Iterative social consolidations

Santos, Yuri David; Kooi, Barteld; Verbrugge, Rineke

Published in:
Journal of Logic and Computation

DOI:
10.1093/logcom/exac030

Publication date:
2022

Iterative social consolidations: Forming beliefs from many-valued evidence and peers' opinions

YURI DAVID SANTOS, *Department of Theoretical Philosophy, University of Groningen, 9712 GL, The Netherlands.*

BARTELD KOOI, *Department of Theoretical Philosophy, University of Groningen, 9712 GL, The Netherlands.*

RINEKE VERBRUGGE, *Department of Artificial Intelligence, University of Groningen, 9747 AG, The Netherlands.*

Abstract

Recently, several logics modelling evidence have been proposed in the literature. These logics often also feature beliefs. We call the process or function that maps evidence to beliefs *consolidation*. In this paper, we use a four-valued modal logic of evidence as a basis. In the models for this logic, agents are represented by nodes, peer connections by edges and the *private* evidence that each agent has by a four-valued valuation. From this basis, we propose methods of consolidating the beliefs of the agents, taking into account both their private evidence as well as their peers' opinions. To this end, beliefs are computed iteratively. The final consolidated beliefs are the ones in the point of stabilization of the model. However, it turns out that some consolidation policies will not stabilize for certain models. Finding the conditions for stabilization is one of the main problems studied here, along with other properties of such consolidations. Our main contributions are twofold: we offer a new dynamic perspective on the process of forming evidence-based beliefs, in the context of evidence logics, and we set up and address some mathematically challenging problems, which are related to graph theory and practical subject areas such as belief/opinion diffusion and contagion in multi-agent networks.

1 Introduction

With the advent of social media platforms, the way information spreads and is consumed has shifted drastically in the past decades. As many as 42% of news consumers in countries like Chile, Brazil and Malaysia prefer to get informed via social media, whereas in countries such as the US, Canada and Australia this figure is about 25%.[1] While this new way of sharing information has advantages, it also facilitates the spread of misinformation. When asked whether they are worried about what is 'real or fake' on the internet, a majority of people in Brazil (85%), the UK (70%) and the US (67%) answered positively, whereas in other countries the figure is lower but still significant, e.g. 38% for Germany and 31% for the Netherlands.[2]

Amidst such a deluge of information of variable quality, perhaps one of the most important skills in the 21st century is to use the proper epistemic machinery, i.e. to know how to fetch, filter and

[1] Digital News Report 2019, Reuters Institute. Accessed via: https://ora.ox.ac.uk/objects/uuid:18c8f2eb-f616-481a-9dff-2a479b2801d0/.

[2] Ibid.

Vol. 32, No. 6, © The Author(s) 2022. Published by Oxford University Press.

Advance Access Publication on 8 August 2022 https://doi.org/10.1093/logcom/exac030

aggregate information, to separate reliable from unreliable sources, to combine pieces of evidence appropriately and to consolidate evidence into beliefs.

In artificial intelligence, epistemic and doxastic logics are used as tools to model the knowledge and belief of agents [26, 45]. In practical, real-world scenarios, however, these intelligent agents often have to rely on inconsistent or incomplete data to build up their representation of the world. We can think of these data as *evidence*, a looser and more general concept than that of *justification* as featured in justification logics [2–4, 20, 27]. Here, evidence may be quite weak, e.g. a source that says that a proposition is the case; think of a newspaper article that claims that drinking two cups of coffee a day is healthy.

Recently, a series of logics have emerged with the purpose of modelling agents who possess evidence [13, 21, 29, 34, 37, 38, 41–44]. Given this setting, then, we pose the following problem: how to *consolidate* this evidence into beliefs? We highlighted the relevance of this problem in [33], where we used the term *consolidation*[3] to refer to the process of forming beliefs from evidence—formally represented by functions from evidence to doxastic models. The complexity of certain epistemic tasks has been studied (e.g. in [16, 39]), and, in the same vein, looking at consolidations as processes enables us to ask questions about the complexity of such operations.

As in our previous papers [33, 35], we use the four-valued epistemic logic (FVEL) that we developed in [32, 34] as a base. The resulting system is reminiscent of [6, 7]: agents are represented as nodes and peer relations are represented as as edges, while belief is decided iteratively via modal operators B_0, B_1, etc. In a first moment, $B_0\varphi$ is decided for each formula φ, based solely on the agent's own evidence. These are the agent's initial beliefs or B_0-beliefs. Next, B_1-beliefs are decided based on the agent's evidence again plus the B_0-beliefs of its neighbors. Then, B_2-beliefs are decided similarly but also taking into account B_1-beliefs of the neighbors, and so on. The reason for this choice is to make evidence *private* to each agent. So an agent can only access its own evidence plus its neighbors' *opinions* (or beliefs), but their evidence is not public, in contrast to our previous paper [35]; this public evidence allowed for belief consolidation in a single iteration.

The main question that we investigate in this article is the following: what are the conditions for stabilization of the iterative process of consolidating agents' beliefs given their private evidence and their peers' opinions? We define three consolidation policies and study their (in-)stability properties in different circumstances.

The rest of this article is structured as follows. Section 2 provides the logical background, based on a combination of many-valued and modal logic. Section 3 introduces three consolidation policies and definitions of some desired properties. In Section 4, we attempt to answer the main question of this paper, namely, when is stability attained under these policies? Section 5 discusses some related research, while in Section 6 we sketch some possible avenues of future research.

2 Logical language

In this section we explore a variant of FVEL [34] proposed in [35]. The only difference here is our new definitions for belief.

[3] Borrowed from belief revision [23, 24], where it has the meaning of transforming a potentially inconsistent belief base into a consistent one.

2.1 Syntax

Let At be a countable set of atoms. Below, $p \in At$; the classical part of the language is given by $\mathcal{L}_0$; the propositional part is given by $\mathcal{L}_1$; and the complete language is given by $\mathcal{L}$:

$$\begin{aligned} \mathcal{L}_0 \quad & \psi ::= p \mid \sim \psi \mid (\psi \wedge \psi) \\ \mathcal{L}_1 \quad & \chi ::= \psi \mid \sim \chi \mid (\chi \wedge \chi) \mid \neg\chi \\ \mathcal{L} \quad & \varphi ::= \chi \mid \sim \varphi \mid (\varphi \wedge \varphi) \mid \Box\varphi \mid B_i\psi \end{aligned}$$

where $i \in \mathbb{N}$. We abbreviate $\varphi \vee \psi \stackrel{\text{def}}{=} \sim(\sim \varphi \wedge \sim \psi)$ and $\Diamond\varphi \stackrel{\text{def}}{=} \sim\Box \sim \varphi$.

We restrict belief to classical propositional formulas ($\mathcal{L}_0$) because formulas with $\neg$ refer to evidence, and we are not interested here in expressing agents holding beliefs about evidence, only about facts. As we will see below, the semantics are not quite compositional since, for instance, whether $\neg p$ is true or not is not determined by whether p is true or not. This is because in the four-valued semantics, truth and falsity are not defined as complements but may overlap or leave gaps.

Let us give some examples on how to read formulas of $\mathcal{L}_1$. Here, literals such as p are read as *the agent has evidence for p*, whereas $\neg p$ is read as *the agent has evidence against p*. Similarly, formulas such as $p \wedge q$ are read as *the agent has evidence for* $p \wedge q$, and $\neg(p \wedge q)$ is read as *the agent has evidence against* $p \wedge q$. Here, $\sim \chi$, on the other hand, is read as *it is not the case that* φ, and so $\sim p$ is read as *it is not the case that the agent has evidence for p*.

In $\mathcal{L}$, the reading of formulas is inherited from $\mathcal{L}_1$. Additionally, we read $\Box\varphi$ as φ *holds for all peers* and $B_i\varphi$ as *the agent believes* φ *in iteration i*. Here, our reading of belief formulas deviates from how we read formulas in $\mathcal{L}_1$: $B_i p$ is not read as *the agent believes that she has evidence for p* (at iteration i), but simply as *the agent believes p* (at iteration i).

2.2 Semantics

Models are tuples $M = (S, R, V)$, where S is a finite non-empty set of agents, R is a binary irreflexive[4] relation on S representing 'peerhood' and $V : At \times S \to \mathcal{P}(\{0, 1\})$ is a four-valued valuation representing agents' evidence: $\{1\}$ is *true* (t), $\{0\}$ is *false* (f), $\{0, 1\}$ is *both* (b) and $\emptyset$ is *none* (n).[5] A satisfaction relation is defined as follows:

$M, s \models p$	iff $1 \in V(p, s)$
$M, s \models \neg p$	iff $0 \in V(p, s)$
$M, s \models \sim \varphi$	iff $M, s \not\models \varphi$
$M, s \models (\varphi \wedge \psi)$	iff $M, s \models \varphi$ and $M, s \models \psi$
$M, s \models \neg(\varphi \wedge \psi)$	iff $M, s \models \neg\varphi$ or $M, s \models \neg\psi$
$M, s \models \Box\varphi$	iff for all $t \in S$ such that sRt, it holds that $M, t \models \varphi$
$M, s \models \neg \sim \varphi$	iff $M, s \models \varphi$
$M, s \models \neg\neg\varphi$	iff $M, s \models \varphi$

[4] Baltag, Christoff, Rendsvig and Smets [7] work with symmetric, serial and irreflexive relations. Irreflexivity here means that the agents are not peers of themselves.

[5] We stick to the standard [9] in the naming of truth values. In our context, however, n is better understood as *no evidence about* φ, t as *only evidence for* φ (or *positive* evidence), f as *only evidence against* φ (or *negative* evidence), and b as *evidence both for and against* φ.

An extended valuation function can be defined differently for each type of formula. If $\varphi \in \mathcal{L}_1$, then: $1 \in \overline{V}(\varphi, s)$ iff $M, s \models \varphi$; $0 \in \overline{V}(\varphi, s)$ iff $M, s \models \neg\varphi$. Otherwise: $1 \in \overline{V}(\varphi, s)$ iff $M, s \models \varphi$ iff $0 \notin \overline{V}(\varphi, s)$.

We can also define formulas discriminating which of the four truth values formula $\varphi \in \mathcal{L}_1$ has:

$$\varphi^n \stackrel{\text{def}}{=} (\sim \varphi \wedge \sim \neg\varphi);$$

$$\varphi^f \stackrel{\text{def}}{=} \sim\sim (\sim \varphi \wedge \neg\varphi);$$

$$\varphi^t \stackrel{\text{def}}{=} \sim\sim (\varphi \wedge \sim \neg\varphi);$$

$$\varphi^b \stackrel{\text{def}}{=} \sim\sim (\varphi \wedge \neg\varphi).$$

If φ has value $x \in \{t, f, b, n\}$, i.e. $\overline{V}(\varphi, s) = x$, then φ^x has value t, otherwise φ^x has value f. If a formula $\varphi \in \mathcal{L}_1$ has valuation t or f, we say that the evidence for φ is *unambiguous*, otherwise we say it is *ambiguous*. If $V(p, s) = \{1\}$, we say that *s is a t-agent* (with respect to p, but this will usually be omitted as we will mostly be thinking of a fixed atom p); if $V(p, s) = \{0\}$, we say that *s is an f-agent*, otherwise we say *s is a b/n-agent*.[6]

Notice that the semantics for belief was left open. Our goal in this paper is to discuss a number of possible definitions for the semantics of belief, taking into account that evidence is private to each agent; therefore, belief can only be defined from each agent's own evidence plus their neighbors' beliefs.

3 Iterative social consolidations: preliminaries

The following definition will be employed throughout the paper:

DEFINITION 1 (Attitude).
Let $\text{Att}_i : \mathcal{L}_0 \times S \to \{1, 0, -1\}$ be a function such that:

- $\text{Att}_i(\varphi, s) = 1$ iff $M, s \models B_i\varphi$;
- $\text{Att}_i(\varphi, s) = -1$ iff $M, s \models B_i \sim \varphi$;
- otherwise $\text{Att}_i(\varphi, s) = 0$.

The function Att_i also depends on a model M, but this will be left implicit. We will write Att'_i if we are referring to a modified model M'. Moreover, as it will become clear later, $M, s \models B_i\varphi$ and $M, s \models B_i \sim \varphi$ are mutually exclusive.

How to define beliefs from evidence, i.e. how to *consolidate*? Before defining any consolidations, we will present the following notion, which is similar to bisimulation:

DEFINITION 2 (n-Equivalence).
We say that $(M, s) \leftrightharpoons_n (M', s')$ iff (M, s) and (M', s') satisfy exactly the same $\mathcal{L}_1$ formulas and exactly the same formulas of the form $\Box B_i\varphi$, $\Diamond B_i\varphi$, $\Box \sim B_i\varphi$, $\Diamond \sim B_i\varphi$, where $i \leq n$.

[6] For this paper we could have used only three values (t, f and b/n), but since this is part of a larger project involving FVEL, we chose to keep the four values.

Now we employ this equivalence to limit the space of possibilities. All consolidations have to conform to the following condition:

DEFINITION 3 (Consolidation Definability Condition (CDC)).
If $(M,s) \leftrightharpoons_n (M',s')$, then, for all $\varphi \in \mathcal{L}_0$, $M,s \models B_{n+1}\varphi$ iff $M',s' \models B_{n+1}\varphi$.

What the Consolidation Definability Condition (CDC) does is to make consolidations behave as functions whose input is the initial evidence and the *belief history* of peers. It will be clear later why we want to consider the history instead of just the last iteration of peers' beliefs. Even in this limited space, there are many possibilities. In this paper, we will limit ourselves to consolidations that obey the *regularity* condition, defined as follows:

DEFINITION 4
We define *regular consolidations* to be those policies that respect, for all $i \in \mathbb{N}$:

$$
\begin{array}{lll}
M,s \models B_0 p & \text{iff} & M,s \models p^t \\
M,s \models B_0 \sim p & \text{iff} & M,s \models p^f \\
M,s \models B_i \sim\!\sim \varphi & \text{iff} & M,s \models B_i\varphi \\
M,s \models B_i(\varphi \wedge \psi) & \text{iff} & M,s \models B_i\varphi \text{ and } M,s \models B_i\psi \\
M,s \models B_i \sim (\varphi \wedge \psi) & \text{iff} & M,s \models B_i \sim \varphi \text{ or } M,s \models B_i \sim \psi
\end{array}
$$

Behind Definition 4 is the idea that only beliefs in literals have to be consolidated, and from those basic beliefs others can be built by simple propositional reasoning. Moreover, the first two clauses say that if the evidence for an atom p is only positive (there is only evidence for p but not against p) or only negative, then the agent will initially believe p or $\sim p$, respectively.

Before the first iteration, the agents have not formed any beliefs, so each agent can only use their own private evidence. In the next iterations, however, every agent may have formed beliefs, and therefore, in order to use all information they have available, the agents can now combine their own evidence with the opinion of their peers to form more robust beliefs.[7] We remark that here the iterations are not intended to model the passage of time, but are only a necessary technical device used to circumvent the lack of peers' opinions in the beginning. If the goal were to realistically model time, it would make more sense to have asynchronous updates, where one agent updates in each iteration, but we will leave this variant for future work. The beliefs under $B_0, B_1, \ldots$ here do not really mean that the agent is convinced about such beliefs at any moment; these are just steps towards the agent's actual beliefs, which we will denote by the operator B (without index); this B represents the beliefs of the agent in her *point of stabilization*. So if the agent does not stabilize its beliefs with respect to a formula φ, we cannot say that it has actually formed any (stable) beliefs on φ.

DEFINITION 5 (Stabilization).
An agent s in a model M is said to be *stable* at iteration $i \in \mathbb{N}$ with respect to $\varphi \in \mathcal{L}_0$ if, for all $j \geq i$: $\mathrm{Att}_i(\varphi,s) = \mathrm{Att}_j(\varphi,s)$. A model $M = (S,R,V)$ is said to be *stable* at iteration $i \in \mathbb{N}$ with

[7] This iterative process might remind one of Google's famous PageRank algorithm [30].

respect to $\varphi \in \mathcal{L}_0$ if for all $s \in S$, s is stable at iteration i with respect to φ. If a model/agent is not stable (with respect to a formula) it is *unstable*. The smallest i such that agent s is stable at iteration i with respect to φ is called the *stabilization point* of agent s with respect to φ. The largest stabilization point among all agents in M with respect to φ is called the *stabilization point* of M with respect to φ. If the stabilization point of a model/agent (with respect to φ) is 0, it is called *static* (with respect to φ).

4 Consolidation policies

In this section we will study three regular consolidation policies. Good policies follow some general principles such as not wasting information, being neither too gullible nor too skeptical, etc. We will highlight the qualities and flaws of each policy as we discuss them.

4.1 Policy I: monotonic belief diffusion

Below we define our first consolidation (a complete definition of belief):

DEFINITION 6 (Policy I).
Policy I is the regular consolidation with $M, s \models B_{n+1}p$ iff:
$M, s \models p^t$ or ($M, s \models p^b \vee p^n$ and $M, s \models \Diamond B_n p$ and $M, s \models \Box B_n p$).
And analogously for $B_{n+1} \sim p$.

Now, similarly to [7], we are faced with the question of what are the conditions under which this specific policy eventually stabilizes. In most cases we will only talk about stability referring to some arbitrary atom p, as the dynamics are similar for all formulas.

LEMMA 1
For regular consolidations, if a model/agent is stable at iteration i with respect to all atoms $p \in At$, then this model/agent is stable at iteration i with respect to all formulas $\varphi \in \mathcal{L}_0$.

PROOF. Follows directly from Definition 4. □

Actually, Policy I is guaranteed to stabilize, as we prove below.

PROPOSITION 1
Under Policy I, for any model M and $\varphi \in \mathcal{L}_0$, the stabilization point of M with respect to φ is at most k, where k is the length of the longest directed path in M without repeated edges.

PROOF. First we prove the proposition for an arbitrary atom p, which implies the general proposition due to Lemma 1. Let k be the length of the longest directed path without repetition, and suppose s is an unstable agent at iteration k. If $k = 0$, it is immediately obvious that this cannot be the case, so let us assume that $k > 0$.

If all peers of s were stable at iteration $k - 1$, s would be stable by iteration k (Definition 6). So there is an agent s_1 such that sRs_1 and s_1 is unstable at iteration $k - 1$. Similar reasoning applies to s_1: she has a neighbor s_2 who is unstable at iteration $k - 2$, and so on, until we reach agent s_k who is unstable at iteration 0. But if agent s_k is unstable, there must be an agent s_{k+1} such that $s_k R s_{k+1}$ (which could make the beliefs of s_k change in the next iterations). But, from s to s_{k+1} there is a path of length $k + 1$, which by our assumption (regarding k) means that there is at least one repeated

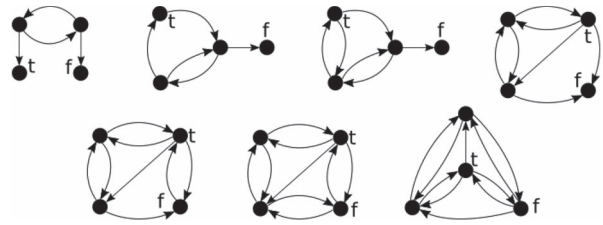

FIGURE 1 All unstable models of size 4 under Policy II have between 4 and 10 edges. This figure shows only some of them.

edge in this path, and therefore one repeated agent. This, in turn, implies that we have a cycle with at most k agents (otherwise the length of the longest path without repetition would exceed k). If s_{k+1} is one of the repeated agents, then $s_{k+1} \in \{s, s_1, ..., s_k\}$; otherwise, the repeated agents are all in $\{s, s_1, ..., s_k\}$. In any case, there is a cycle whose members s_i are all in $\{s, s_1, ..., s_k\}$. But, since all $s_i \in \{s, s_1, ..., s_k\}$ are unstable, they are all b/n-agents. But, if that is the case, then it is not hard to see that $\mathrm{Att}_j(p, s_i) = 0$, for all $j \in \mathbb{N}$. But this means that all $s_i \in \{s, s_1, ..., s_k\}$ are static. Contradiction□

Notice that for consolidations in general, due to the CDC, we cannot talk about fixpoints in the traditional sense, i.e., an iteration i where the beliefs (the output) are the same as in iteration $i-1$. In Policy I, though, it is the case that if all beliefs are the same in iteration i and $i+1$, then the model is stable at i. In this policy, if the evidence is unambiguous, the agent immediately forms belief or disbelief, and never changes. Stabilization is explained by the following:

PROPOSITION 2
In Policy I, the spread of belief is *monotonic*: let $l \in \{p, \sim p\}$ for some $p \in At$; and for all $i \in \mathbb{N}$, let $A_{i,l} = \{s \in S \mid M, s \models B_i l\}$; then for all $i \in \mathbb{N}$, $A_{i,l} \subseteq A_{i+1,l}$.[8]

PROOF. Informally: once an agent adopts belief/disbelief, it means that all her peers have also adopted such attitude (or that she had belief/disbelief from the start, due to unambiguous evidence), which in turn implies that all *their* peers have also done so, and so on... □

Policy I is also very restrictive: the only possible change in attitude for an agent ('in time', or relative to the progression of iterations) is from abstention to belief/disbelief.[9] This leads us to our next definition: a more flexible consolidation.

4.2 Policy II: unstable consolidations

DEFINITION 7 (Policy II).
Policy II is the regular consolidation with $M, s \models B_{n+1}p$ iff:
$M, s \models p^t$ or ($M, s \models p^b \vee p^n$ and $M, s \models \Diamond B_n p$ and $M, s \models \Box \sim B_n \sim p$).
And analogously for $B_{n+1} \sim p$.

[8] A similar monotonicity holds for the Threshold Model Update in [7,Definition 2.4].

[9] Similar in spirit to our work, but different in many technical aspects, Liu, Seligman and Girard [25] call a change from belief to disbelief (or vice-versa) *revision* and a change from belief/disbelief to abstention *contraction*, adopting the classical terms from belief revision [1]. Likewise, the change from abstention to belief/disbelief could be named *expansion*.

What changes now is that peers that abstain are ignored (unless all of them abstain), i.e. the agents are less cautious about forming belief/disbelief when their evidence is ambiguous. Policy II is not guaranteed to stabilize. For example, the models of Figure 1 do not stabilize—in that figure, agents where p has value t are marked with a t, and similarly for f; for the other agents, the evidence for p is ambiguous. For the first model of Figure 1 (top left), note that the agents with unambiguous evidence adopt belief and disbelief immediately, but a (agent on the top left of the model) and b (top right) keep changing between belief (disbelief) and abstention. First, B_0p holds for the agent marked with a t. Then, since b abstains (neither B_0p nor $B_0 \sim p$ hold), the only non-abstaining neighbor of a believes p, therefore we get B_1p for a, and similarly $B_1 \sim p$ for b. In the next iteration, however, a has a neighbor with B_1p (the one marked with t) and one with $B_1 \sim p$ (agent b), so she abstains—similarly for b. The cycle repeats indefinitely; therefore, Policy II is not monotonic, in contrast to Policy I, as shown in Proposition 2.

Instability is undesirable for consolidations. Even though it might be rational to be always open to changing our minds, especially upon the discovery of new evidence, our models are finite and they receive no new information input during the consolidation process. Therefore, rational agents are expected to decide, in a finite amount of time (or number of iterations) what are their final belief states. Now, however, all possible attitude changes between belief, disbelief and abstention are possible.

PROPOSITION 3
Stability in Policy II is decidable.

PROOF. For n agents, a model has 3^n possible belief states (sets of attitudes of the agents) at each iteration. Since belief in one iteration depends only on the fixed evidence and on the belief state in the previous iteration, if the belief state in iteration 3^n differs from that of iteration $3^n - 1$, the model is unstable. □

PROPOSITION 4
Let R^+ be the transitive closure of R. Under Policy II, for every $s \in S$ that remains unstable, there is a t such that sR^+t and t is in a cycle.[10]

PROOF. Consider an agent s such that: (*) there is no t such that sR^+t and t is in a cycle. Then all directed paths starting from this s are finite (forming a rooted directed acyclic graph). We will prove the proposition by induction on the length k of the longest path starting at s. I.H: For each agent s such that (*) holds and whose longest path starting from it has size $k \leq n - 1$, agent s is stable. Base: $k = 0$, then obviously s is stable. Step: $k = n$. Consider any of the longest paths from s: $(s, s_1, ..., s_{n-1})$. Then, $(s_1, ..., s_{n-1})$ has length $n - 1$ and (*) holds for s_1, which by I.H. gives us that s_1 is stable. The other peers of s that belong to smaller paths are also stable, due to the I.H. Therefore, all peers of s are stable and thus so is s. □

PROPOSITION 5
A model with only one b/n-agent s is stable under Policy II.

[10]Here it might be possible to draw some connection to abstract argumentation frameworks [17], because in that theory odd cycles result in the inexistence of stable extensions.

PROOF. If s is the only b/n agent, all peers of s are stable (their attitude does not change). Then there are three possibilities, as follows:

1. The set of peers of s is empty. In that case s is static.
2. All peers have the same attitude. In that case, s will also adopt this attitude.
3. There are at least two peers with different attitudes. In that case, s will remain static.

In all three cases, stabilization occurs at the first iteration. □

PROPOSITION 6
In any model without a t-agent (or without an f-agent), the spread of belief is monotonic (under Policy II).

PROOF. Let M be a model without any f-agents. By Definition 7, under Policy II it is impossible for any agent to have attitude -1 (disbelief). Moreover, the only way that an agent s who 'adopted' (changed attitude from 0 to 1) can unadopt (revert back to 0) is when:

1. all its peers who previously had attitude 1 also unadopted; or
2. one of its peers changed attitude to -1.

Since item (ii) is impossible, the only way is via item (i), but then for that to happen it is necessary that all the peers of the peers of s unadopted. This recursion cannot go on forever because our models are finite, and since we do not have an f-agent, there cannot be a first unadopter. Therefore, unadoption is impossible and thus belief spread is monotonic. Similar reasoning applies for the case of no t-agent. □

DEFINITION 8 (Submodel).
We say $M' = (S', R', V')$ is a *submodel* of $M = (S, R, V)$ if $S' \subseteq S$, $R' \subseteq R$, and for all $p \in At$ and $s \in S'$: $V'(p, s) = V(p, s)$.

DEFINITION 9 (Model Restriction).
Let $M = (S, R, V)$. A restriction M_Z of M to $Z \subseteq S$ is the submodel $M_Z = (Z, R', V')$ of M with $R' = R \cap (Z \times Z)$.

PROPOSITION 7
Let $M = (S, R, V)$, let $s \in S$, let R^* be the reflexive and transitive closure of R, and let $R^*(s) = \{t \in S \mid sR^*t\}$. Then for all $t \in R^*(s)$, all $\varphi \in \mathcal{L}_0$ and all $i \in \mathbb{N}$: $M_{R^*(s)}, t \models B_i\varphi$ iff $M, t \models B_i\varphi$.

PROOF. This is an instance of a theorem that holds for so-called *generated submodels* (see Proposition 2.6 in [12]). □

COROLLARY 1
Any unstable model (under Policy II) has at least one cycle $(s_1, ..., s_n)$ with $n \geq 2$ and such that for all $s_i \in \{s_1, ..., s_n\}$: s_i is a b/n-agent; there are $a, b \in S$ such that s_iR^+a and s_iR^+b, a is a t-agent and b is an f-agent.

PROOF. First we modify the proof of Proposition 4 to show that: for any unstable agent s there is a t such that sR^+t and t is in a cycle consisting only of unstable b/n-agents. We prove the contrapositive

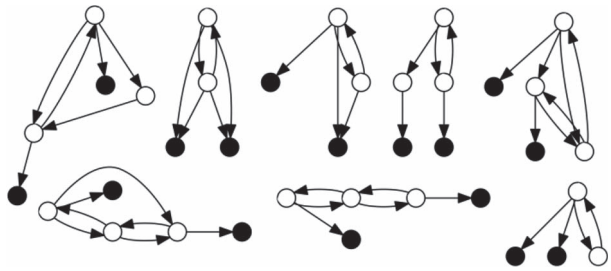

FIGURE 2 Some stable models satisfying the conditions of Corollary 1. In each model here, the b/n-agents are white, any one of the black agents can be taken as a t-agent and the other as an f-agent.

by induction as before, by assuming that no such cycle exists, and therefore any path from s to any unstable b/n-agent is finite. The rest of the induction is similar. In the base case, if the agent has no unstable b/n-peer, then all its peers are stable and therefore it is stable.

Now we just have to show that for all members s_i of this cycle, there are agents a and b as in the corollary statement. Note that for any two agents s_i, s_j in a cycle, $R^+(s_i) = R^+(s_j)$. Now assume there is no t-agent a such that $s_i R^+ a$ is true. We know by Proposition 7 that the beliefs of s_i are the same as in $M_{R^*(s_i)}$, which has no t-agent. But by Proposition 6, we know that in such a model the spread of belief is monotonic and therefore the model stabilizes. So there has to be a t-agent a with $s_i R^+ a$. Analogous reasoning applies for f-agent b. □

COROLLARY 2
The first model of Figure 1 (top left) is the smallest (in number of agents and edges) unstable model under Policy II.

Corollary 1 gives necessary but not sufficient conditions for instability (see Figure 2)[11] :

OPEN PROBLEM 1
What is the set of unstable models under Policy II?

The set of such models is obviously infinite, but enumerable. For each $k \in \mathbb{N}$, we just need to generate all models of size k (which is a finite number of models), with each possible valuation (assuming here only one atom: $At = \{p\}$) and combination of edges, and compute whether the model is stable or not (decidable, by Proposition 3). This algorithm works as a finite description of the set of unstable models (under Policy II). It would be more interesting, however, to have a more 'structural' description, such as the one of Corollary 1.

Unfortunately, we have not managed to find such a simple structural characterization of unstable models (and actually we do not know if such a characterization is even possible), but the following is our attempt at finding 'simplifications' that could hopefully yield models that capture the 'essence' of instability.

[11] For example, Christoff and Grossi [14] solve this problem for a different logic. Van Benthem [40] provides a more abstract study of *oscillations* in logics, corresponding to the phenomenon that we call instability.

DEFINITION 10 (Reduction).
Let $\mathbb{M}$ be the class of all models. A relation $T \subseteq \mathbb{M} \times \mathbb{M}$ is called a *semi-reduction* if for all models $M = (S, R, V)$, $M' = (S', R', V')$, we have MTM' iff: M' is stable iff M is stable; and $S' \subseteq S$. Moreover, if M' is a submodel of M, then T is called a *reduction*.

DEFINITION 11 (Faithful reduction).
A semi-reduction T is called *faithful* if for all models $M = (S, R, V)$, $M' = (S', R', V')$, we have MTM' iff: for all $i \in \mathbb{N}$, all $\varphi \in \mathcal{L}_0$ and all $s \in S'$, $M', s \models B_i\varphi$ iff $M, s \models B_i\varphi$.

Note that if for all M, M', MTM' only if M' is a restriction of M, then T is a faithful reduction. Below, let us consider arbitrary models $M = (S, R, V)$ and $M' = (S', R', V')$.

DEFINITION 12
Below we define $T_1, T_2, T_3, T_4, T_5, T_6 \subseteq \mathbb{M} \times \mathbb{M}$ such that:

- MT_1M' iff: M' is the submodel of M such that $R \setminus R' = \{(s, t)\}$, where $s, t \in S$ and s is a t-agent or an f-agent, and $S' = S$.
- MT_2M' iff: there is an $s \in S$ such that there is no $t \in S$ with sRt or tRs, and M' is the restriction of M to $S \setminus \{s\}$.
- MT_3M' iff: there is a b/n-agent s, a t-agent a and an f-agent b in S such that sRa and sRb, and M' is a restriction of M to $S \setminus \{s\}$.
- MT_4M' iff: there is a b/n-agent $s \in S$ for which there is no t-agent $a \in S$ with sR^+a, and no f-agent $b \in S$ with sR^+b and M' is a restriction of M to $S \setminus \{s\}$.
- MT_5M' iff: there are at least two distinct t-agents (or f-agents) $a, b \in S$; $S' = S$, $V' = V$ and $R' = R \cap ((S \setminus \{b\}) \times (S \setminus \{b\})) \cup Q$, with $Q = \{(a, s) \mid (b, s) \in S \times S\} \cup \{(s, a) \mid (s, b) \in S \times S\}$.
- MT_6M' iff: there is a b/n-agent s for which there is no cycle $(s_1, ..., s_n)$ consisting only of b/n-agents in M such that for an s_i in $(s_1, ..., s_n)$, a t-agent $a \in S$ and an f-agent $b \in S$, it holds that s_iR^+a, s_iR^+b and s_iR^+s; and M' is a restriction of M to $S \setminus \{s\}$.

Note that T_5 is the only of the relations T_i above in which MT_iM' does not require M' to be a submodel of M, which means that one has to apply it wisely if one wants to actually simplify a model (basically, one t-agent and one f-agent have to be chosen to concentrate all incoming arrows). It is called a semi-reduction because it does not necessarily yield simpler models.

PROPOSITION 8
The relations T_1, T_2, T_3, T_4 of Definition 12 are faithful reductions, T_5 is a faithful semi-reduction, and T_6 is a (non-faithful) reduction.

PROOF. This proof is straightforward. In some cases, one just has to use Proposition 7 and have in mind that peers with attitude 0 do not affect any agent's beliefs. □

Now one can apply arbitrary sequences of the reductions above to obtain, from an arbitrary model, less cluttered counterparts that are stable if and only if the original was (see Figure 3). In the three cases in Figure 3, the models have certain features that allow us to enlarge them or add nodes and connections that do not change anything stability-wise. For example, if an agent is connected to a t-peer, adding another t-peer will not make any difference. Also, if in the reduced example to the right of the arrow in Figure 3(b), one added some b/n-agents between the t-agent and the cycle, as

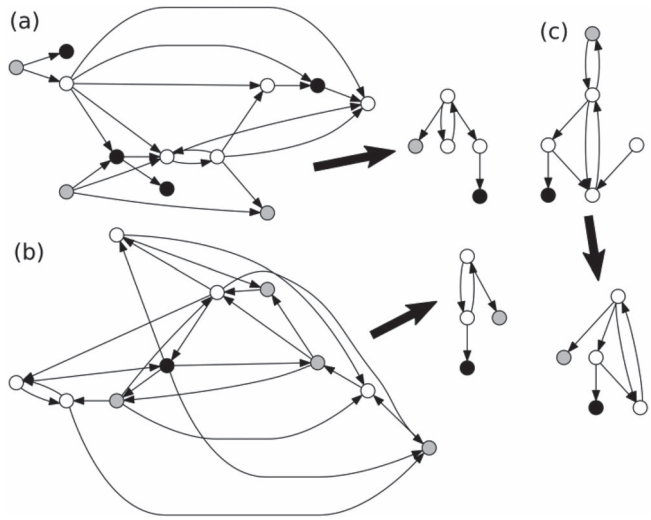

FIGURE 3 Here b/n-agents are white, t-agents are gray and f-agents are black. Many stable models reduce to a single agent model (after applying $T_1 - T_6$ as much as possible), but there are cases like (a) above where it is not reduced completely. Likewise, many unstable models reduce to the smallest unstable model, like case (b), but some do not, as in case (c). These reduced models highlight features affecting a model's stability or instability.

long as one added the same number of b/n-agents between the f-agent and the cycle in the same direction, it would just take one more iteration for the beliefs to propagate.

In Figure 3(a), if the path between the t-agent and the cycle has a different length than the path between the f-agent and the cycle, then this will stabilize the cycle. Finally, in Figure 3(c), only some arrows and agents that do not affect stability have been removed when going from the top picture to the bottom one.

One can also use only faithful reductions to obtain a simplification where all agents have exactly the same belief history. Using reductions might be a more efficient way of checking whether a model is stable, and it certainly is an easier method for humans in many cases. A more formal comparison of the complexity between checking stabilization in the standard way versus using reductions is left for future work.

To get a sense of the prevalence of instability, we randomly generated (using the Erdoős-Rényi method [19]) and tested 100,000 models of size 4 to 60 and found the percentages of unstable ones, as shown in Figure 4. It is clear that we can expect a zero-one law here (e.g. as in [46]), with the percentage going to zero in the limit when the size tends to infinity. This mitigates the problem of instability for Policy II: in large enough models (such as a big network of scientists, for example), 'almost never' will the agents incur an irrational consolidation infinite loop. A possible informal explanation is that, as the size increases, the number of possible structures that can be built grows much faster than the number of possibilities for 'instability-inducing' structures, which are very specific: they have to respect the conditions of Corollary 1 plus other unknown conditions (Open Problem 1). Another interesting question is: why does the percentage of unstable models peak at size 11?

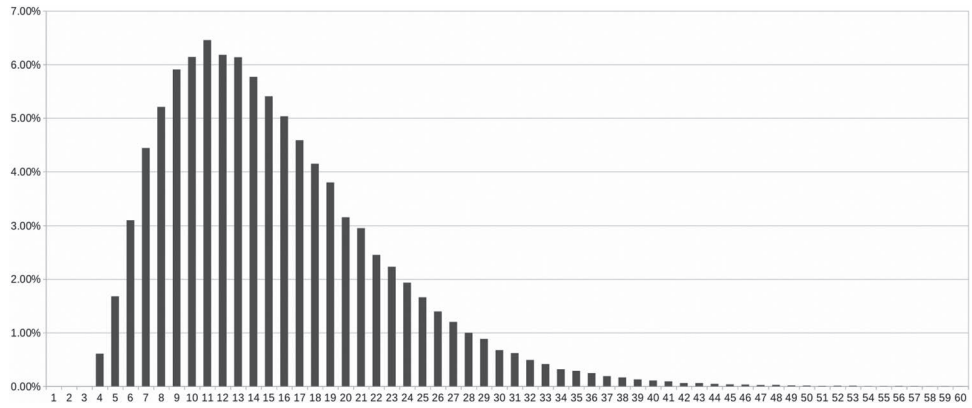

FIGURE 4 Percentage of unstable models per model size, for Policy II.

4.3 Policy III: ignoring unstable peers

Our next consolidation policy will try to tackle the instability problem by temporarily ignoring agents who have not been stable for the last λ iterations. Formally, we define the following abbreviation:

$$M, s \models \mathbf{stable}^n_{\lambda,p} \text{ (with } n \geq \lambda \geq 1 \text{and } p \in At)$$

with the meaning: $Att_{n-1}(p,s) = Att_{n-2}(p,s) = ... = Att_{n-\lambda}(p,s)$. (*Agent s has been stable about p in the last* λ *iterations preceding iteration n*), and set that if $\lambda \leq 1$, then $M, s \models \mathbf{stable}^n_{\lambda,p}$; and if $\lambda > n$, then $M, s \models \mathbf{stable}^n_{\lambda,p}$ is defined as $M, s \models \mathbf{stable}^n_{n,p}$. This abbreviation does not increase expressivity, because it can always be defined by a finite propositional combination of conditions. For example, $M, s \models \mathbf{stable}^{10}_{2,p}$ is defined as the following disjunctive condition: ($M, s \models B_9 p$ and $M, s \models B_8 p$) or ($M, s \models B_9 \sim p$ and $M, s \models B_8 \sim p$) or ($M, s \models \sim B_9 p \wedge \sim B_9 \sim p$ and $M, s \models \sim B_8 p \wedge \sim B_8 \sim p$).

From the above, we conclude that, as expected, $M, s \models \Box\mathbf{stable}^n_{\lambda,p}$ means that for all $t \in S$ such that sRt, $M, t \models \mathbf{stable}^n_{\lambda,p}$. Dually, $M, s \models \Diamond\mathbf{stable}^n_{\lambda,p}$ means that there is a $t \in S$ such that sRt and $M, t \models \mathbf{stable}^n_{\lambda,p}$. To restrict the modal operators to stable peers only, we can define $M, s \models \Box^n_{\lambda,p}\varphi$ as $M, s \models \Box(\sim \mathbf{stable}^n_{\lambda,p} \vee \varphi)$, and we can define $M, s \models \Diamond^n_{\lambda,p}\varphi$ as $M, s \models \Diamond(\mathbf{stable}^n_{\lambda,p} \wedge \varphi)$. Now we are ready to define the family of policies Policy III-λ:

DEFINITION 13 (Policy III-λ).
Let $1 \leq \lambda \in \mathbb{N}$. Policy III-λ is the regular consolidation with $M, s \models B_{n+1}p$ iff:
$M, s \models p^t$ or ($M, s \models p^b \vee p^n$ and $M, s \models \Diamond^{n+1}_{\lambda,p} B_n p$ and $M, s \models \Box^{n+1}_{\lambda,p} \sim B_n \sim p$).
And analogously for $B_{n+1} \sim p$.

It is not hard to see that Definition 13 is compliant with the CDC (Definition 3), which is also the reason why we defined the CDC based on the history of peers' beliefs and not only on the last iteration. Note also that if the parameter $\lambda = 1$, Policy III-λ coincides with Policy II, so the former is a generalization of the latter. Figure 5 shows the evolution of belief using Policy III-λ on the model of Figure 1, with different values of λ.

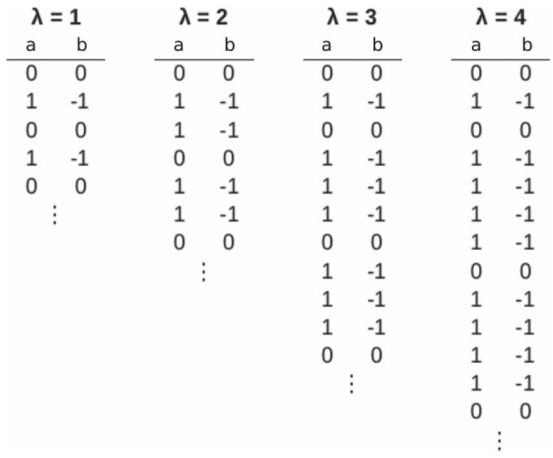

FIGURE 5 Iterations of belief for agents a and b in the first model of Figure 1 (top left). Abstention is represented by 0, belief by 1 and disbelief by -1 (as in Definition 1).

The question that immediately surfaces is whether larger values of λ improve stability. Looking at Figure 5, we notice that larger values of λ make iterations of abstention less frequent. We have to formally define what 'improving stability' means here.

DEFINITION 14 (Stability Measure).
A stability measure $<_{M,\varphi}$ for consolidations (with respect to a fixed model M and an arbitrary $\varphi \in \mathcal{L}_0$), where $C_2 <_{M,\varphi} C_1$ is read as C_1 *is more stable than* C_2 *for* φ *in* M, has to respect the following principles:

1. If C_1 makes M static but C_2 does not, then $C_2 <_{M,\varphi} C_1$;
2. if C_1 makes M stable but C_2 does not, then $C_2 <_{M,\varphi} C_1$;
3. if M stabilizes with C_1 at iteration i and with C_2 at $j > i$, then $C_2 <_{M,\varphi} C_1$;

If for all models M and all $\varphi \in \mathcal{L}_0$ it is the case that $C_2 <_{M,\varphi} C_1$, then we say that $C_2 < C_1$, i.e., C_1 is more stable than C_2.

For an arbitrary measure of stability, our initial hypothesis is as follows:

HYPOTHESIS 1
Let $1 \leq \lambda, \kappa \in \mathbb{N}$. If $\lambda < \kappa$, then Policy III-λ < Policy III-κ.

We can start by asking whether Policy III with $\lambda = 1$ can, in some case, be more stable than with $\lambda = 2$. Definition 14-i cannot be used to violate Hypothesis 1, when changing λ from 1 to 2. If a model is static, changing λ will not produce any changes in belief. Perhaps surprisingly, though, our Hypothesis 1 can be violated by Definition 14-ii, i.e. there is a model that stabilizes when $\lambda = 1$ but does not when $\lambda = 2$, namely the model of Figure 6 (right). For clarity, the value of p is not shown when it is ambiguous.

The models of Figure 6 (left) and Figure 6 (right) are the smallest models (considering number of agents and arrows) that feature, respectively: (a) a change from unstable with $\lambda = 1$ to stable with $\lambda = 2$ and (b) the opposite. We tested computationally and verified that phenomena (a) and (b)

FIGURE 6 Left: unstable with $\lambda = 1$, stable with $\lambda = 2$. Right: stable with $\lambda = 1$, unstable with $\lambda = 2$.

λ = 1

a	b	c
0	0	0
-1	1	0
0	1	-1
0	0	0
⋮		

λ = 2

a	b	c
0	0	0
-1	1	0
-1	1	0
0	1	-1
0	1	0

stabilised

λ = 1

a'	b'	c'
0	0	0
-1	1	0
0	0	-1
-1	0	0
-1	0	-1

stabilised

λ = 2

a'	b'	c'
0	0	0
-1	1	0
-1	1	0
0	0	-1
-1	1	0
-1	1	0
0	0	-1
⋮		

FIGURE 7 Iterations of belief for the models of Figure 6.

do not happen in models with 4 or less agents. Another surprising result is that, among models of size 5, there are exactly the same number of models where (a) and (b) occur. Moreover, we tested with $\lambda = 1, ..., 10$, running for at most 1000 iterations, and all models fitting (a) were stable with $\lambda = 2, ..., 10$, and all models fitting (b) were unstable with $\lambda = 2, 3$ and stable otherwise. We suspect that phenomena (a) and (b) occur due to the qualitative difference between Policy III-λ with $\lambda = 1$, which equals Policy II, and with $\lambda \geq 2$. Moreover, on the basis of the experiments we can conjecture that increasing λ does have a positive effect in terms of stability for the vast majority of models and changes in λ, with possible exceptions when λ is raised from 1 to 2 or 3 (although Hypothesis 1 is false). This is because the (b)-type models became stable with larger values of λ, despite becoming unstable with $\lambda = 2, 3$.

4.4 Comparing the policies

In the beginning of Section 4, we mentioned some properties of good policies such as not wasting information and being neither too gullible nor too skeptical. With respect to these properties, we make a few observations:

- Policy II is less skeptical than Policy I, because in Policy II, in contrast to Policy I, an agent does not need all neighbors to believe p in order to start believing p.
- Similarly, Policy III seems to be less skeptical than Policy I.
- It is unclear whether Policy III is more or less skeptical than Policy II. Policy III seems at first sight to create fewer iterations in which agents are undecided, but we have not proven any relation formally.

From a certain perspective, skepticism has to do with wasting information: if an agent never forms any belief, then it is maximally skeptical but also wastes maximum information. From another perspective, an agent ignores information in order to form more beliefs, e.g. an agent using Policy III ignores unstable peers for that purpose.

4.5 Other policies

Other policies that have not yet been explored include the following ideas:

1. Stopping the consolidation at a fixed iteration defined by a parameter λ. This policy would guarantee stability in a forceful manner. A drawback is that it would be *too sensitive* to the parameter λ, especially in the case of unstable models (under Policy III);
2. Limiting the number of times an agent can change its attitude, e.g. after going from abstention to belief/disbelief, it cannot go back to abstention. We probably will not have problems to define this consolidation respecting the CDC, for even though the agents *do not* take into account their own belief histories directly, these are definable from their peers' previous beliefs;
3. Defining belief based on the number of peers holding a certain attitude. In [35], we show that by introducing a dynamic operator (which increases expressivity), we can count peers with certain attitudes. This would probably be a more realistic way of consolidating beliefs, but demands a language richer than the modal logic used here.
4. Allowing t and f-agents to change attitudes. For this, item (iii) above might be helpful. Alternatively, b/n-agents could have a policy similar to Policy II, whereas t and f-agents could be more resistant to change, adopting a strategy in line with Policy I.

5 Related work

Researchers in social epistemology and agent-based modeling have investigated agents in social networks that combine their own evidence with the beliefs they know their peers to have. Beliefs in social networks sometimes indeed stabilize over time. This phenomenon can be rational and positive, see e.g. the *correct consensus* described by Zollman [47]. Stabilization can also have undesired effects, such as the *mere (incorrect) consensus* described by Zollman [47], the *echo chambers* described by Nguyen [28], and the *polarization* described by Dykstra and colleagues [18].

From a logical point of view, perhaps the closest work to ours is being done by Sedlár and Majer [36]. They use a four-valued paraconsistent logic that is very close to the one we used previously in [32, 34] and on which the current formalism is based. Their work features a source-related interpretation of the modal accessibility relation. In this way, inconsistencies and conflicting information are explained by information with origins in different sources. The motivation of their paper differs from ours in many ways and is more closely related to our previous work (see, e.g. [32, 34]).

Grandi, Lorini and Perrussel [22] study the way opinions on a set of issues evolve in a social network in a setting that is similar to ours and they study the notion of convergence, which is very similar to our notion of stability. Their model has a set of agents who are linked in a graph and who try to form an opinion on a set of atomic propositions, which may be restricted by so-called integrity constraints. The process of how opinions may change in a network involves iterative application of well-known aggregation procedures from the literature on social choice theory. They obtain a number of results on convergence. The main difference between their approach and ours is that the opinions of agents in the system of [22] are binary, so an agent cannot abstain on an issue. The other difference is that [22] has *integrity constraints* that capture the possible dependencies that atomic propositions may have. We do not consider these in our framework.

An important societal problem is *pluralistic ignorance*: each agent in a group believes that their private attitude deviates from that of the others, while in fact everyone acts identically, in opposition

to their private attitudes. For example, in Hans-Christian Andersen's famous fairy-tale, a whole group of adults acts as if they believe that the emperor is showing off his beautiful new clothes, until a child expresses their true beliefs, namely, that the emperor is not wearing any clothes at all. Christoff and Hansen [15] present a dynamic hybrid framework that is sufficiently fine grained to model situations of pluralistic ignorance, which could not yet be modeled in [25]. It would be interesting to extend the framework of our current paper, in which agents' evidence is always private, with the possibility of public evidence as in [35] in order to be able to model such situations of pluralistic ignorance.

Bjorndahl and Özgün [11] study a generalization of the topological frameworks of [5, 29]. In their new framework, evidence is taken to be factual: each piece of evidence corresponds to a set of possible worlds. However, this correspondence is itself subject of uncertainty, i.e. the agents do not know what the evidence actually entails. This is similar to our approach in that the agents have to find out what to believe based on their evidence (i.e., they have to consolidate), but differs from it in the factivity of evidence. What the evidence entails is not necessarily clear in our models. Their logic allows for the modelling of, among other things, calibration errors, giving it a quantitative aspect in the treatment of evidence. This is something we hinted at in the end of [35], although the quantitative aspects of these two papers are manifested in very different ways.

6 Conclusions and future work

In this paper we used a many-valued modal logic (FVEL) to represent a multi-agent network of peers and their evidence, and defined belief based on this network and on the evidence. This paper is an alternative view to our previous paper [35], where the agents can access one another's evidence, therefore allowing for consolidations done in one step. By making the evidence private, we triggered an iterative process, through which the agents always (in Policy I) or almost always (Policy II) approximate a final belief. The exception to this is in the problematic cases of unstable models, which were one of the main topics explored here.

For the consolidation policies with public evidence that we defined in [35], we extensively assessed them against a set of rationality postulates inspired by social choice theory, which we call 'consistency', 'modesty', 'no gurus', 'atom independence', 'logical omniscience' and 'monotonicity'. For precise definitions of these rationality postulates and the relation between belief consolidation and judgment aggregation based on voting, see [35]. A similar assessment of rationality postulates for our current policies based on private evidence remains for future work. At first sight, within the abstractions and limitations of our modelling (i.e., respecting the CDC), the policies sketched here seem to be moderately rational. At least, they respect some of the rationality postulates that we defined for general consolidation functions in [33, footnote 4]:

- If evidence is only positive, then you should not disbelieve;
- If evidence is only negative, then you should not believe;
- If only positive (negative) evidence is not enough to generate belief (disbelief), then nothing is.

One philosophical point that has to be better defended is the non-temporal aspect of the iterations in the consolidation process. We are not trying to represent agents who are updating their beliefs as the days and years go by. This time can be seem as just a 'processing time'. It can also be viewed as passage of time in a very restricted situation where agents cannot do anything besides communicating with their peers—they do not have the opportunity to consult other sources of information or to make deep reflections—as if they are deliberating in an isolated room. This

deliberation, of course, would be of a very simple kind, in which all they are allowed to do is to ask the opinions of their peers, at discrete moments of time, or turns.

In the problem of informational cascades [8, 10], rational individual behaviour might lead to bad doxastic outcomes. In that case, this happens due to the evidence being private to each agent, who can only access the others' final judgements, which is also a feature of our models. The instability, but possibly other irrational behaviours in our system, is partially explained by that feature. A comparison with the system in [35], where others' evidence is public, would help to elucidate the impact of private evidence.

There is still much work to be done on these consolidations, especially Policy III, which has not yet been explored in great depth. It would also be interesting to study the properties of the stabilized belief operators B of the stable policies that we studied so far. Other policies such as the ones in Section 4.5 can also be studied. Moreover, as remarked earlier, here a three-valued logic (such as the ones in [31, Ch. 7]) of evidence would suffice, but consolidations that distinguish n and b could be developed in future work.

Funding

This research was funded by Ammodo KNAW project "Rational Dynamics and Reasoning".

References

[1] C. E. Alchourrón, P. Gärdenfors and D. Makinson. On the logic of theory change: partial meet contraction and revision functions. *Journal of Symbolic Logic*, **50**, 510–530, 1985.
[2] S. Artemov. Logic of proofs. *Annals of Pure and Applied Logic*, **67**, 29–59, 1994.
[3] S. Artemov. Operational modal logic. *Technical Report MSI 95–29*. Cornell University, 1995.
[4] S. Artemov. Explicit provability and constructive semantics. *Bulletin of Symbolic Logic*, **7**, 1–36, 2001.
[5] A. Baltag, N. Bezhanishvili, A. Özgün and S. Smets. Justified belief and the topology of evidence. In *Logic, Language, Information, and Computation*, pp. 83–103. Springer, 2016.
[6] A. Baltag, Z. Christoff, R. Rendsvig and S. Smets. Dynamic epistemic logics of diffusion and prediction in social networks. In *Proceedings of the 12th Conference on Logic and the Foundations of Game and Decision Theory*, 2016.
[7] A. Baltag, Z. Christoff, R. Rendsvig and S. Smets. Dynamic epistemic logics of diffusion and prediction in social networks. *Studia Logica*, **107**, 489–531, 2019.
[8] A. V. Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, **107**, 797–817, 1992.
[9] N. Belnap. A useful four-valued logic. In *Modern Uses of Multiple-valued Logic*, M. Dunn and G. Epstein., eds, pp. 5–37. Springer, 1977.
[10] S. Bikhchandani, D. Hirshleifer and I. Welch. A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy*, **100**, 992–1026, 1992.
[11] A. Bjorndahl and A. Özgün. Uncertainty about evidence. In *Proceedings Seventeenth Conference on Theoretical Aspects of Rationality and Knowledge, TARK 2019, Toulouse, France, 17–19 July 2019*, L. S. Moss., ed. EPTCS, vol. **297**, pp. 68–81, 2019.
[12] P. Blackburn, M. de Rijke and Y. Venema. *Modal Logic*. Cambridge University Press, 2002.
[13] W. Carnielli and A. Rodrigues. An epistemic approach to paraconsistency: a logic of evidence and truth. *Synthese*, **196**, 1–25, 2017.

[14] Z. Christoff and D. Grossi. Stability in binary opinion diffusion. In *Logic, Rationality, and Interaction*, pp. 166–180. Springer, 2017.
[15] Z. Christoff and J. U. Hansen. A logic for diffusion in social networks. *Journal of Applied Logic*, **13**, 48–77, 2015.
[16] C. Dégremont, L. Kurzen and J. Szymanik. Exploring the tractability border in epistemic tasks. *Synthese*, **191**, 371–408, 2014.
[17] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, **77**, 321–357, 1995.
[18] P. Dykstra, C. Elsenbroich, W. Jager, G. R. Renardel de Lavalette and R. Verbrugge. Put your money where your mouth is: DIAL, a dialogical model for opinion dynamics. *Journal of Artificial Societies and Social Simulation*, **16**, 2013.
[19] P. Erdős and A. Rényi. On random graphs I. *Publicationes Mathematicae*, **6**, 290–297, 1959.
[20] M. Fitting. The logic of proofs, semantically. *Annals of Pure and Applied Logic*, **132**, 1–25, 2005.
[21] M. Fitting. Paraconsistent logic, evidence and justification. *Studia Logica*, **105**, 1149–1166, 2017.
[22] U. Grandi, E. Lorini and L. Perrussel. Propositional opinion diffusion. In *Proceedings of the 14th International Conference in Autonomous Agents and Multiagent Systems (AAMAS-2015)*, 2015. Errata: Theorem 9 is false.
[23] S. O. Hansson. *Belief Base Dynamics*. PhD Thesis, Uppsala University, 1991.
[24] S. O. Hansson. Semi-revision. *Journal of Applied Non-Classical Logics*, **7**, 151–175, 1997.
[25] F. Liu, J. Seligman and P. Girard. Logical dynamics of belief change in the community. *Synthese*, **191**, 2403–2431, 2014.
[26] J.-J. Meyer and van der Hoek. *Epistemic Logic for AI and Computer Science*. Cambridge University Press, 1995.
[27] A. Mkrtychev. Models for the logic of proofs. In *Logical Foundations of Computer Science*, pp. 266–275. Springer, 1997.
[28] C.T. Nguyen. Echo chambers and epistemic bubbles. *Episteme*, **17**, 141–161, 2020.
[29] A. Özgün. *Evidence in Epistemic Logic: A Topological Perspective*. PhD Thesis, University of Amsterdam, 2017.
[30] L. Page, S. Brin, R. Motwani and T. Winograd. The pagerank citation ranking: bringing order to the web. *Technical Report 1999–66*, Stanford InfoLab, 1999. Previous number = SIDL-WP-1999-0120.
[31] G. Priest. *An Introduction to Non-Classical Logic: From If to Is*, 2nd edn. Cambridge University Press, 2008.
[32] Y. D. Santos. A dynamic informational-epistemic logic. In A. Madeira and M. Benevides., eds, *Dynamic Logic. New Trends and Applications*. Lecture Notes in Computer Science, vol. **10669**, pp. 64–81. Springer, 2018.
[33] Y. D. Santos. Consolidation of belief in two logics of evidence. In P. Blackburn, E. Lorini and M. Guo., eds, *International Conference on Logic, Rationality and Interaction (LORI)*. Lecture Notes in Computer Science, vol. **11813**, pp. 57–70. Springer, 2019.
[34] Y. D. Santos. A four-valued dynamic epistemic logic. *Journal of Logic, Language and Information*, **29**, 451–489, 2020.
[35] Y. D. Santos. Social consolidations: rational belief in a many-valued logic of evidence and peerhood. In *Foundations of Information and Knowledge Systems: 11th International Symposium, FoIKS 2020, Dortmund, Germany, February 17–21, 2020, Proceedings*, A. Herzig and J. Kontinen., eds, pp. 58–78. Springer, 2020.

[36] I. Sedlár and O. Majer. Modelling sources of inconsistent information in paraconsistent modal logic. In *New Essays on Belnap-Dunn Logic*, H. Omori and H. Wansing., eds, pp. 293–310. Springer, 2019.
[37] C. Shi, S. Smets and F. R. Velázquez-Quesada. Beliefs based on evidence and argumentation. In *Logic, Language, Information, and Computation*, pp. 289–306. Springer, 2018.
[38] Chenwei Shi, Sonja Smets and Fernando R. Velázquez-Quesada. Beliefs supported by binary arguments. *Journal of Applied Non-Classical Logics*, **28**, 165–188, 2018.
[39] J. Szymanik and R. Verbrugge. Tractability and the computational mind. In M. Sprevak and M. Colombo. eds. *The Routledge Handbook of the Computational Mind*, pp. 339–354. Routledge, 2018.
[40] J. van Benthem. Oscillations, logic, and dynamical systems. In S. Ghosh and J. Szymanik, eds, *The Facts Matter. Essays on Logic and Cognition in Honour of Rineke Verbrugge*, chapter 1, pp. 9–22. College Publications, 2015.
[41] J. van Benthem, D. Fernandez-Duque and E. Pacuit. Evidence and plausibility in neighborhood structures. *Annals of Pure and Applied Logic*, **165**, 106–133, 2014.
[42] J. van Benthem, D. Fernández-Duque and E. Pacuit., et al. Evidence logic: a new look at neighborhood structures. In *Advances in Modal Logic*, T. Bolander, T. Braüner, S. Ghilardi and L. Moss, eds, vol. **9**, pp. 97–118. College Publications, 2012.
[43] J. van Benthem and E. Pacuit. Dynamic logics of evidence-based beliefs. University of Amsterdam, ILLC, 2011.
[44] J. van Benthem and E. Pacuit. Dynamic logics of evidence-based beliefs. *Studia Logica*, **99**, 61–92, 2011.
[45] H. van Ditmarsch, W. van der Hoek and B. Kooi. *Dynamic Epistemic Logic*. Springer Science & Business Media, 2007.
[46] R. Verbrugge. Zero-one laws for provability logic: axiomatizing validity in almost all models and almost all frames. In *36th Annual ACM/IEEE Symposium on Logic in Computer Science, LICS 2021, Rome, Italy, June 29–July 2, 2021*, pp. 1–13. IEEE, 2021.
[47] K. J. S. Zollman. Social network structure and the achievement of consensus. *Politics, Philosophy & Economics*, **11**, 26–44, 2012.

Received 4 March 2021