opinion update $\triangle C$, with $\triangle E$, agents are eager to support issues in F, and will only refrain from doing so if they are certain that few (i.e. less than τ) of their neighbors do the same.

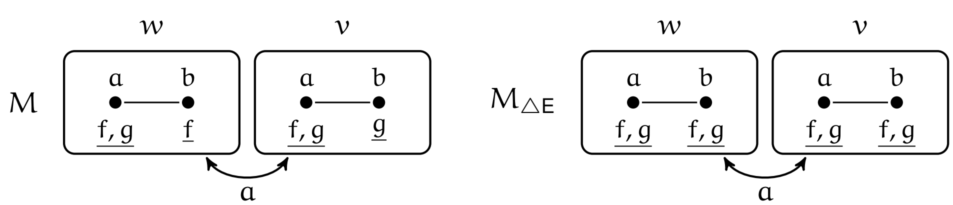

Figure 13: A succession of models, with $\tau = 1$, $F = \{f, g\}$ and initial sight $(0, 1)$.

Figure 13 shows how this changes the opinion adoption behavior of agents: Starting from the same model that was shown in Figure 11, after an eager epistemic opinion update with sight $(1, 0)$, agent a continues to support both f and g, even though she is not sure which of these agent b supports.

By changing the universal quantifier to an existential quantifier in the definition, we also obtain an eager version of the epistemic relation update, defined formally below.

Definition 16 (Eager Epistemic Relation Update). *Given an epistemic SNM* M*, applying the* (n, m) *sight epistemic relation update* $\Box E$ *results in a model* $M_{\Box E} = (W, A, N_{\Box E}, F, V, \omega, \tau, \{\sim_b^{\Box E}\}_{a \in A})$ *where, for any* $a, b \in A$ *and any* $w, w' \in W$*:*

- $b \in N_{\Box E}(w, a)$ *iff* $a \neq b$ *and* $\exists v \sim_a w : \frac{|Smlty(v)(a,b)|}{|F|} \geqslant \omega$ *and* $\exists u \sim_b w : \frac{|Smlty(u)(a,b)|}{|F|} \geqslant \omega$,
- $w \sim_a^{\Box E} w'$ *if and only if*
  - *(i)* $w \sim_a w'$ *and*
  - *(ii)* $\forall c \in \underline{N}_{\Box E}^{n}(w)(a) : V(w, c) = V(w', c)$ *and*
  - *(iii)* $\forall d \in \underline{N}_{\Box E}^{m}(w)(a) : N_{\Box E}(w, d) = N_{\Box E}(w', d)$.

Where $\underline{N}_{\Box E}^{n}(w)(a)$ *is the set of n-reachable agents from agent* a *in state w known from Definition 11, but constructed using* $N_{\Box E}$ *from the updated model* $M_{\Box E}$.

After an update with $\Box E$, two agents will be neighbors if neither of them is certain that they are not similar enough in opinions. Figure 14 illustrates how this influences the network change after an update: After applying the eager epistemic relation update to M, agents a and b only connect to each other in states where both of them are certain that they agree on the issue f.

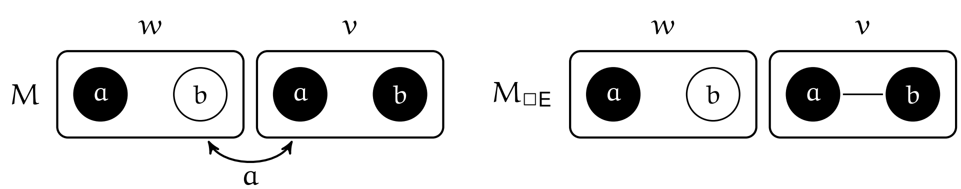

Figure 14: A succession of models, with $\omega = 1$, $F = \{f\}$ and initial sight $(0, 1)$.

Note that after the update, it is not the case anymore that $w \sim_a^{\Box E} v$; agent a has learned to distinguish between both states. We will come back to the topic of learning in epistemic SNMs – and to the model shown in Figure 14 – in Subsection 4.4. Before we do so, it will be helpful to compare our epistemic update types on a more general level.

4.3 Comparing Updates

How does the epistemic dimension affect the spread of behaviors and the development of relations in social network models? In the present section, we will compare epistemic and non-epistemic SNMs in their development under update streams. To be able to do so, we will give a definition of *update stream* that is slightly modified with respect to the definition given earlier in Section 2. Further, we will introduce the notion of *state-generated models*.

Definition 17 (Epistemic Update Stream)**.** *An epistemic update stream $\vec{\dagger}$ is an infinite sequence of updates $(\dagger_n)_{n\in N}$ (with $\dagger \in \{\triangle\mathsf{C}, \Box\mathsf{C}, \triangle\mathsf{E}, \Box\mathsf{E}\}$). A repeated epistemic update is an epistemic update stream that only consists of one type of update.*

An epistemic SNM M stabilizes under an update stream $\vec{\dagger}$ if there exists some $n \in N$ such that $M_n = M_m$ for all $m > n$ and with $M_n, M_m \in \vec{\dagger}(M)$. A model M is stable under an update stream $\vec{\dagger}$ if $M_0 = M_n$ for all $n \geqslant 0$, i.e. if the point of stabilization has already been reached.

Definition 18 (State-Generated Social Network Model (adapted from [4]))**.** *Let $M = (W, A, N, F, V, \omega, \tau, \{\sim_a\}_{a\in A})$ be an epistemic social network model. For any state $w \in W$, the state-generated, non-epistemic social network model $M(w) = (A, N^{M(w)}, F, V^{M(w)}, \omega, \tau)$ is defined as (for all $a \in A$):*

$$N^{M(w)}(a) = N(w, a)$$
$$V^{M(w)}(a) = V(w, a)$$

4.3.1 Opinion Change

Using state-generated models allows us to compare the development of non-epistemic SNMs and epistemic SNMs in a formally precise way: Given an epistemic social network model, we can generate a non-epistemic model from one of its states and then compare both models' behavior under update streams. For instance, this technique allows us to prove that the epistemic opinion updates $\triangle\mathsf{C}$ and $\triangle\mathsf{E}$ differ from the non-epistemic update $\triangle$ regarding the amount of agents that adopts an issue $f \in F$ after an update is applied.

Proposition 4.3. *Let M be an epistemic SNM, and let $M_{\triangle\mathsf{C}}$ be the model obtained by updating with the cautious epistemic opinion update $\triangle\mathsf{C}$. Then, let $M_{\triangle\mathsf{C}}(w)$ be model state-generated from $M_{\triangle\mathsf{C}}$ for some state $w \in W$.*

Considering again the initial model M*, let* $M(w)$ *be the state-generated model from* w*, and let* $M(w)_{\triangle}$ *be the result of updating with the non-epistemic opinion update* $\triangle$*.*

Now, we define $A^f_{\triangle C} = \{b \in A : f \in V^{M_{\triangle C}(w)}\}$ *and* $A^f_{\triangle} = \{a \in A : f \in V^{M(w)}_{\triangle}\}$*. Then:*

$$A^f_{\triangle C} \subseteq A^f_{\triangle}$$

Proof. Suppose some agent a is in $A^f_{\triangle C}$. Then $f \in V^{M_{\triangle C}(w)}(a)$, so $f \in V_{\triangle C}(w, a)$ by Definition 18, so by Definition 13, one of three things must be the case in the original model M:

(i) For all $v \sim_a w$, it holds that $f \in V(v, a)$ and $N(v, a) = \emptyset$. Then $f \in V(w, a)$ and $N(w, a) = \emptyset$, so $f \in V^{M(w)}(a)$ and $N^{M(w)}(a) = \emptyset$ in the state-generated model $M(w)$, so $f \in V^{M(w)}_{\triangle}(a)$. So then $a \in A^f_{\triangle}$.

(ii) For all $v \sim_a w$, it holds that $\tau = 0$. Since τ does not change with any update or in the process of state-generation, it follows that $f \in V^{M(w)}_{\triangle}(a)$, so $a \in A^f_{\triangle}$.

(iii) For all $v \sim_a w$, it holds that $\frac{|N_f(v,a)|}{|N(v,a)|} \geqslant \tau$. This also holds for w then. So since $N^{M(w)} = N(w, a)$ and $V^{M(w)}(a) = V(w, a)$, we get that $\frac{|N^{M(w)}_f(a)|}{|N^{M(w)}(a)|} \geqslant \tau$. It follows that $f \in V^{M(w)}_{\triangle}(a)$, so $a \in A^f_{\triangle}$. □

Comparing state per state for a given issue $f \in F$, an update step with the cautious epistemic opinion update $\triangle C$ will never lead more agents to adopt the issue than an update with the non-epistemic opinion update $\triangle$. Understanding the number of agents that adopt an issue after one opinion update as *adoption speed*, Proposition 4.3 tells us that cautious opinion updates on an epistemic model never lead to a quicker adoption of issues than "basic" opinion updates on a non-epistemic model. As of now, this leaves open the possibility that adoption under both update types always has the same speed, i.e. that the sets $A^f_{\triangle C}$ and $A^f_{\triangle}$ are identical for all initial models and states. The counterexample given in Figure 15 shows that this is not the case: From an initial epistemic SNM M with one issue, a threshold $\tau = \frac{1}{2}$ and sight $(0, 1)$, updating the model with $\triangle C$ and then obtaining a state-generated model $M_{\triangle C}(w)$ leads to a set of f-adopters $A^f_{\triangle C} = \{a \in A : f \in V^{M_{\triangle C}(w)}(a)\} = \{b\}$. If we first state-generate the model $M(w)$ and then apply the non-epistemic opinion update $\triangle$, we instead get a set $A^f_{\triangle} = \{a \in A : f \in V^{M(w)}_{\triangle}(a)\} = \{a, b, c\}$.

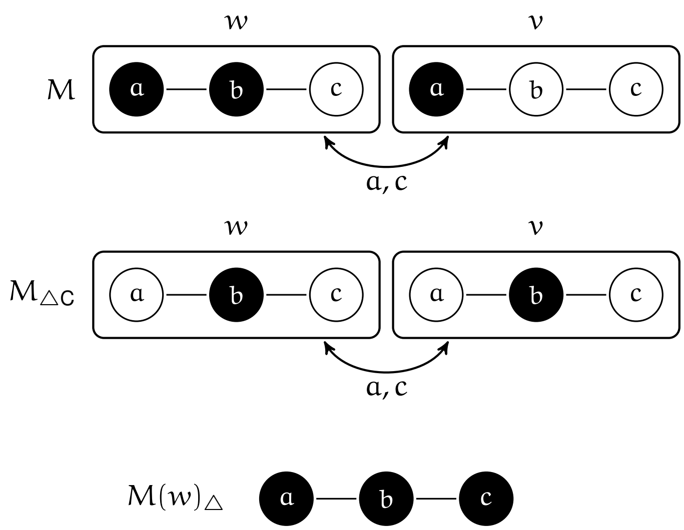

Figure 15: Three models. M is an epistemic SNM with $\tau = \frac{1}{2}$, $F = \{f\}$ and initial sight $(0, 1)$; $M_{\triangle C}$ is the model resulting from updating M with a cautious epistemic opinion update. $M(w)_{\triangle}$ is obtained by taking the model generated from state w of M and applying a non-epistemic opinion update.

Thus, there are models where $\triangle C$ leads to a slower adoption of a given issue than $\triangle$ does. In a sense, Proposition 4.3 and Figure 15 give us the confirmation that $\triangle C$ really does model a more *cautious* attitude towards adopting or supporting issues.

We now turn to eager epistemic opinion updates. It turns out that we can prove a very similar result with respect to them and the non-epistemic opinion update $\triangle$.

Proposition 4.4. *Let M be an epistemic SNM. Fix some state $w \in W$, and let $M(w)$ be the corresponding state-generated non-epistemic SNM. Now consider the model $M_{\triangle E}$, obtained by updating M with the eager epistemic opinion update $\triangle E$; its state-generated model $M_{\triangle E}(w)$; and the model $M(w)_{\triangle}$, obtained by updating the initial state-generated model $M(w)$ with the non-epistemic opinion update $\triangle$.*

Let $A^f_{\triangle E}(w) = \{a \in A : f \in V^{M_{\triangle E}(w)}(a)\}$ be the set of f-supporting agents in the model obtained by first updating with $\triangle E$ and then state-generating. Let $A^f_{\triangle} = \{a \in A : f \in V^{M(w)}_{\triangle}\}$ be the set of f-supporting agents in the model that *was first state-generated and then updated with $\triangle$. Then:*

$$A^{f}_{\triangle} \subseteq A^{f}_{\triangle\mathsf{E}}$$

Proof. Take some $a \in A^{f}_{\triangle}$: We know that $f \in V^{M(w)}_{\triangle}$, so one of three things holds:

(i) $f \in V^{M(w)}$ and $N^{M(w)} = \emptyset$. Then, by Definition 18, we know that in the original epistemic model M, $N(w, a) = \emptyset$ and $f \in V(w, a)$. So then, by Definition 15, we know that $f \in V_{\triangle\mathsf{E}}(w, a)$, so $f \in V^{M_{\triangle\mathsf{E}}(w)}$, so $a \in A^{f}_{\triangle\mathsf{E}}$.

(ii) $\tau = 0$. Since τ does not change with any update or with state-generation, it follows that $f \in V_{M_{\triangle\mathsf{E}}(w)}$, so $a \in A^{f}_{\triangle\mathsf{E}}$.

(iii) $\frac{|N^{M(w)}_{f}|}{|N^{M(w)}(a)|} \geqslant \tau$. Then, by Definition 18, we know that $\frac{|N_f|}{|N(a)|} \geqslant \tau$ in the original epistemic model M. Thus, again by Definition 15, $f \in V_{\triangle\mathsf{E}}(w, a)$, so $f \in V^{M_{\triangle\mathsf{E}}(w)}$, so $a \in A^{f}_{\triangle\mathsf{E}}$. $\square$

Again, the opposite inclusion does not hold, Figure 16 shows a counterexample. Both states of M become an f-consensus after one eager epistemic opinion update $\triangle\mathsf{E}$ with sight $(0, 1)$. Compare this to the state-generated model $M(v)$, which, after one opinion update $\triangle$, would have only agent b supporting f.

We now see that eager epistemic opinion updates never lead to a slower adoption than non-epistemic opinion updates, and that they sometimes cause a genuinely quicker adoption speed. Once again, this can be taken as confirmation that the adoption behavior modeled by $\triangle\mathsf{E}$ really is more *eager*.

The difference between the non-epistemic SNMs of Section 2 and the epistemic SNMs of this section is that the latter allow agents to be *uncertain* about the state of the network. The two types of epistemic opinion updates we defined above differ in how they handle situations in which an agent is uncertain about aspects of the model that are relevant to her own opinion. Propositions 4.3 and 4.4 and Figures 15 and 16 spell out that these differences lead to differences in the speed with which a given opinion spreads through a social network.

The models given as examples of slower or quicker adoption speed in Figure 15 and 16 both have initial sight $(0, 1)$. This is not a coincidence: Only models with initial sight (n, m) with $n < 1$ are models in which agents can be uncertain about the opinions of their neighbors – and their direct neighbors' opinions are the only relevant factor for their decision whether to

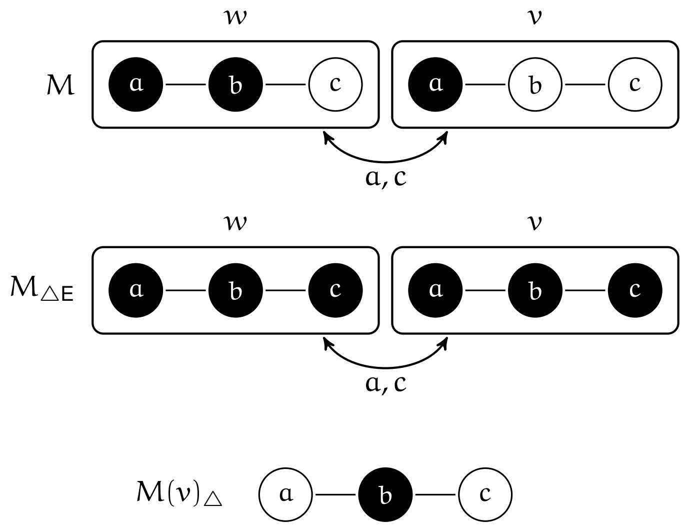

Figure 16: Three models. M is an epistemic SNM with $\tau = \frac{1}{2}$, $F = \{f\}$ and initial sight $(0, 1)$; $M_{\triangle E}$ is the model resulting from updating M with an eager epistemic opinion update. $M(v)_{\triangle}$ is obtained by taking the state-generated model $M(v)$ and applying a non-epistemic opinion update.

support an issue after an opinion update. In other words, with $n \geqslant 1$ agents can still be uncertain about other parts of the model, but they are certain about all factors relevant to opinion updates. This means that in models with sight $(n \geqslant 1, m)$, non-epistemic opinion updates, cautious epistemic opinion updates and eager epistemic opinion updates produce the exact same adoption behavior for opinions:

Proposition 4.5. *Let* $M = (W, A, N, F, V, \omega, \tau, \{\sim_a\}_{a \in A})$ *be an epistemic SNM with initial sight* (n, m)*, where* $n \geqslant 1$ *(and* $m \geqslant 0$*). For any state* $w \in W$*, let* $M(w)$ *be the state-generated model; let* $M(w)_{\triangle}$ *be the model obtained by updating with the non-epistemic opinion update. Let* $M_{\triangle C}$ *be the original model, updated with a cautious epistemic opinion update with sight* (n, m)*; let* $M_{\triangle C}(w)$ *be the state-generated model of that. Let* $M_{\triangle E}$ *be the original model, updated with an eager epistemic opinion update with sight* (n, m)*; let* $M_{\triangle E}(w)$ *be the state-generated model of that. Then:*

$$M(w)_{\triangle} = M_{\triangle E}(w) = M_{\triangle C}(w)$$

Proof. Note that A, F, ω and τ stay the same between all models by definition.

Also, we know by Definition 18 that $N_{\triangle}^{M(w)} = N^{M_{\triangle C}(w)} = N^{M_{\triangle C}(w)}$. So it suffices to prove that the valuations are the same between the three models to prove the claim.

Fix some state $w \in W$. Let $a \in A$ be some agent and $f \in F$ be any issue. Since M has sight $(n \geqslant 1, m)$ we know that:

$$\left(\forall v \sim_a w : \frac{|N^f(v,a)|}{|N(v,a)|} \geqslant \omega\right) \Leftrightarrow \left(\exists v \sim_a w : \frac{|N^f(v,a)|}{|N(v,a)|} \geqslant \omega\right)$$

And by the fact that $\sim_a$ is an equivalence relation, we also know that:

$$\left(\exists v \sim_a w : \frac{|N^f(v,a)|}{|N(v,a)|} \geqslant \omega\right) \Leftrightarrow \left(\frac{|N^f(w,a)|}{|N(w,a)|} \geqslant \omega\right)$$

Using Definition 18, and the update Definitions 3, 13 and 15, it then follows that:

$$f \in V_{\triangle}^{M(w)}(a) \Leftrightarrow f \in V^{M_{\triangle C}(w)}(a) \Leftrightarrow f \in V^{M_{\triangle C}(w)}(a)$$

This holds for any agent and issue, so $V_{\triangle}^{M(w)} = V^{M_{\triangle C}(w)} = V^{M_{\triangle E}(w)}$, and thus we know that all three state-generated models are equal. $\square$

4.3.2 Relation Change

We have gained some insight into how epistemic and non-epistemic opinion updates differ. Since we have defined similar update types on relations, the natural next step is to also compare those. Towards this, we essentially use the same techniques used in the previous section. This allows us to come to a better understanding of the differences in development of the network N between the updates $\Box$, $\Box C$ and $\Box E$.

Proposition 4.6. *Let M be an epistemic SNM, and let $w \in W$ be some state in M. Again, consider the epistemic model obtained by updating with the cautious epistemic relation update, $M_{\Box C}$; the model generated from $M_{\Box C}$ and state w, named $M_{\Box C}(w)$; the state-generated model obtained from w and the original model, named $M(w)$; and the non-epistemically opinion-updated version of that, $M_{\Box}$. Then, for each agent $a \in A$:*

$$N^{M_{\Box C}(w)}(a) \subseteq N_{\Box}^{M(w)}(a)$$

Proof. Suppose that $b \in N^{M_{\Box C}(w)}(a)$. Then, by Definition 18, we know that $b \in N_{\Box C}(w,a)$ in the model $M_{\Box C}$. This means that in the original model M, it must be the case that for all $\nu \in W$ such that $\nu \sim_a w$, $\frac{|\text{Smlty}(\nu)(a,b)|}{|\mathsf{F}|} \geqslant \omega$. So then also $\frac{|\text{Smlty}(w)(a,b)|}{|\mathsf{F}|} \geqslant \omega$, which means that in the state-generated model $M(w)$, we have:

$$\frac{|(V^{M(w)}(a) \cap V^{M(w)}(b)) \cup (\overline{V^{M(w)}(a)} \cap \overline{V^{M(w)}(b)})|}{|\mathsf{F}|} \geqslant \omega$$

Thus, $b \in N_{\Box}^{M(w)}(a)$. □

The opposite direction does not hold for all agents, states and epistemic models, as witnessed by the counterexample given in Figure 17. Here, we see that for agent a and state ν, $N^{M_{\Box C}(\nu)}(a) = \emptyset$ and $N_{\Box}^{M(\nu)}(a) = \{b\}$, so $N^{M_{\Box C}(\nu)}(a) \subsetneq N_{\Box}^{M(\nu)}(a)$.

What does this mean? Given a social network and some agent a, applying the cautious epistemic relation update to the network will make a have at most as many neighbors as she would have as a result of applying a non-epistemic relation update to the network – and for some networks and agents, the cautious epistemic version of the update will lead to a smaller set of neighbors than the non-epistemic one. Like in the previous section, this result can be seen as a kind of sanity check: *Cautious* epistemic updates on relations really do model a more cautious approach to engaging in social ties than the basic, non-epistemic updates. Under $\Box\mathsf{C}$, agents take uncertainty about the similarity of their opinions as a reason to avoid entering a social relation.

To complete the picture, we prove an analogous result for $\Box$ and $\Box\mathsf{E}$:

Proposition 4.7. *Let M be an epistemic SNM, and let $w \in W$ be some state in M. Again, consider the epistemic model obtained by updating with the eager epistemic relation update, $M_{\Box E}$; the model generated from $M_{\Box E}$ and state w, named $M_{\Box E}(w)$; the state-generated model obtained from w and the original model, named $M(w)$; and the non-epistemically opinion-updated version of that, $M_{\Box}$. Then, for each agent $a \in A$:*

$$N_{\Box}^{M(w)}(a) \subseteq N^{M_{\Box E}(w)}(a)$$

Proof. Suppose that $b \in N_{\Box}^{M(w)}(a)$. Then

$$\frac{|(V^{M(w)}(a) \cap V^{M(w)}(b)) \cup (\overline{V^{M(w)}(a)} \cap \overline{V^{M(w)}(b)})|}{|\mathsf{F}|} \geqslant \omega$$

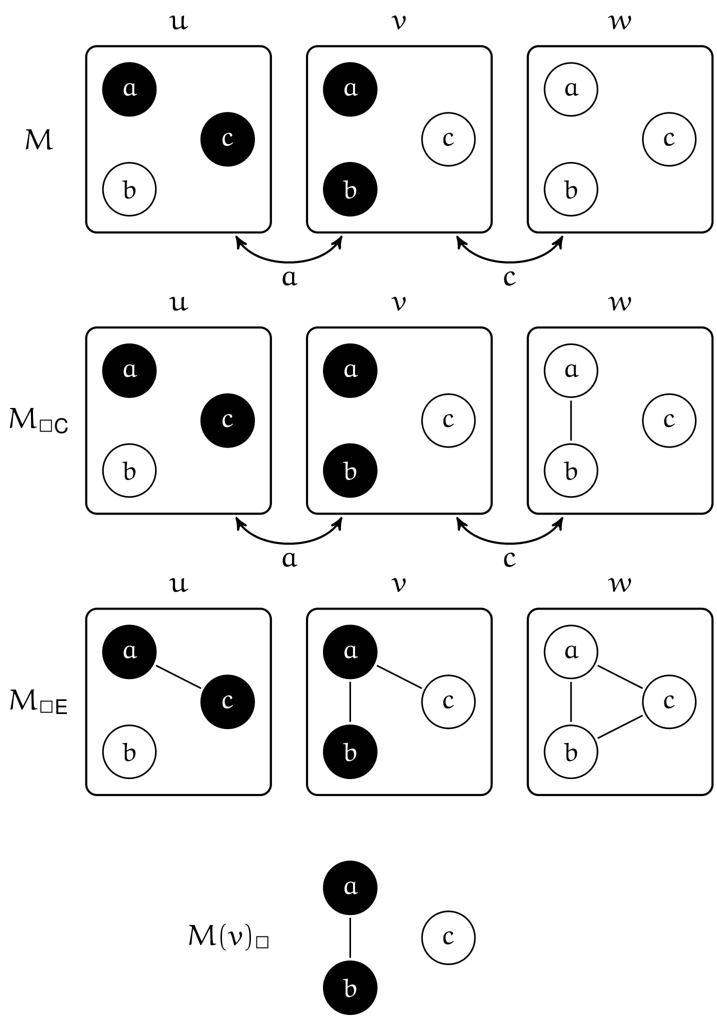

Figure 17: Four models, all with $\omega = 1$. M has initial sight $(1, 0)$. $M_{\square C}$ and $M_{\square E}$ are obtained by applying a cautious or eager epistemic relation update to M. $M(v)_{\square}$ is obtained by taking the model $M(v)$, state-generated from v, and then applying the non-epistemic relation update $\square$.

So by Definition 18, $\frac{|\text{Smlty}(w)(a,b)|}{|F|} \geqslant \omega$ in M. Then, by Definition 16, we know that $b \in N_{\square E}(w, a)$. Again by Definition 18, we then get that $b \in N^{M_{\square E}(w)}(a)$. $\square$

The opposite inclusion does not hold. To see this, consider again Figure 17: For agent a, we have $N_{\Box}^{M(v)}(a) = \{b\}$ and $N^{M(v)_{\Box E}}(a) = \{b, c\}$, so $N_{\Box}^{M(v)}(a) \subsetneq N^{M(v)_{\Box E}}(a)$.

With eager epistemic relation updates, agents will connect as neighbors as long as they are not certain that they do not have enough in common. In effect, this means that agents take uncertainty about their overlap in opinions as a cause to enter social relations. In this way, eager epistemic relation updates on a social network lead, for a given agent, to at least as many neighbors as that same agent would have after a non-epistemic update; and for some models and agents, their set of neighbors after an eager update is larger than after a non-epistemic update.

One question remains to be answered. In the previous section, we had seen that our various types of opinion updates only differ, regarding the models they produce, for models with sight $(n = 0, m)$. Is the existence of differences between the models produced by $\Box$, $\Box C$ and $\Box E$ also dependent on the sight of the model?

It is not. For any sight an epistemic SNM can have, it is possible to find example networks, agents and states for which the three types of opinion updates differ, and example networks for which all three update types give the exact same outcome. This independence, formally stated in Proposition 4.8, reflects the fact that *sight* is a condition on agents' knowledge of their (extended) neighborhood. But for relation updates, the structure of an agents' (extended) neighborhood is irrelevant – the distribution of opinions among agents is the only factor determining the outcome of the update.

Proposition 4.8. *Consider a model* M *with initial sight* (n, m) *(where* $n, m \in \mathbb{N}$ *and* $n \leqslant m + 1$*).*

For each initial sight (n, m) *(where* $n, m \in \mathbb{N}$ *and* $n \leqslant m + 1$*), there exist epistemic social network models* M *and* M' *with that sight and states* $w \in W \in M$*,* $v \in W' \in M'$ *such that:*

$$N^{M(w)_{\Box C}} \subsetneq N_{\Box}^{M(w)} \subsetneq N^{M(w)_{\Box E}} \tag{1}$$

$$N^{M'(v)_{\Box C}} = N_{\Box}^{M'(v)} = N^{M'(v)_{\Box E}} \tag{2}$$

Where the neighborhood sets are defined, based on state-generated and updated models from M *and* M'*, like in Proposition 4.6 and 4.7.*

Proof. Again, consider Figure 17, specifically model M. The caption specifies that M has initial sight $(1, 0)$. But no matter the initial sight of M, the models $M_{\Box\mathsf{C}}$, $M(\nu)_{\Box}$ and $M_{\Box\mathsf{E}}$ always stay the same. Thus, M acts as an example of claim (1) for any initial sight.

As an example for claim (2), consider the following epistemic SNM: $M' = (W, A, N, F, V, \omega, \tau, \{\sim_a\}_{a \in A})$ with $W = \{w\}$, $A = \{a, b\}$, $N(w, a) = N(w, b) = \emptyset$, $F = \{f\}$, $V(w, a) = \{f\}$ and $V(w, b) = \emptyset$, $\omega = 1$ and $\tau = 1$ and $\sim_a = \sim_b = \{(w, w)\}$. In words, this is a model with one state and two neighborless agents that do not have any overlap in opinions. Since $\omega = 1$, neither of the agents will gain a new neighbor after $\Box$, $\Box\mathsf{C}$ or $\Box\mathsf{E}$. This is true regardless of the initial sight we assign to the model. Thus, M' can be used as an example of claim (2) for any initial sight. $\Box$

Thus, a social network model's initial sight does not have an effect on the model's network structure after updating with $\Box$, $\Box\mathsf{C}$ or $\Box\mathsf{E}$. Initial sight does, however, influence the equivalence relations between states after an update – and these equivalence relations are closely connected to mechanisms of *learning* which can take place through updates. In the following sections, we will take a brief look at the topic of learning, and we will see what influence relation updates and initial sight have on it.

4.4 Learning with Epistemic Updates

By including an epistemic dimension in our social network models, we can formally represent situations in which – through an update on the model – agents come to know something they did not know before. This subsection will survey different learning processes that can occur within our formal framework.

Before we start, our formal setting should be distinguished from settings found in two related fields: First, we are only modelling "hard" knowledge using indistinguishability relations for all agents, so our models are not fit to represent more nuanced belief changes known from plausibility models.[25] Further, we are not considering inductive learning in the long term, as is done in formal learning theory.[26] Instead, our notion of learning is more simplistic. Whenever an update of an epistemic SNM leads to an agent being certain of a feature of the network of which she was not certain before,

[25] For an introduction to plausibility models, see [6].

[26] [26] gives a philosophical introduction to the topic. A more logic-focused treatment of formal learning theory can be found in [5].

we say that she has learned something. Keeping this in mind, one can distinguish two basic mechanisms for learning.

4.4.1 Learning by Convergence

Consider an initial epistemic SNM in which an agent cannot distinguish between two states in W that differ with respect to N or V. If an appropriate stream of updates is applied to the model, it might happen that the states' differences regarding the network structure N or valuation V subside. In this case, the aforementioned agent gains knowledge about the network. Since this mechanism enables learning through states developing "to be the same", we will call it *learning by convergence*.

Learning by convergence can occur with any of the epistemic updates we define in Section 4. Figures 11 and 13 show examples of this mechanism: In both figures, we start with a model M in which agent a does not know whether agent b supports issue f or g. After applying $\triangle\mathsf{C}$ or $\triangle\mathsf{E}$, respectively, agent a comes to know the state of the model, formally reflected in the fact that the worlds w and v do not differ w.r.t. V or N anymore. Both examples also illustrate another characteristic of learning by convergence: Even though agent a knows the state of the network after the update, she does not know what state she was initially in, i.e. she does not learn which issue agent b supported in the initial model.

4.4.2 Learning by Distinction

Along with the network's current state, the second learning mechanism allows agents to also learn the initial state of the network after an update stream.

As before, consider an epistemic SNM M, where an agent cannot distinguish between two states $w, v \in W$ that differ w.r.t. N or V. If some update $\dagger$ with sight (n, m) leads w and v to differ in (a) the assignment of opinions to an agent within the n-reachable radius around a or (b) the assignment of neighbors to an agent within the m-reachable radius of a, then a will be able to distinguish w from v in $\mathsf{M}_\dagger$. This is ensured by our inclusion of sight-preserving clauses in the update Definitions 13, 14, 15 and 16 – through these clauses, a's relation $w \sim_a v$ is "cut" if (a) or (b) apply after the update. This can be interpreted as reasoning from agent a: While in M, she might not know if w or v represents the actual state of things, but she can reason that *if* w *is the actual state, it would develop in a way that is different from* v *under the update* †. When this development then takes place, a learns what was initially the case. Since, with this interpretation, this learning mechanism depends on agents' ability to distinguish between states, we will name is *learning by distinction*.

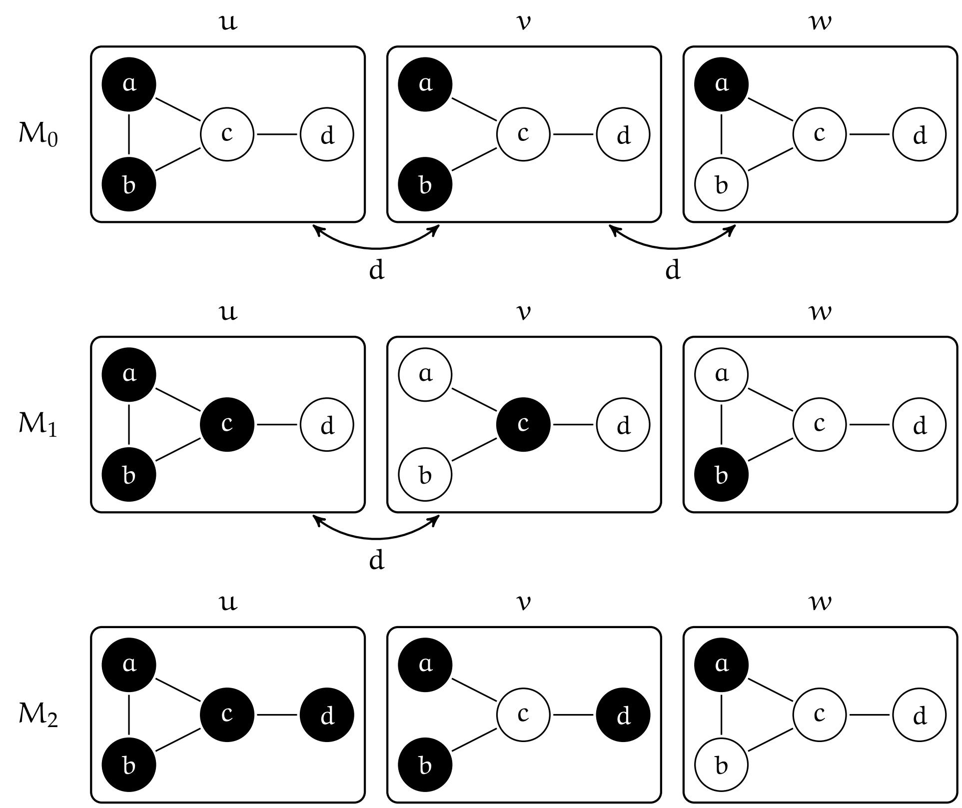

Figure 18: A sequence of SNMs: $M_0, M_1, M_2 \in \triangle\vec{\mathsf{C}}(M)$ with $F = \{f\}$, $\tau = \frac{1}{2}$ and initial sight $(1, 1)$.

Figure 18 shows an example of learning by distinction. In M, agent d does not know if agents a and b are neighbors or if agent b supports f. After one update with $\triangle\mathsf{C}$, d learns to distinguish state w, where agent c does not support f, from states u and v, where c does. Applying $\triangle\mathsf{C}$ once more enables d to also distinguish between u and v: In the former, c still supports f, in the latter, c has stopped her support. Thus, after two applications of

$\triangle \mathsf{C}$, agent d knows which of the three states u, v and w she was in initially.

While Figure 18 shows a sequence of models under the repeated update $\triangle\vec{\mathsf{C}}$, the same sequence could be obtained under the repeated update $\triangle\vec{\mathsf{E}}$. In general, learning by distinction via opinion updates is only possible if the initial model has a sight of $(n \geqslant 1, m)$.[27] Since $\triangle \mathsf{C}$ and $\triangle \mathsf{E}$ only change the opinion distribution in each state, and agents know their own opinion in models with any sight, the only way to distinguish different states is by the opinions in their (extended) neighborhood. But with sight $(0, m)$, agents are not required to be certain about their neighbors opinions, so they cannot make any new distinctions between states after $\triangle \mathsf{C}$ or $\triangle \mathsf{E}$ are applied to the model.

Epistemic relation updates allow for learning by distinction, too – but the specifics of the mechanism differ from learning under opinion updates. Figure 19 shows an example of agent a learning b's opinion after an update with $\square \mathsf{C}$. Figure 20 shows a learning b's opinion through $\square \mathsf{E}$.

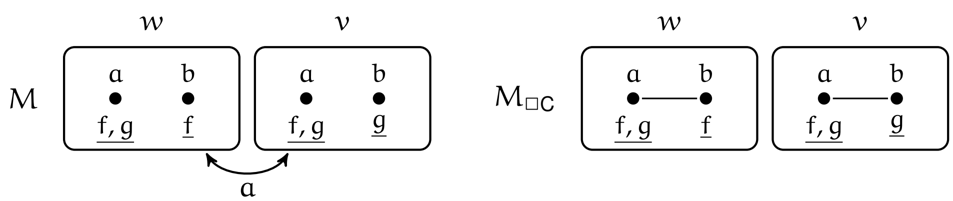

Figure 19: A pair of SNMs with $\mathsf{F} = \{f, g\}$, $\omega = \frac{1}{2}$ and initial sight $(1, 1)$.

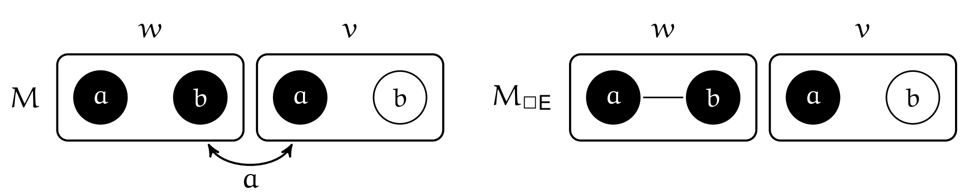

Figure 20: A pair of single-issue SNMs with $\omega = 1$ and initial sight $(1, 1)$.

[27] Together with Proposition 4.5 we thus know that learning by distinction under $\triangle \mathsf{C}$ happens in an epistemic SNM *if and only if* it happens under $\triangle \mathsf{E}$.

Two things should be noted here. First, learning by distinction under either type of relation update is possible even if the update is done with sight $(0,0)$ – Figure 20 shows an example with $\square\mathsf{E}$ that would develop the same if the update were applied with sight $(0,0)$. This is as expected, since relation updates change the network structure, and even with sight $(0,0)$, agents are still required to know who their neighbors are – meaning they can use differences regarding their neighborhoods between states to distinguish those states.

Second, Figure 19 shows how epistemic relation updates can turn *de dicto* knowledge of opinion similarities possessed by an agent into *de re* knowledge: In M, agent a knows that she and agent b have the same opinion on $\omega = \frac{1}{2}$ of the issues in F, but a does not know which of the issues she and b actually agree on. After an update with $\square\mathsf{C}$ (updating with $\square\mathsf{E}$ would give the same result), a learns, by distinction, which issue she and b jointly support. As argued on page 4.1, this is possible because both $\triangle\mathsf{C}$ and $\triangle\mathsf{E}$ only require agents to have *de dicto* knowledge of their sufficient similarity to connect after an update.

Conclusion and Further Work

In this thesis we formally modelled the two tendencies contributing to homophily, social influence and social selection. We used our model to clarify the connection between homophily, fragmentation and polarization, and to investigate the effect of uncertainty on homophily-driven development of social networks.

In Section 1, we introduced homophily as a plausible organizational principle for social networks, and we gave a brief introduction to threshold models. In Section 2, we introduced a formal model of social networks and defined two model-transforming updates: Opinion updates, which correspond to social influence, and social relation updates, which correspond to social selection. We showed that social relation updates are closely connected to cluster-splits and argued that this suggests a connection between homophily and social polarization. This was followed by a number of stabilization results – among them the result that cluster-split models with non-zero thresholds τ and ω will stabilize in a cluster-split under any update stream. We take this to show that the group fragmentation induced by social selection cannot easily be resolved through homophily-based network changes. Section 3 served to introduce the basic ideas and methods of epistemic logic. We put these to work in Section 4, defining epistemic social networks models which allow agents to be unsure about the opinions and social relations of other agents. We defined a number of epistemic model transforming updates and compared them. Our results suggest that the effect of uncertainty on social network development depends on the agents' attitude: Eager behavior can lead to quicker opinion adoption and more strongly socially connected networks; but cautious behavior can make social relations more sparse and slow down the spread of opinions. We finished the section by describing two types of learning phenomena that can take place through changes in the social network.

What could be done next? To conclude this thesis we sketch a number of directions for further research based on the work presented here.

Fragmentation and polarization in multi-issue models. Our results on the connections between homophily-based updates, cluster-splits and polarization were obtained for single-issue models. For multi-issue models, the picture is likely to be more complex, as opinion differences two agents have with respect to one issue can be "bridged" by agreement on other issues. A further investigation of the interaction of opinion updates and social relation updates on multi-issue models could provide a more nuanced account of the connection between homophily, fragmentation and polarization.

Characterizing oscillation. Subsection 2.4 saw us give sufficient conditions for single-issue social network models to oscillate. A proper characterization by way of sufficient and necessary conditions remains to be given, as does a characterization result for oscillation in multi-issue models.

Assigning cautious and eager types to agents. In our epistemic setting, either all agents update cautiously or all agents update eagerly. One could instead assign a C- or E-type to each agent, marking them as a cautious or eager character. It would then be possible to define opinion and relation updates that make each agent update in line with her own type. This would lead to network behavior different from the one we have observed in this thesis.

Mixing cautious and eager relation updates. With the epistemic relation updates defined in Section 4, agents do not differentiate between neighbors and non-neighbors: They connect to another agent if they are certain that there is enough "opinion overlap" (or not certain of the opposite), and they disconnect if they are not certain (or certain that there isn't enough overlap). From a social perspective, it might be more plausible to have stricter epistemic conditions for two agents to *become* neighbors than to *stay* neighbors. With a relation update like this, the network structure of a "pre-update" model would have a greater effect on the "post-update" model.

Prediction updates. Both in Section 2 and Section 4, agents update their opinions and social relations based solely on the present state of the network. Following the work done in [4], one could define *predictive* versions of opinion and social relation updates, allowing agents to incorporate information on the future development of their social network into their update decisions.

Defining a formal language for our models. Analogously to other work in social network logic, it would be possible to define a formal language and an axiom system to reason about the models defined in this thesis. The formalism could follow the examples provided in [10, 3, 4, 2], adding a modal operator for social selection in a way that is similar to [28]. This would e.g. allow for characterization of network properties in the formal language.

Bibliography

[1] Robert J. Aumann. "Agreeing to Disagree". In: *Readings in Formal Epistemology*. Springer International Publishing, 2016, pp. 859–862. DOI: 10.1007/978-3-319-20451-2_40. URL: https://doi.org/10.1007/978-3-319-20451-2_40.

[2] Edoardo Baccini, Zoé Christoff, and Rineke Verbrugge. "Opinion Diffusion in Similarity-Driven Networks". Unpublished manuscript. 2022.

[3] Alexandru Baltag, Zoé Christoff, Jens Ulrik Hansen, and Sonja Smets. "Logical Models of Informational Cascades". In: *Logic Across the University: Foundations and Applications*. Ed. by Johan Van Benthem and Fenrong Lui. College Publications, 2013, pp. 405–432.

[4] Alexandru Baltag, Zoé Christoff, Rasmus K. Rendsvig, and Sonja Smets. "Dynamic Epistemic Logics of Diffusion and Prediction in Social Networks". In: *Studia Logica* 107.3 (July 2018), pp. 489–531. DOI: 10.1007/s11225-018-9804-x. URL: https://doi.org/10.1007/s11225-018-9804-x.

[5] Alexandru Baltag, Nina Gierasimczuk, Aybüke Özgün, Ana Lucia Vargas Sandoval, and Sonja Smets. "A dynamic logic for learning theory". In: *Journal of Logical and Algebraic Methods in Programming* 109 (Dec. 2019), p. 100485. DOI: 10.1016/j.jlamp.2019.100485. URL: https://doi.org/10.1016/j.jlamp.2019.100485.

[6] Alexandru Baltag and Bryan Renne. "Dynamic Epistemic Logic". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Winter 2016. Metaphysics Research Lab, Stanford University, 2016.

[7] Alexandru Baltag and Sonja Smets. "A Qualitative Theory of Dynamic Interactive Belief Revision". In: *Logic and the Foundations of Game and Decision Theory (LOFT 7)*. Amsterdam University Press. 2008, pp. 11–58.

[8] Alexandru Baltag and Sonja Smets. "Learning by Questions and Answers: From Belief-Revision Cycles to Doxastic Fixed Points". In: *Logic, Language, Information and Computation*. Springer Berlin Heidelberg, 2009, pp. 124–139. DOI: 10.1007/978-3-642-02261-6_11. URL: https://doi.org/10.1007/978-3-642-02261-6_11.

[9] Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic*. Cambridge University Press, 2001. DOI: 10.1017/cbo9781107050884. URL: https://doi.org/10.1017/cbo9781107050884.

[10] Zoé Christoff. "Dynamic logics of networks: information flow and the spread of opinion". PhD thesis. University of Amsterdam, 2016. URL: https://eprints.illc.uva.nl/id/eprint/2132.

[11] Hans van Ditmarsch, Wiebe van der Hoek, and Barteld Kooi. *Dynamic Epistemic Logic*. Springer Netherlands, 2008. DOI: 10.1007/978-1-4020-5839-4. URL: https://doi.org/10.1007/978-1-4020-5839-4.

[12] David Easley and Jon Kleinberg. *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge university press, 2010.

[13] Ronald Fagin, Joseph Y. Halpern, Yoram Moses, and Moshe Vardi. *Reasoning About Knowledge*. MIT Press, 1995.

[14] Mark Granovetter. "Threshold Models of Collective Behavior". In: *American Journal of Sociology* 83.6 (May 1978), pp. 1420–1443. DOI: 10.1086/226707. URL: https://doi.org/10.1086/226707.

[15] Pelle G. Hansen, Vincent F. Hendricks, and Rasmus K. Rendsvig. "Infostorms". In: *Metaphilosophy* 44.3 (Apr. 2013), pp. 301–326. DOI: 10.1111/meta.12028. URL: https://doi.org/10.1111/meta.12028.

[16] Sven Ove Hansson. "Logic of Belief Revision". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Spring 2022. Metaphysics Research Lab, Stanford University, 2022.

[17] Kaarlo Jaakko Juhani Hintikka. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Ithaca, NY, USA: Cornell University Press, 1962.

[18] Daniel J. Isenberg. "Group polarization: A critical review and meta-analysis." In: *Journal of Personality and Social Psychology* 50.6 (June 1986), pp. 1141–1151. DOI: 10.1037/0022-3514.50.6.1141. URL: https://doi.org/10.1037/0022-3514.50.6.1141.

[19] David Kempe, Jon Kleinberg, and Éva Tardos. "Maximizing the spread of influence through a social network". In: *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, Aug. 2003. DOI: 10.1145/956750.956769. URL: https://doi.org/10.1145/956750.956769.

[20] Fenrong Liu, Jeremy Seligman, and Patrick Girard. "Logical dynamics of belief change in the community". In: *Synthese* 191.11 (Mar. 2014), pp. 2403–2431. DOI: 10.1007/s11229-014-0432-3. URL: https://doi.org/10.1007/s11229-014-0432-3.

[21] Miller McPherson, Lynn Smith-Lovin, and James M. Cook. "Birds of a Feather: Homophily in Social Networks". In: *Annual Review of Sociology* 27.1 (Aug. 2001), pp. 415–444. DOI: 10.1146/annurev.soc.27.1.415. URL: https://doi.org/10.1146/annurev.soc.27.1.415.

[22] Stephen Morris. "Contagion". In: *Review of Economic Studies* 67.1 (Jan. 2000), pp. 57–78. DOI: 10.1111/1467-937x.00121. URL: https://doi.org/10.1111/1467-937x.00121.

[23] Michael Nelson. "Propositional Attitude Reports". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Spring 2023. Metaphysics Research Lab, Stanford University, 2023.

[24] Eric Pacuit and Olivier Roy. "Epistemic Foundations of Game Theory". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2017. Metaphysics Research Lab, Stanford University, 2017.

[25] Rasmus Rendsvig and John Symons. "Epistemic Logic". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta. Summer 2021. Metaphysics Research Lab, Stanford University, 2021.

[26] Oliver Schulte. "Formal Learning Theory". In: *The Stanford Encyclopedia of Philosophy*. Ed. by Edward N. Zalta and Uri Nodelman. Fall 2023. Metaphysics Research Lab, Stanford University, 2023.

[27] Sonja Smets and Fernando R. Velázquez-Quesada. "A Closeness- and Priority-Based Logical Study of Social Network Creation". In: *Journal of Logic, Language and Information* 29.1 (Jan. 2020), pp. 21–51. DOI: 10.1007/s10849-019-09311-5. URL: https://doi.org/10.1007/s10849-019-09311-5.

[28] Sonja Smets and Fernando R. Velázquez-Quesada. "A Logical Analysis of the Interplay Between Social Influence and Friendship Selection". In: *Dynamic Logic. New Trends and Applications: Second International Workshop, DaLi 2019, Porto, Portugal, October 7–11, 2019, Proceedings 2*. Springer. 2020, pp. 71–87.

[29] Sonja Smets and Fernando R. Velázquez-Quesada. "A logical study of group-size based social network creation". In: *Journal of Logical and Algebraic Methods in Programming* 106 (2019), pp. 117–140.

[30] Sonja Smets and Fernando R. Velázquez-Quesada. "How to make friends: a logical approach to social group creation". In: *Logic, Rationality, and Interaction: 6th International Workshop, LORI 2017, Sapporo, Japan, September 11-14, 2017, Proceedings 6*. Springer. 2017, pp. 377–390.

[31] Cass R. Sunstein. "The Law of Group Polarization". In: *SSRN Electronic Journal* (1999). DOI: 10.2139/ssrn.199668. URL: https://doi.org/10.2139/ssrn.199668.

[32] Liang Zhen and Jeremy Seligman. "A Logical Model of the Dynamics of Peer Pressure". In: *Electronic Notes in Theoretical Computer Science* 278 (Nov. 2011), pp. 275–288. DOI: 10.1016/j.entcs.2011.10.021. URL: https://doi.org/10.1016/j.entcs.2011.10.021.