Frequentist belief update under ambiguous evidence in social networks*

Michel Grabisch[1] and M. Alperen Yasar[1, 2]

[1]Applied Mathematics, Université Paris 1 Panthéon-Sorbonne, France
[2]Economics, Ca' Foscari University of Venice, Italy

February 4, 2024

Abstract

In this paper, we study a frequentist approach to belief updating in the framework of Dempster-Shafer Theory (DST). We propose several mechanisms that allow the gathering of possibly ambiguous pieces of evidence over time to obtain a belief mass assignment. We then use our approach to study the impact of ambiguous evidence on the belief distribution of agents in social networks. We illustrate our approach by taking three representative situations. In the first one, we suppose that there is an unknown state of nature, and agents form belief on the set of possible states. Nature constantly sends a signal which reflects the true state with some probability but which can also be ambiguous. In the second situation, there is no ground truth, and agents are against or in favor of some ethical or societal issues. In the third situation, there is no ground state either, but agents have opinions on left, center, and right political parties. We show that our approach can model various phenomena often observed in social networks, like polarization, echo chambers, and bounded confidence effects.

Keywords: Dempster-Shafer theory, agent-based modeling, social networks, ambiguous evidence, subjective belief update, opinion dynamics.

1 Introduction

Emerging behaviors of human beliefs are complex by nature. Both Bayesian and heuristic approaches have been used to study these emerging phenomena in social networks (e.g., polarization, echo chambers, bounded confidence, ...). However, empirical observations suggest that such models, while robust, do not fully encapsulate the nuances of real-world belief evolution (Chandrasekhar et al., 2020) (see Acemoglu and Ozdaglar (2011) about a theoretical paper comparing Bayesian models to heuristic models).

Agents with similar priors may have vastly different posteriors after observing the same piece of evidence (Acemoglu et al., 2010). One of the explanations for this phenomenon is the *ambiguity* of evidence. Updating beliefs under ambiguity has been challenging in the Bayesian setting. For example, Fryer Jr et al. (2019) assumes that an agent with a bias towards one of the events interprets ambiguous evidence as unequivocally supporting that favored event. It is within this context that the Dempster-Shafer Theory (DST) emerges as a potent alternative, incorporating the rationality of Bayesian models while still allowing for the intricacies of heuristics. An advantage of DST is that it need not model singleton evidence solely but rather can model evidence that might support multiple states of nature simultaneously. Therefore, two agents with identical betting probabilities, the same priors on the surface, might have two different mental models leading to different posteriors (Smets & Kennes, 2008) (see Gilboa and Schmeidler (1993) for an explanation of how non-additive probability approaches are suitable for modeling ambiguous evidence).

*This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 956107, "Economic Policy in Complex Environments (EPOC)".

DST has two mainstream belief update methodologies: combination rules (Shafer, 1976) and conditioning rules (Shafer, 1981) (see Dubois and Denoeux (2012) for a comparison of both). Both of these methods can cause radical changes in one's belief under completely opposite evidence, which might violate *belief persistence* (see Dominiak and Lee (2017) for a discussion on *belief persistence*). For this reason, we propose a frequentist update methodology for DST. Our frequentist rule acts as an evidence aggregation mechanism where every atom of evidence only nudges an agent towards a direction, which is similar to heuristic approaches (Allahverdyan & Galstyan, 2014). Using this methodology, we create several update rules for different cases and then show the emergent behavior of these updates in three different agent-based models.

The contribution of this paper is three-fold. Our first and foremost contribution is that we create a frequentist update framework for DST, addressing the challenges posed by the violation of *belief persistence*. Second, we propose several update rules for different scenarios and study their properties. Finally, we explore the application of DST in modeling belief dynamics within social networks, providing insights into emerging phenomena such as polarization or echo chambers. We demonstrate the efficacy of DST in managing ambiguous evidence, showing its unique ability to capture and process uncertainty.

The rest of the paper will be organized as follows: in Section 2, we will present the basic notation of DST that will be used later. In Section 3, we will create a frequentist framework for DST, suggest several update rules for this framework, and study the properties of these rules. Section 4 will study the emergent phenomena of DST in social networks using agent-based modeling, explaining the differences between the combination rule and the frequentist rule. Finally, Section 5 will conclude the paper with a summary of our results.

2 Basic concepts and notation

2.1 The Dempster-Shafer framework

We present basic notions of Dempster-Shafer theory (DST), which are necessary for our exposition. A detailed presentation can be found in Grabisch et al. (2016). Let Ω be a finite set of n *outcomes* (or states of nature, etc.), and 2^Ω be the power set denoting all possible subsets of outcomes, called *events*. In Dempster-Shafer Theory, an agent's belief about events is modeled by a *basic mass assignment* or *mass distribution*, abbreviated hereafter by `bma`, which is a function $m : 2^\Omega \to [0,1]$, satisfying $\sum_{X\subseteq\Omega} m(X) = 1$. Given some evidence received by the agent, $m(X)$ represents the (quantity of) belief committed exactly to the event X and not to any proper subset of X. The amount of belief committed to Ω represents the level of ignorance of the agent, while the amount of belief committed to the empty set represents the belief that Ω does not contain all possible outcomes (open world hypothesis). Throughout this paper, we assume $m(\varnothing) = 0$. We denote by $\mathcal{M}(\Omega)$ the set of all `bma`s on Ω satisfying the latter property.

We give in Table 1 an example of a mass distribution, where $\Omega = \{a, b\}$.

Events	Masses
$\varnothing$	0
$\{a\}$	0.3
$\{b\}$	0.4
Ω	0.3

Table 1: An example of a mass distribution.

Following Shafer (1976), we introduce, for a given mass distribution m of some agent,

- The *belief function* $Bel : 2^\Omega \to [0,1]$, assigning to every event X the quantity $Bel(X) = \sum_{Y\subseteq X} m(Y)$, interpreted as the degree of certainty/belief that event X realizes, i.e., the certainty that the true outcome or state of nature lies in X.

- The *plausibility function* $Pl : 2^\Omega \to [0,1]$, assigning to every event X the quantity $Pl(X) = \sum_{Y\cap X\neq\varnothing} m(Y)$, interpreted as the degree of plausibility that event X realizes, i.e., the plausibility that the true outcome or state of nature lies in X.

Therefore, following the example above, we can derive Table 2:

	Mass	Belief	Plausibility
$\varnothing$	0	0	0
$\{a\}$	0.3	0.3	0.6
$\{b\}$	0.4	0.4	0.7
Ω	0.3	1	1

Table 2: Belief and plausibility functions of the agent.

2.2 Combination rule

Consider two agents i, j, with `bmas` m_i, m_j, coming from two sets of evidence supposed to be both trustable and independent. *Dempster's rule of combination* allows to combine these two sets of evidence and to produce a new mass distribution $m_i \oplus m_j$ defined as follows:

$$(m_i \oplus m_j)(X) = \frac{\sum_{Y \cap Z = X} m_i(Y) m_j(Z)}{1 - \kappa}, \quad \varnothing \neq X \subseteq \Omega, \tag{1}$$

and $(m_i \oplus m_j)(\varnothing) = 0$, provided $1 - \kappa \neq 0$. The quantity $\kappa = \sum_{Y \cap Z = \varnothing} m_i(Y) m_j(Z)$ is the *level of conflict*, and quantifies to what extent the two pieces of evidence contradict each other*. It is easy to check that $m_i \oplus m_j$ is indeed a `bma`.

As an illustration, consider two pieces of evidence represented by mass distributions m_1, m_2 on $\Omega = \{a, b\}$ given by

$$m_1(\{a\}) = 0.6, \quad m_1(\{b\}) = 0.3, \quad m_1(\Omega) = 0.1$$
$$m_2(\{a\}) = 0.1, \quad m_2(\{b\}) = 0.5, \quad m_2(\Omega) = 0.4.$$

Table 3 gives the result of the combination with and without normalization.

	$\varnothing$	$\{a\}$	$\{b\}$	Ω
without normalization	0.33	0.31	0.32	0.04
with normalization	0	0.4627	0.4776	0.0597

Table 3: Combination of m_1, m_2

2.3 Transferable belief model

Shafer's modelling of the beliefs of an agent is done through the belief function and the plausibility function, which can be interpreted as giving an interval where the (unknown) true probability lies. The following question arises: What if an agent is forced to give a probability to make a decision? Smets and Kennes (2008) answered this question by proposing their *transferable belief model*: they assume that the modeling stage of the beliefs of the agent, given a number of evidence, is done in the framework of Shafer (they call this the *credal level*), while this belief information is transferred to the *pignistic level*, where the decision has to be made, via a betting probability, denoted by `BetP`, indicating that this is the probability P the agent would bet on. `BetP` is defined as:

$$\mathtt{BetP}(x) = \sum_{A \ni x} \frac{m(A)}{|A|} \tag{2}$$

It is well known that `BetP` is nothing other than the Shapley value (Shapley, 1953) of Bel, the belief function of the credal level.

*Smets and Kennes (2008) do not perform normalization by the denominator, since they do not assume that the mass given to $\varnothing$ has to be equal to 0. This approach can be used when the open-world hypothesis is true.

3 Frequentist belief updates in DST

We consider as above a finite set Ω of n possible states of nature or outcomes. Each agent has a prior belief on the true state of nature, which is represented by a `bma` denoted by m_0. At each time step $t = 1, 2, \ldots$, agents receive a signal or evidence e_t giving information about the true state of nature under the form of a non-empty subset of Ω. If $e_t = \{\omega\}$ for some $\omega \in \Omega$, the evidence is said to be *non-ambiguous*, while if e_t is not a singleton, then the evidence is said to be *ambiguous*. We do not discard the case $e_t = \Omega$, considered as fully ambiguous, and bringing in principle no information.

Upon receiving a signal, each agent updates her `bma` by some mechanism, and we denote by m_t the `bma` of the agent at time t, after having received signal e_t and updated.

In the sequel, we consider several updating rules, which are based on a frequentist view rather than on the Dempster's rule of combination, mimicking the way a probability is estimated by repeated observations via the law of large numbers. To the best of our knowledge, we did not find any previous attempt in this direction, as all updating mechanisms in DST are based either on Dempster's rule of combination or one of its variants or by conditioning like in Bayesian updating.

The presence of ambiguous evidence, together with different possible attitudes of the agent towards new information, opens the field of many possible updating rules. We present some examples of rules in the following subsections, ending with a general form encompassing all the given examples.

3.1 The unbiased rule

This rule is the most natural one in the DST framework. If an evidence $e_t = X$ is received, only the value of the `bma` at X is modified, nothing else. Formally:

$$m_t(X) = \begin{cases} \dfrac{t \cdot m_{t-1}(X) + 1}{t+1}, & \text{if } e_t = X \\ \dfrac{t \cdot m_{t-1}(X)}{t+1}, & \text{otherwise} \end{cases}, \quad (X \subseteq \Omega, X \neq \varnothing). \tag{3}$$

A more compact form is the following:

$$m_t(X) = \frac{t \cdot m_{t-1}(X) + \delta_{e_t}(X)}{t+1} \tag{4}$$

with $\delta_{e_t}(X) = 1$ if $e_t = X$, and 0 otherwise. It can be easily checked that

$$m_t(X) = \frac{m_0(X) + n_t(X)}{t+1}, \quad \varnothing \neq X \subseteq \Omega, \tag{5}$$

where $n_t(X)$ is the number of occurrences of X in the sequence of evidence $e_0, \ldots, e_t$. This formula shows that the order in the sequence of signals is unimportant. Also, as t tends to infinity, the influence of the prior belief tends to 0, and the mass distribution tends to the frequency of the events in the sequence, as for the law of large numbers.

Let us illustrate the rule by the following running example.

Example 1. Consider $\Omega = \{a, b\}$ with two states of nature, e.g., a means that there is a climate change, and b means that there is no climate change. An evidence e_t is brought by regular meteorological bulletins, supporting either a or b, or being ambiguous ($e_t = \Omega$).

Consider 3 agents, with the following prior `bmas` m_0^1, m_0^2, m_0^3:

	$\{a\}$	$\{b\}$	Ω
m_0^1	1	0	0
m_0^2	0	1	0
m_0^3	0	0	1

Agent 1 is fully convinced that climate change is real, while Agent 2 is fully convinced that this is a conspiracy theory, and Agent 3 comes from the Moon and has absolutely no opinion on the topic.

Consider first the following sequence of non-ambiguous evidence (omitting braces and commas for short): a, a, b, a, b, where there are slightly more pieces of evidence of a than b. Applying the unbiased rule 5 times yields the following (using (5)):

	$\{a\}$	$\{b\}$	Ω
m_5^1	$2/3$	$1/3$	0
m_5^2	$1/2$	$1/2$	0
m_5^3	$1/2$	$1/3$	$1/6$

One can see how the mass distribution of the agents tends to the frequentist distribution of the events: $3/5$ for a and $2/5$ for b. Observe also how the mass of ignorance diminishes for Agent 3.

Let us now consider that in the previous sequences, the two last pieces of evidence become ambiguous: a, a, b, ab, ab. We obtain the following:

	$\{a\}$	$\{b\}$	Ω
m_5^1	$1/2$	$1/6$	$1/3$
m_5^2	$1/3$	$1/3$	$1/3$
m_5^3	$1/3$	$1/6$	$1/2$

Ambiguity raises the level of ignorance and slows down the convergence to the frequency of appearances of a and b.

The next proposition gives two important properties of this rule.

Proposition 1. *The unbiased rule has the following properties:*

1. *Assuming m_{t-1} is a* `bma`*, m_t is also a* `bma`*.*
2. *If $e_t = X$, $m_t(X) \geqslant m_{t-1}(X)$, and more precisely*

$$m_t(X) - m_{t-1}(X) = \frac{1 - m_t(X)}{t} \geqslant 0.$$

On the other hand, $m_t(Y) \leqslant m_{t-1}(Y)$ for $Y \neq X$, more precisely

$$m_t(Y) - m_{t-1}(Y) = -\frac{m_t(Y)}{t} \leqslant 0.$$

Also, $\frac{m_t(Y)}{m_{t-1}(Y)} = \frac{t}{t+1}$.

Proof. 1. See Proposition 4.
2. From the definition of m_t, we get

$$(t+1)m_t(X) = t \cdot m_{t-1}(X) + 1$$

from which we get the desired expression (and similarly for $m_t(Y)$). Finally, the inequalities follow from the fact that m_t is a `bma` because m_0 is a `bma` and of property (1). □

The first property ensures that the updating process remains in the DST framework. The second property shows a desirable feature: the mass of the received event increases while the other ones decrease.

3.2 Distributive belief update

A distinctive feature of the unbiased rule is that every event is treated similarly, and the fact that the evidence is ambiguous does not matter. In particular, receiving the totally ambiguous signal Ω just augments the mass of Ω. However, we might argue that $e_t = \Omega$ does not bring any information, so why m_t should be different from m_{t-1}?

The distributive rule considers that when an ambiguous evidence X is received, it could have been any subset of X, and the agent updates the masses of every subset of X *in proportion of her belief on these subsets*: the more the agent believes that Z is true, the more she is inclined to "interpret" any new evidence $X \supseteq Z$ as a manifestation of Z. Hence, ambiguous evidence serves to reinforce the belief on less ambiguous evidence.

We use the notation $\hat{m}$ to denote this updating rule, defined as follows:

$$\hat{m}_t(X) = \begin{cases} \dfrac{t \cdot \hat{m}_{t-1}(X) + \hat{m}_{t-1}(X : e_t)}{t+1}, & \text{if } X \subseteq e_t \\ \dfrac{t \cdot \hat{m}_{t-1}(X)}{t+1}, & \text{otherwise} \end{cases}, \quad (X \subseteq \Omega, X \neq \varnothing), \tag{6}$$

with

$$\hat{m}(X : Z) := \begin{cases} \dfrac{\hat{m}(X)}{\sum_{Y \subseteq Z} \hat{m}(Y)}, & \text{if } \sum\limits_{Y \subseteq Z} \hat{m}(Y) > 0 \\ 1, & \text{otherwise} \end{cases}, \quad (\varnothing \neq X \subseteq Z). \tag{7}$$

A simpler formulation is:

$$\hat{m}_t(X) = \frac{t \cdot \hat{m}_{t-1}(X) + \hat{m}_{t-1}(X : e_t)\delta_{\subseteq e_t}(X)}{t+1} \tag{8}$$

with $\delta_{\subseteq e_t}(X) = 1$ if $X \subseteq e_t$, and 0 otherwise.

We give a number of properties of the distributive rule.

Proposition 2. *The distributive rule has the following properties:*

1. *If $\hat{m}_{t-1}$ is a* `bma`*, then $\hat{m}_t$ is a* `bma` *too.*
2. *If e_t is a singleton, then $\hat{m}_t = m_t$, i.e., the distributive rule coincides with the unbiased rule.*
3. *If $e_t = \Omega$, then $\hat{m}_t = \hat{m}_{t-1}$.*
4. *Suppose $Z \subseteq X \subseteq Y$. Then*

$$\hat{m}_t(Z \mid e_t = X) \geqslant \hat{m}_t(Z \mid e_t = Y),$$

where the conditioning bar indicates the observed evidence.

Proof. 1. See Proposition 4.

2. Suppose $e_t = \{\omega\}$. Then $\hat{m}_{t-1}(X : \{\omega\})$ reduces to $\hat{m}_{t-1}(\{\omega\} : \{\omega\}) = 1$. Therefore

$$\hat{m}_t(X) = \begin{cases} \dfrac{t \cdot \hat{m}_{t-1}(\{\omega\}) + 1}{t+1}, & \text{if } X = \{\omega\} \\ \dfrac{t \cdot \hat{m}_{t-1}(X)}{t+1}, & \text{otherwise} \end{cases},$$

which coincides with the unbiased rule.

3. Suppose $e_t = \Omega$. Then for every $\varnothing \neq X \subseteq \Omega$

$$\hat{m}_{t-1}(X : \Omega) = \frac{\hat{m}_{t-1}(X)}{\sum_{Y \subseteq \Omega} \hat{m}_{t-1}(Y)} = \hat{m}_{t-1}(X)$$

since $\hat{m}_{t-1}$ is a `bma` (because m_0 is a `bma` and of property (1)). Then

$$\hat{m}_t(X) = \frac{t \cdot \hat{m}_{t-1}(X) + \hat{m}_{t-1}(X)}{t+1} = \hat{m}_{t-1}(X)$$

for every $\varnothing \neq X \subseteq \Omega$.

4. Suppose $Z \subseteq X \subseteq Y$. Observe that

$$\hat{m}_{t-1}(Z : X) = \frac{\hat{m}_{t-1}(Z)}{\sum_{K \subseteq X} \hat{m}_{t-1}(K)} \geqslant \frac{\hat{m}_{t-1}(Z)}{\sum_{K \subseteq Y} \hat{m}_{t-1}(K)} = \hat{m}_{t-1}(Z : Y)$$

because $\hat{m}_{t-1}$ is nonnegative as it is a `bma`. The desired result follows from this inequality. $\square$

The first property shows that this rule also remains in the DST framework, and the last property shows that belief increases more when the evidence is less ambiguous.

We may wonder if a kind of monotonicity property holds, as for the unbiased rule (see Proposition 1 (1)), but this is not true. Indeed, supposing $e_t = Z$ is received and $\sum_{Y \subseteq Z} \hat{m}_{t-1}(Y) > 0$, it is not possible to tell whether $\hat{m}_t(X)$ is greater or smaller than $\hat{m}_{t-1}(X)$ for $X \subseteq Z$, since we get from (6)

$$\hat{m}_t(X) - \hat{m}_{t-1}(X) = \frac{1}{t}\Big(\frac{\hat{m}_{t-1}(X)}{\sum_{Y \subseteq e_t} \hat{m}_{t-1}(Y)} - \hat{m}_t(X)\Big)$$

when $\sum_{Y \subseteq e_t} \hat{m}_{t-1}(Y) \neq 0$. The sign of the expression in parentheses may be positive or negative, depending on the values taken by $\hat{m}_t$ and $\hat{m}_{t-1}$. However, as for the unbiased rule, if $X \not\subseteq e_t$, then $\hat{m}_t(X)$ decreases.

Example 2 (Example 1 cont'd)**.** Let us apply the distributive rule to our previous example. Since for the sequence a, a, b, a, b of non-ambiguous events, the result is identical (see Proposition 2 (2)), we turn to the second sequence a, a, b, ab, ab. We obtain:

	$\{a\}$	$\{b\}$	Ω
$\hat{m}_5^1$	$3/4$	$1/4$	0
$\hat{m}_5^2$	$1/2$	$1/2$	0
$\hat{m}_5^3$	$1/2$	$1/4$	$1/4$

From Proposition 2 (2) and (3), we get in fact $\hat{m}_5^i = m_3^i$ for $i = 1, 2, 3$. We observe that in contrast to the unbiased rule, agents 1 and 2 have no ignorance, as the ambiguous evidence are not taken into account.

We present a second example to illustrate the computation with three states of nature.

Example 3. Let $\Omega = \{a, b, c\}$ be the set of all possible outcomes. Let

Event X	$\{a\}$	$\{b\}$	$\{c\}$	$\{a,b\}$	$\{a,c\}$	$\{b,c\}$	Ω
$m_0(X)$	0.3	0.1	0.2	0.1	0.1	0.1	0.1

be the prior distribution m_0 of an agent. Let $e_1 = \{a, b\}$ be the evidence at time 1. Then, the distributive update is obtained as follows (we omit braces and commas):

$$\begin{aligned}
\hat{m}_1(a) &= \frac{m_0(a) + \frac{m_0(a)}{m_0(a)+m_0(b)+m_0(ab)}}{2} = \frac{0.3 + \frac{0.3}{0.3+0.1+0.1}}{2} = 0.45 \\
\hat{m}_1(b) &= \frac{0.1 + \frac{0.1}{0.3+0.1+0.1}}{2} = 0.15 \\
\hat{m}_1(c) &= \frac{m_0(c)}{2} = 0.1 \\
\hat{m}_1(ab) &= \frac{0.1 + \frac{0.1}{0.3+0.1+0.1}}{2} = 0.15 \\
\hat{m}_1(ac) &= \frac{m_0(ac)}{2} = 0.05 \\
\hat{m}_1(bc) &= \frac{m_0(bc)}{2} = 0.05 \\
\hat{m}_1(abc) &= \frac{m_0(abc)}{2} = 0.05
\end{aligned} \tag{9}$$

Notice that the mass of a increased more than the mass of b, while the relative ratio remained the same. Now let us calculate the $\texttt{BetP}$ of these masses for the previous turn: $\texttt{BetP}_{t-1}(a) = 0.4333$, $\texttt{BetP}_{t-1}(b) = 0.2333$, $\texttt{BetP}_{t-1}(c) = 0.3333$. And the updated betting probabilities are: $\texttt{BetP}_t(a) = 0.56666$, $\texttt{BetP}_t(b) = 0.26666$, $\texttt{BetP}_t(c) = 0.1666$. Even though the evidence supported both states of nature a and b, the agent's prior caused the agent to support evidence a more. This result parallels with Fryer Jr et al. (2019) in terms of ambiguous evidence.

3.3 The biased rule when $\Omega = \{a, b\}$

We place ourselves in the case of two possible states of nature, as in our running example with climate change. The idea behind this rule is the following: when the agent is faced with ambiguous evidence Ω, she updates her belief on a and b in proportion to her current belief as in the distributive rule, but also in proportion to her level of certainty, that is, $1 - m_t(\Omega)$ (recall that $m_t(\Omega)$ is the level of ignorance).

Let us denote by $\tilde{m}_{t+1}$ the `bma` updated by the biased rule. When $e_t = \{a\}$ or $e_t = \{b\}$, the formula is identical to the unbiased one: $\tilde{m}_t(X) = m_t(X)$ for all $X \subseteq \Omega$, $X \neq \varnothing$. When $e_t = \Omega$, the formula is:

$$\tilde{m}_t(X) = \begin{cases} \dfrac{t \cdot \tilde{m}_{t-1}(X) + \tilde{m}_{t-1}(X)(1 - \tilde{m}_{t-1}(\Omega))}{t+1}, \text{ if } X = \{a\} \text{ or } X = \{b\} \\ \tilde{m}_{t-1}(X)\Big(\dfrac{t + 2 - \tilde{m}_{t-1}(X)}{t+1}\Big), \text{ if } X = \Omega. \end{cases} \tag{10}$$

This rule has a number of interesting properties, given in the next proposition below.

Proposition 3. *The biased rule has the following properties:*

1. *When the evidence is a singleton, all three rules coincide.*
2. *If $\tilde{m}_{t-1}$ is a `bma`, then $\tilde{m}_t$ is a `bma`.*
3. *Suppose $e_t = \Omega$ and $0 < \tilde{m}_{t-1}(\Omega) < 1$. Then $\tilde{m}_t(\Omega) > \tilde{m}_{t-1}(\Omega)$, implying that $\tilde{m}_t(\Omega)$ converges to 1 when $e_t = \Omega$ for all $t \geqslant t_0$.*
4. *Supposing $e_t = \Omega$ for all t and $0 < m_0 < 1$, we have $\tilde{m}_t(\Omega) < m_t(\Omega)$ for all t. Consequently, $\tilde{m}$ converges more slowly than m.*
5. *Suppose $e_t = \Omega$. Then*
$$\frac{\tilde{m}_t(\{a\})}{\tilde{m}_t(\{b\})} = \frac{\tilde{m}_{t-1}(\{a\})}{\tilde{m}_{t-1}(\{b\})}$$

Proof. 1. Clear.

2. See Proposition 4.
3. We have from the definition
$$\tilde{m}_t(\Omega) - \tilde{m}_{t-1}(\Omega) = \frac{\tilde{m}_{t-1}(\Omega)}{t+1}(1 - \tilde{m}_{t-1}(\Omega)).$$
As $0 \leq \tilde{m}_{t-1}(\Omega) \leq 1$ for every t, the sequence $\tilde{m}_t(\Omega)$ is nondecreasing, and is strictly increasing iff $0 < \tilde{m}_{t-1}(\Omega) < 1$ for every t. Observe that $m_0(\Omega) = 0$ entails $\tilde{m}_t(\Omega) = 0$ for every t, and $m_0(\Omega) = 1$ entails $\tilde{m}_t(\Omega) = 1$ for every t. It follows that, under that assumption $0 < m_0(\Omega) < 1$, $\tilde{m}_t(\Omega)$ converges to 1.
4. Let us show by induction that starting from the same prior $0 < m_0 < 1$, $\tilde{m}_t(\Omega) < m_t(\Omega)$ for every t. From (3) and (10)) with $t = 1$, we obtain:
$$\begin{aligned}\tilde{m}_1(\Omega) = m_0(\Omega)\frac{3 - m_0(\Omega)}{2} &= \frac{3m_0(\Omega) - (m_0(\Omega))^2}{2} \\ &= \frac{m_0(\Omega) + 1 - (m_0(\Omega) - 1)^2}{2} < \frac{m_0(\Omega) + 1}{2} = m_1(\Omega).\end{aligned}$$
Let us assume the hypothesis till $t - 1$. We have from (10) again
$$\begin{aligned}\tilde{m}_t(\Omega) = \tilde{m}_{t-1}(\Omega)\Big(\frac{t + 2 - \tilde{m}_{t-1}(\Omega)}{t+1}\Big) &= \frac{\tilde{m}_{t-1}(\Omega)(t+2) - (\tilde{m}_{t-1}(\Omega))^2}{t+1} \\ &= \frac{-(\tilde{m}_{t-1}(\Omega) - 1)^2 + 1 + \tilde{m}_{t-1}(\Omega)t}{t+1} \\ &< \frac{\tilde{m}_{t-1}(\Omega) + 1}{t+1} < \frac{m_{t-1}(\Omega)t + 1}{t+1} = m_t(\Omega),\end{aligned} \tag{11}$$
where the last inequality comes from the induction hypothesis.

5. Let us assume a piece of ambiguous evidence, the updated beliefs are:

$$\begin{aligned}\tilde{m}_t(\{a\}) &= \frac{\tilde{m}_{t-1}(\{a\}) \cdot t + \tilde{m}_{t-1}(\{a\}) \cdot (1 - (1 - \tilde{m}_{t-1}(\{a\}) - \tilde{m}_{t-1}(\{b\})))}{t+1} \\ &= \frac{\tilde{m}_{t-1}(\{a\}) \cdot t + \tilde{m}_{t-1}(\{a\})(\tilde{m}_{t-1}(\{a\}) + \tilde{m}_{t-1}(\{b\}))}{t+1}\end{aligned}$$

and

$$\begin{aligned}\tilde{m}_t(\{b\}) &= \frac{\tilde{m}_{t-1}(\{b\}) \cdot t + \tilde{m}_{t-1}(\{b\}) \cdot (1 - (1 - \tilde{m}_{t-1}(\{a\}) - \tilde{m}_{t-1}(\{b\})))}{t+1} \\ &= \frac{\tilde{m}_{t-1}(\{b\}) \cdot t + \tilde{m}_{t-1}(\{b\})(\tilde{m}_{t-1}(\{a\}) + \tilde{m}_{t-1}(\{b\}))}{t+1}.\end{aligned}$$

The ratio between new masses is:

$$\begin{aligned}\frac{\tilde{m}_t(\{a\})}{\tilde{m}_t(\{b\})} &= \frac{\frac{\tilde{m}_{t-1}(\{a\}) \cdot t + \tilde{m}_{t-1}(\{a\})(\tilde{m}_{t-1}(\{a\}) + \tilde{m}_{t-1}(\{b\}))}{t+1}}{\frac{\tilde{m}_{t-1}(\{b\}) \cdot t + \tilde{m}_{t-1}(\{b\})(\tilde{m}_{t-1}(\{a\}) + \tilde{m}_{t-1}(\{b\}))}{t+1}} \\ &= \frac{\tilde{m}_{t-1}(\{a\}) \cdot t + \tilde{m}_{t-1}(\{a\})(\tilde{m}_{t-1}(\{a\}) + \tilde{m}_{t-1}(\{b\}))}{\tilde{m}_{t-1}(\{b\}) \cdot t + \tilde{m}_{t-1}(\{b\})(\tilde{m}_{t-1}(\{a\}) + \tilde{m}_{t-1}(\{b\}))} \\ &= \frac{\tilde{m}_{t-1}(\{a\})(t + \tilde{m}_{t-1}(\{a\}) + \tilde{m}_{t-1}(\{b\}))}{\tilde{m}_{t-1}(\{b\})(t + \tilde{m}_{t-1}(\{a\}) + \tilde{m}_{t-1}(\{b\}))} \\ &= \frac{\tilde{m}_{t-1}(\{a\})}{\tilde{m}_{t-1}(\{b\})} \qquad (12)\end{aligned}$$

□

Suppose $e_t = \Omega$. As observed, the distributive rule yields $\hat{m}_t = \hat{m}_{t-1}$, which is different from the biased rule. Observe that if $\tilde{m}_{t-1}(\Omega) = 1$ (total uncertainty) or $\tilde{m}_{t-1}(\Omega) = 0$ (no uncertainty), then $\tilde{m}_t = \tilde{m}_{t-1}$. In particular, $\tilde{m}_t(\Omega)$ remains 1 or 0, respectively.

Example 4 (Example 1 cont'd)**.** Let us apply the biased rule to our previous example. As for the sequence a, a, b, a, b of non-ambiguous events, the result is identical to the previous ones (see Proposition 3 (1)), we turn to the second sequence a, a, b, ab, ab. We obtain:

	$\{a\}$	$\{b\}$	Ω
$\tilde{m}_5^1$	$3/4$	$1/4$	0
$\tilde{m}_5^2$	$1/2$	$1/2$	0
$\tilde{m}_5^3$	0.4522	0.2261	0.3216

One can see that as expected, $\tilde{m}_5(\Omega)$ remains smaller than $m_5(\Omega)$ (see Proposition 3 (4), and recall that $\tilde{m}_3 = m_3$).

3.4 Generalized formula

All previous rules can be cast into the following general rule:

$$m_t^f(X) = \frac{t \cdot m_{t-1}^f(X) + f(m_{t-1}^f, X, e_t)}{t+1}, \quad (\varnothing \neq X \subseteq \Omega), \qquad (13)$$

with $f : \mathcal{M}(\Omega) \times (2^\Omega \setminus \{\varnothing\}) \times (2^\Omega \setminus \{\varnothing\}) \to [0,1]$ is a function satisfying the property

$$\sum_{\varnothing \neq X \subseteq \Omega} f(m, X, Y) = 1.$$

Such a function f satisfying the above condition is called an *updating function* or *updator*. Indeed, we have:

- For the unbiased rule,
$$f(m_{t-1}, X, e_t) = \delta_{e_t}(X).$$
- For the distributive rule,
$$f(\hat{m}_{t-1}, X, e_t) = \delta_{\subseteq e_t}(X)\hat{m}_{t-1}(X : e_t).$$
- For the biased rule,
$$f(\tilde{m}_{t-1}, X, e_t) = \begin{cases} \delta_{e_t}(X), & \text{if } e_t = \{a\} \text{ or } e_t = \{b\} \\ \tilde{m}_{t-1}(X)(1 - \tilde{m}_{t-1}(\Omega)), & \text{if } e_t = \Omega \text{ and } X \neq \Omega \\ \tilde{m}_{t-1}(\Omega)\Big(\frac{t+2-\tilde{m}_{t-1}(\Omega)}{t+1}\Big), & \text{if } e_t = X = \Omega. \end{cases}$$

It is easy to check that all the above functions f are indeed updating functions.

The next proposition gives some basic properties of this class of updating rules. First, we say that an updating rule based on an updating function f is *strictly monotone* if for all $t > 0$, $m^f_t(X) > m^f_{t-1}(X)$ whenever $e_t = X$ and $m^f_{t-1}(X) < 1$.

Proposition 4. *Let f be an updating function and its corresponding updating rule. The following holds:*

1. *If m^f_{t-1} is a* `bma`*, then m^f_t is a* `bma`*.*
2. *The updating rule is strictly monotone iff*
$$f(m^f_{t-1}, X, X) > m^f_{t-1}(X), \forall \varnothing \neq X \subseteq \Omega, \forall t = 1, 2, \ldots$$

Proof. 1. As f and m^f_{t-1} take values in $[0, 1]$, we have clearly $m^f_t \geqslant 0$. Also, for every $X \subseteq \Omega$, $X \neq \varnothing$,
$$m^f_t(X) \leqslant \frac{t \cdot m^f_{t-1}(X) + 1}{t+1} \leqslant \frac{t+1}{t+1} = 1.$$
Let us prove that $\sum_{X\subseteq\Omega} m^f_t(X) = 1$. We have
$$\sum_{X\subseteq\Omega} m^f_t(X) = \frac{1}{t+1}\left(t \sum_{X\subseteq\Omega} m^f_{t-1}(X) + \sum_{X\subseteq\Omega} f(m^f_{t-1}, X, e_t)\right) = \frac{1}{t+1}(t+1) = 1.$$

2. By definition we have
$$m^f_t(X) - m^f_{t-1}(X) = \frac{1}{t+1}\Big[-m^f_{t-1}(X) + f(m^f_{t-1}, X, X)\Big]$$
which is positive iff $f(m^f_{t-1}, X, X) > m^f_{t-1}(X)$. □

Let us study the convergence properties of the family of updating rules. Let f be an updating function and m_0 some initial `bma`. Given a sequence of evidence $e_1, e_2, \ldots, e_t$, we have the following, for every $\varnothing \neq X \subseteq \Omega$:
$$\begin{aligned} m^f_1(X) &= \frac{m_0(X) + f(m_0, X, e_1)}{2} \\ m^f_2(X) &= \frac{m_0(X) + f(m_0, X, e_1) + f(m^f_1, X, e_2)}{3} \\ m^f_3(X) &= \frac{m_0(X) + f(m_0, X, e_1) + f(m^f_1, X, e_2) + f(m^f_2, X, e_3)}{4} \\ &\vdots \quad \vdots \\ m^f_t(X) &= \frac{m_0(X) + \sum_{k=0}^{t-1} f(m^f_k, X, e_{k+1})}{t+1} \leqslant \frac{m_0(X) + t}{t+1}. \end{aligned}$$

It follows that

$$\lim_{t\to\infty} m_t^f(X) = \lim_{t\to\infty} \frac{1}{t}\sum_{k=1}^{t} f(m_{k-1}^f, X, e_k), \tag{14}$$

provided the limit exists. Let us see some particular cases.

1. Suppose $e_t = X$, $t = 1, 2, \ldots$, and consider the unbiased rule. We have in this case $f(m_k, X, X) = 1$, therefore by (14),

$$\lim_{t\to\infty} m_t(X) = \lim_{t\to\infty} \frac{1}{t}\sum_{k=1}^{t} 1 = 1.$$

2. Suppose e_t is a random sequence where $e_t = X$ with probability p, and consider the unbiased rule again. Then, by the strong law of large numbers, $m_t(X)$ converges to p almost surely.

3. Suppose again that $e_t = X$, $t = 1, 2, \ldots$, and consider the distributive rule. The following can be shown.

Proposition 5. *Suppose $m_0(X) > 0$ and that $e_t = X$ for $t = 1, 2, \ldots$. Then*

$$\lim_{t\to\infty} \hat{m}_t(X) = m_0(X : X).$$

The proof relies on the following lemma.

Lemma 1. *Suppose $m_0(X) > 0$ and that $e_t = X$ for $t = 1, 2, \ldots$. Then, by the distributive rule, for any $t = 1, 2, \ldots$,*

$$\hat{m}_t(X : X) = m_0(X : X), \quad \hat{m}_t(Y : X) = m_0(Y : X) \textit{ for every } Y \subset X.$$

Proof. We show the result by induction. Let us put

$$\begin{aligned} m_0(X : X) &= \frac{m_0(X)}{\sum_{Y\subset X} m_0(Y) + m_0(X)} =: \alpha, \\ m_0(Y : X) &= \frac{m_0(Y)}{\sum_{Y\subset X} m_0(Y) + m_0(X)} =: \alpha_Y \text{ for any } Y \subset X. \end{aligned}$$

Note that $\alpha_Y = 0$ iff $m_0(Y) = 0$ and that $\sum_{Y\subset X} \alpha_Y + \alpha = 1$. We have

$$\hat{m}_1(X) = \frac{1}{2}\big(m_0(X) + \alpha\big), \quad \hat{m}_1(Y) = \frac{1}{2}\big(m_0(Y) + \alpha_Y\big).$$

Therefore

$$\begin{aligned} m_1(X : X) &= \frac{\hat{m}_1(X)}{\sum_{Y\subset X} \hat{m}_1(Y) + \hat{m}_1(X)} \\ &= \frac{m_0(X) + \alpha}{\sum_{Y\subset X} m_0(Y) + m_0(X) + 1} \\ &= \frac{m_0(X) + \alpha}{\frac{1}{\alpha} m_0(X) + 1} = \alpha. \end{aligned}$$

Similarly, for any $Y \subset X$:

$$\begin{aligned} m_1(Y : X) &= \frac{\hat{m}_1(Y)}{\sum_{Y\subset X} \hat{m}_1(Y) + \hat{m}_1(X)} \\ &= \frac{m_0(Y) + \alpha_Y}{\sum_{Y\subset X} m_0(Y) + m_0(X) + 1} \\ &= \frac{m_0(Y) + \alpha_Y}{\frac{1}{\alpha_Y} m_0(Y) + 1} = \alpha_Y. \end{aligned}$$

Let us suppose the property true till t and let us prove it for $t+1$. We have, using the induction hypothesis:

$$\hat{m}_{t+1}(X) = \frac{(t+1)\hat{m}_t(X) + \alpha}{t+2},$$
$$\hat{m}_{t+1}(Y) = \frac{(t+1)\hat{m}_t(Y) + \alpha_Y}{t+2}$$

for any $Y \subset X$. Therefore

$$\begin{aligned} m_{t+1}(X:X) &= \frac{\hat{m}_{t+1}(X)}{\sum_{Y\subset X} \hat{m}_{t+1}(Y) + \hat{m}_{t+1}(X)} \\ &= \frac{(t+1)\hat{m}_t(X) + \alpha}{(t+1)\Big(\sum_{Y\subset X} \hat{m}_t(Y) + \hat{m}_t(X)\Big) + 1} \\ &= \frac{(t+1)\hat{m}_t(X) + \alpha}{\frac{t+1}{\alpha}\hat{m}_t(X) + 1} = \alpha. \end{aligned}$$

Similarly, for any $Y \subset X$:

$$\begin{aligned} m_{t+1}(Y:X) &= \frac{\hat{m}_{t+1}(Y)}{\sum_{Y\subset X} \hat{m}_{t+1}(Y) + \hat{m}_{t+1}(X)} \\ &= \frac{(t+1)\hat{m}_t(Y) + \alpha_Y}{(t+1)\Big(\sum_{Y\subset X} \hat{m}_t(Y) + \hat{m}_t(X)\Big) + 1} \\ &= \frac{(t+1)\hat{m}_t(Y) + \alpha_Y}{\frac{t+1}{\alpha_Y}\hat{m}_t(Y) + 1} = \alpha_Y. \end{aligned}$$

□

Proof. (of Proposition 5) By using (14) and Lemma 1, we find, under constant evidence $e_t = X$, $t = 1, 2, \ldots,$:

$$\begin{aligned} \lim_{t\to\infty} \hat{m}_t(X) &= \lim_{t\to\infty} \frac{1}{t} \sum_{k=1}^{t} f(\hat{m}_{k-1}, X, X) \\ &= \lim_{t\to\infty} \frac{1}{t} \sum_{k=1}^{t} \hat{m}_{k-1}(X:X) \\ &= \lim_{t\to\infty} \frac{1}{t} t \cdot m_0(X:X) = m_0(X:X). \end{aligned}$$

Similarly,

$$\lim_{t\to\infty} \hat{m}_t(Y) = m_0(Y:X)$$

for any $Y \subset X$. □

4 Agent-based models

In order to understand how beliefs evolve in a social network via different settings and update rules, we create three agent-based models. In the first setting, there is a true state of nature, like in Example 1. In each turn, nature constantly sends signals. After updating their beliefs, agents discuss with their neighbors. In the second setting, we create a model where agents form opinions about an ethical issue with no true state; therefore, agents form opinions via their discussions with their neighbors. Finally, we extend our second model to a three-party setting, e.g., $\Omega = \{l, c, r\}$.

To study opinion dynamics, we create social networks via the Barabasi-Albert algorithm in all of our models (Barabási & Albert, 1999), which proceeds as follows. In order to generate a network with $N = \{1, 2, \ldots, n\}$, we initially assume $|N| = 1$. Then, in each step t, we add a new agent n_t until

$|N| = n$. Agent n_t connects to agent i of degree k_i with probability $p_i = \frac{k_i}{\sum_j k_j}$. This algorithm gives us a scale-free network whose degree distribution follows a power law. Its main characteristic is that few nodes will have many connections, and many nodes will have only a few connections [†]. The main advantage of scale-free networks is that there are few extremely influential agents, which sometimes poses a challenge for asymptotic learning (Acemoglu et al., 2011).

In this section, we will use three of the rules we described in Section 3: Dempster's combination rule from Equation (1), the distributive rule from Equation (6) when $|\Omega| = 3$, and the biased rule from Equation (10) when $|\Omega| = 2$. The precise way of combination and belief updating will be explained in each subsection. We nevertheless make an essential remark about Dempster's rule of combination. Recall that it is not defined when the level of conflict κ is equal to 1 in order to avoid division by 0. Practically, we do not perform the combination if κ exceeds $d = 0.9$. It means that agents with very different beliefs do not update; hence, their opinions are not affected, as if they would not communicate at all. This reminds the *bounded confidence models* in social networks, where agents delete links with agents with which they have too distant opinions (see Grabisch and Rusinowska (2020) for an extensive literature review about bounded confidence models and the effect of the distance).

In our agent-based model, we use a total of 9 parameters. We run each simulation on average $T = 2000$ turns, where beliefs stabilize. Moreover, we create $R = 1000$ simulations for each parameter combination and then take the average results. In order to facilitate network visualization, we take $N = 200$ agents in the network. We have also tested the robustness of our results with 2000 agents, where we have observed no noticeable difference in results. We create agents with a prior belief mass assignment randomly selected from a Dirichlet distribution in each run, e.g., $m_0^i \sim \mathrm{Dir}(\alpha_\varnothing, \ldots, \alpha_\Omega)$, $\forall i \in N$, where $\alpha_\varnothing, \ldots, \alpha_\Omega$ are Dirichlet parameters that denote the weight of outcome of $\varnothing, \ldots, \Omega$ in the prior (see A for a detailed explanation on our reasoning for using Dirichlet distributions). Table 4 describes the parameters used in the simulation.

Parameter	Value range	Description
T	2000	Number of turns per run;
R	1000	Number of runs per parameter combination;
N	200	Number of agents in the network;
α_H	20	A high weight Dirichlet parameter;
α_M	1	An unbiased agent's Dirichlet parameter;
α_L	0.1	A low weight Dirichlet parameter;
p	$(0,5,1]$	The probability of observing the true state if the evidence is non-ambiguous;
q	$[0,1)$	The probability of observing a non-ambiguous evidence;
π	0.01	The proportion of edges that are selected to communicate;
d	0.9	Bounded confidence parameter;
k	1	Neighbors to be connected in Barabasi-Albert algorithm (scale-free).

Table 4: Parameters of simulations.

4.1 Existence of a true unknown state

This section studies the process of belief update when there is a true state of nature that is unknown, while nature constantly sends a signal revealing the true state with a probability p, the signal being possibly ambiguous. We take as framework Example 1, with $\Omega = \{a, b\}$ where a means that climate

[†]There is a dispute in the literature based on the shape of networks encountered in empirical data. Some studies claim that real data follows a scale-free shape (Mislove et al., 2007), while a recent paper challenges that the real data is rarely scale-free (Broido & Clauset, 2019). However, this challenge is based on how empirical observations are more complex and have underlying unique properties. Therefore, we have also tested our model using a Watts-Strogatz algorithm (Watts & Strogatz, 1998), where agents are in a "small world", i.e., they can reach one another in very few steps. The significance of our results remains the same.

change is real, while b indicates the that there is no climate change. We distinguish two different procedures for updating beliefs, detailed in the following subsections.

4.1.1 Communication via Dempster's rule of combination

Algorithm 1 Dempster's rule with a true state

A network $\mathcal{G}$ is randomly generated via Barabasi-Albert algorithm;
for every time step t **do**
    Nature determines e_t:
    if $e_t = X \neq \Omega$ **then**
        $\mathbb{P}(e_t = X) = p_X * q$; $X \neq \varnothing$;
    else
        $\mathbb{P}(e_t = \Omega) = 1 - q$
    end if
    for every agent i **do**
        $m_t^i(Y)$ is computed by Eq. 10, $\forall Y \subseteq \Omega$;
    end for
    A proportion π of links is selected randomly;
    for every selected link ij **do**
        Compute the level of conflict κ between m_t^i and m_t^j;
        if $\kappa \leqslant d$ **then**
            $m_t^i = m_t^j \leftarrow m_t^i \oplus m_t^j$;
        else state No update occurs;
        end if
    end for
end for

Supposing the true state of nature is a, nature sends ambiguous signal Ω with probability 0.1; otherwise, nature sends a signal a with probability 0.8 and b with probability 0.2. At each time step t:

1. Each agent i updates its `bma` m_{t-1}^i with the signal X sent by Nature by using the biased rule (Eq. (10)).

2. Select a proportion π of links. For each link ij selected, the agents i, j update their `bmas` m_t^i, m_t^j by replacing them by $m_{t-1}^i \oplus m_{t-1}^j$ (Dempster's rule), provided the level of conflict κ does not exceed d (bounded confidence). Otherwise, no update occurs.

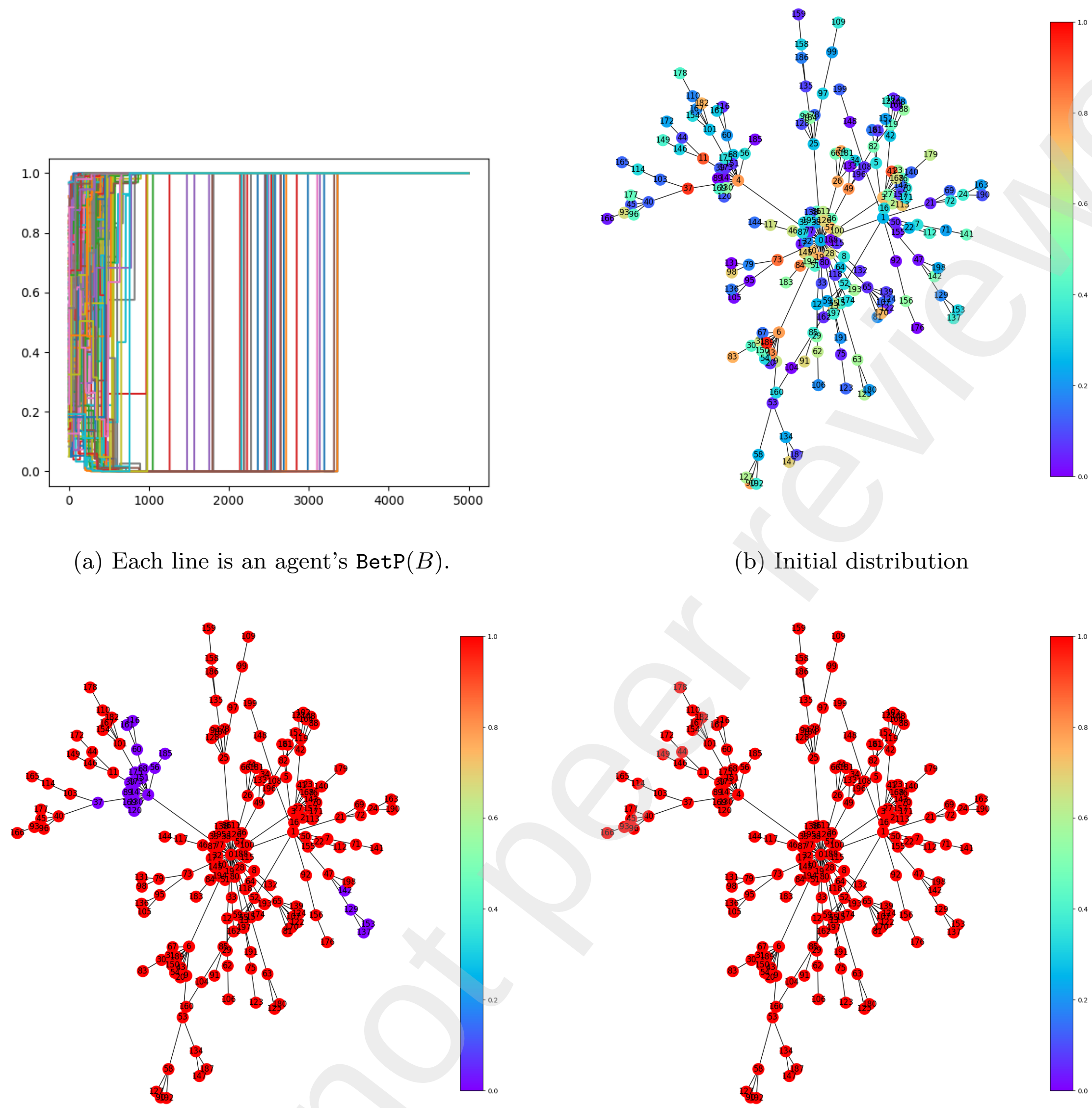
Algorithm 1 summarizes the whole procedure.

Figure 1 shows that agents start fully believing in climate change if nature consistently sends evidence for it. In this case, nature is a constant persuader with a rigid opinion. This persuader parallels the literature, where a constant persuader often prevails. However, we also observe short-term polarization, where agents first polarize in two distinct opinions. Then, the *wrong* side starts losing confidence after constant exposure to nature. Finally, everyone converges to full belief.

What is also interesting is the fact that agents' $m(B)$ values do not converge to the true $p \times q$ value when they communicate via Dempster's combination rule. Instead, it converges to full belief $m(B) = 1$. The combination rule causes agents to believe they are entirely correct and that there is no other probability. This behavior comes from the combination rule being self-reinforcing, meaning that two agents with the same opinion will reinforce their beliefs and not stay in the position. This result is similar to models that have been encountered in opinion dynamics models (Allahverdyan and Galstyan, 2014; Baumann et al., 2020).

4.1.2 Communication via frequentist update

This section assumes that agents communicate via the biased update rule from Equation (10). Frequentist update ensures that agents are not communicating thoroughly as they did via Dempster's combination rule; instead, they communicate with a softer approach where agents change their opinions slowly. More precisely, at each time step t:



(a) Each line is an agent's $\texttt{BetP}(B)$.

(b) Initial distribution

(c) After 20000 turns.

(d) Final distribution

Figure 1: Created via Algorithm 1. Red indicates a belief in climate change ($\texttt{BetP}(B) = 1$). $p = 0.8, q = 0.9$

1. Updating with a signal of nature is identical to the previous case.
2. Select a proportion π of links. For each link ij selected:
   (a) Agent i sends a signal $X \subseteq \Omega$ to Agent j with probability $m_t^i(X)$, while simultaneously Agent j sends a signal $Y \subseteq \Omega$ to Agent i with probability $m_t^j(Y)$.
   (b) Both agents i and j update via the frequentist biased rule (10).

Algorithm 2 summarizes this procedure, and Figure 2 shows our results.

We observe from Figure 2 that agents now converge to a value close to the true pq value, which is 0.72. However, it is not easy to separate the effect of communication when there is a constant persuader: nature. The difference here is that agents take tens of thousands of turns to converge to the true state, while Dempster's rule stabilizes after 3000 turns on average. Furthermore, our time plot from Figure 2a shows no polarization at any given time.

Algorithm 2 Frequentist rule with a true state

A network $\mathcal{G}$ is randomly generated via Barabasi-Albert algorithm;
for every time step t **do**
  Nature determines e_t:
  if $e_t = X \neq \Omega$ **then**
    $\mathbb{P}(e_t = X) = p_X * q$; $X \neq \varnothing$;
  else
    $\mathbb{P}(e_t = \Omega) = 1 - q$
  end if
  for every agent i **do**
    $m_t^i(Y)$ is computed by Eq. 10, $\forall Y \subseteq \Omega$;
  end for
  A proportion π of links is selected randomly;
  for every selected link ij **do**
    Agent i sends signal $X \subseteq \Omega$ to agent j with probability $m_t^i(X)$
    Agent j sends signal $Y \subseteq \Omega$ to agent i with probability $m_t^i(Y)$
    $m_t^i(Z)$ is computed by Eq. 10 with $e_t = X$ and m_t^i, $\forall Z$;
    $m_t^j(Z)$ is computed by Eq. 10 with $e_t = Y$ and m_t^j, $\forall Z$;
  end for
end for

4.2 No evidence from nature

In this section, we study an example where agents try to form opinions about an issue without obtaining evidence from nature. This situation is similar to forming opinions about political or social issues where there is no true state. Let us consider $\Omega = \{a, b\}$ with two ethical propositions, e.g., a indicates a pro-choice opinion, and b indicates a pro-life opinion. We test the belief evolution in a network using the combination and frequentist rules.

4.2.1 Communication via the Dempster's rule of combination

In this model, at each time step t:

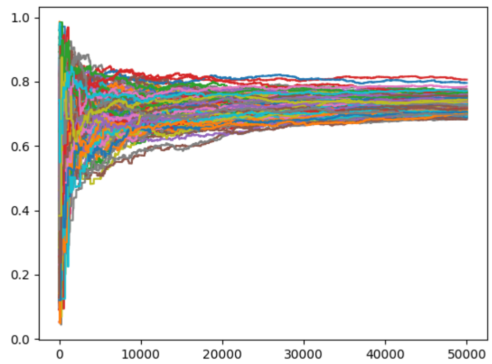
1. Select a proportion π of links.
2. For each link ij selected, the agents i, j update their `bmas` m_t^i, m_t^j by replacing them by $m_{t-1}^i \oplus m_{t-1}^j$ (Dempster's rule), provided the level of conflict κ does not exceed d (bounded confidence). Otherwise, no update occurs.

Algorithm 3 summarizes this procedure.

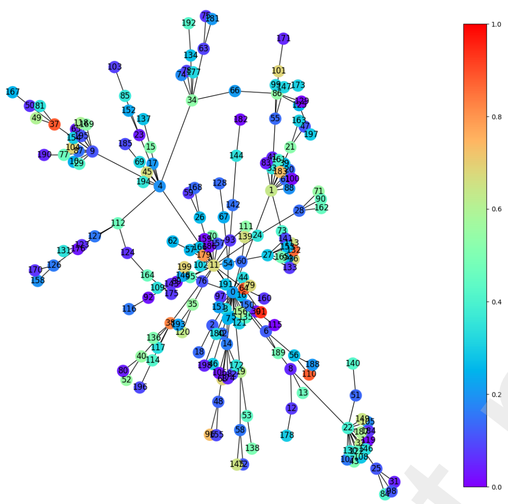
Algorithm 3 Dempster's rule with no true state

A network $\mathcal{G}$ is randomly generated via Barabasi-Albert algorithm;
for every time step t **do**
  A proportion π of links is selected randomly;
  for every selected link ij **do**
    Compute the level of conflict κ between m_t^i and m_t^j;
    if $\kappa \leqslant d$ **then**
      $m_t^i = m_t^j \leftarrow m_t^i \oplus m_t^j$;
    else state No update occurs;
    end if
  end for
end for

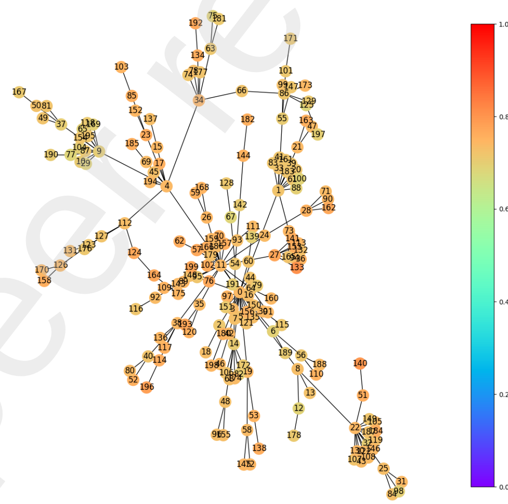
Figure 3 reveals total polarization if agents communicate with their neighbors using Dempster's rule. Surprisingly enough, this polarization remains even when $d = 0.999$, indicating that agents are highly tolerant and communicate with each other unless they are on the opposite side of the spectrum.



(a) Each line is an agent's $\texttt{BetP}(B)$.



(b) Initial distribution



(c) Final distribution

Figure 2: Created via Algorithm 2. Red indicates a belief in climate change ($\texttt{BetP}(B) = 1$). $p = 0.8, q = 9$.

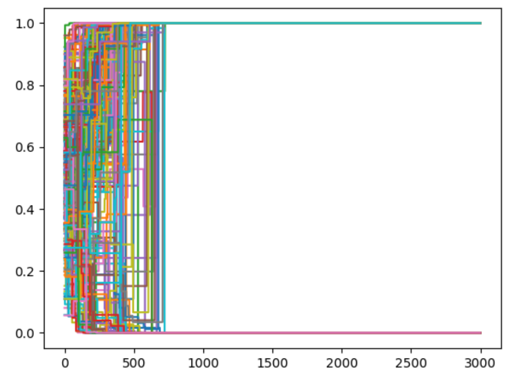
Moreover, we tested the robustness of our results under small-world graph generation, and the results remain robust.

Figure 3 shows a strong polarization result. Moreover, when we compare Figure 3 to Figure 1, we realize that the constant flow of evidence eventually wears down radical agents in our previous model. The lack of a constant persuader, nature, is a potent source of polarization.

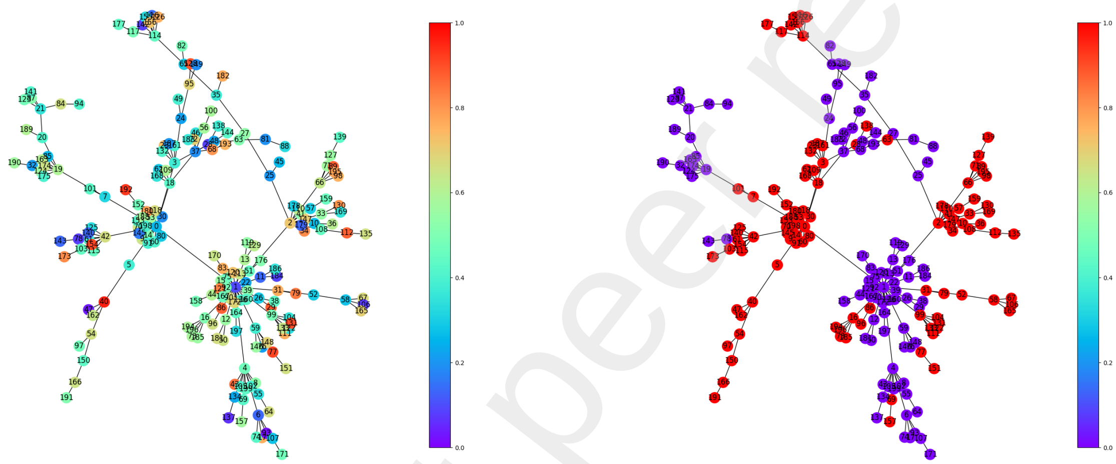
4.2.2 Frequentist communication

The procedure here is very similar to the previous case, except that Agent i sends a signal $X \subseteq \Omega$ with a probability m^i_{t-1} to a randomly chosen neighbor j, who updates its `bma` via the frequentist biased rule. Algorithm 4 describes this process.

Figure 4 shows the smoothest polarization and echo chamber emergence in all of our simulations. However, this polarization is not as substantial as in the previous section, where agents used the combination rule. We see small echo chambers of different opinions. Here, Figure 4a indicates a more chaotic result; however, one can observe that clusters in Figure 4c are indeed from the same opinion. The graph shows the emergence of an echo chamber, where one cluster remains with one opinion, and any outsider opinion is eliminated the turn after. Finally, it is possible to suggest that Dempster's rule will cause strict polarization due to its self-reinforcing nature, while frequentist rule will cause the



(a) Each single line represents an agent's $\texttt{BetP}(B)$.



(b) Initial distribution

(c) Final distribution

Figure 3: Created via Algorithm 3. Red indicates belief in climate change ($\texttt{BetP}(B) = 1$).

Algorithm 4 Frequentist rule with no true state

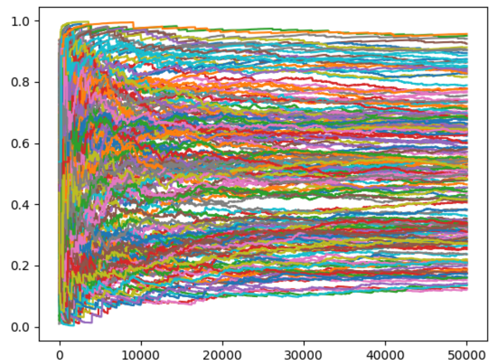
A network $\mathcal{G}$ is randomly generated via Barabasi-Albert algorithm;
for every time step t **do**
    A proportion π of links is selected randomly;
    for every selected link ij **do**
        Agent i sends signal $X \subseteq \Omega$ to agent j with probability $m^i_{t-1}(X)$
        Agent j sends signal $Y \subseteq \Omega$ to agent i with probability $m^i_{t-1}(Y)$
        $m^i_t(Z)$ is computed by Eq. 10 with $e_t = X$ and m^i_{t-1}, $\forall Z$;
        $m^j_t(Z)$ is computed by Eq. 10 with $e_t = Y$ and m^j_{t-1}, $\forall Z$;
    end for
end for

---

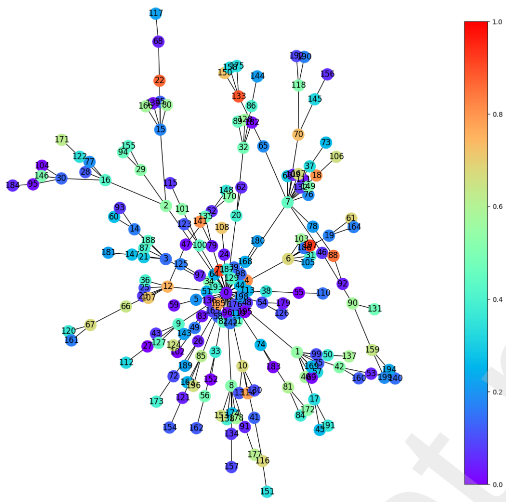
emergence of echo chambers.

4.3 Ambiguous evidence and recognition of ignorance

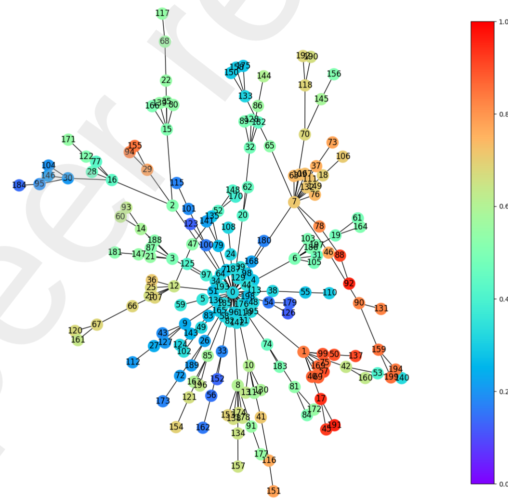
An essential contribution of this paper is that it implements a degree of recognition of ignorance. One question that arises from this fact is what happens to the ambiguity of agents under our belief update rules. To answer this question, we study two cases where agents attempt to learn the true



(a) Each line represents a single agent's $\texttt{BetP}(B)$ value (belief to climate change).



(b) Initial distribution



(c) Final distribution

Figure 4: Created via Algorithm 4. Red indicates belief in climate change ($\texttt{BetP}(B) = 1$).

ambiguity in nature, with and without communication. Algorithm 2 is used to create the case where agents communicate via the Biased update rule (Equation 10. Then, we create another model where agents receive evidence from nature and update their beliefs without the information from neighbors. Algorithm 5 summarizes the model.

Moreover, we increase the ambiguity of evidence to 0.9, meaning that $q = 0.1$. Only one piece of evidence out of ten is significant, while the rest is ambiguous.

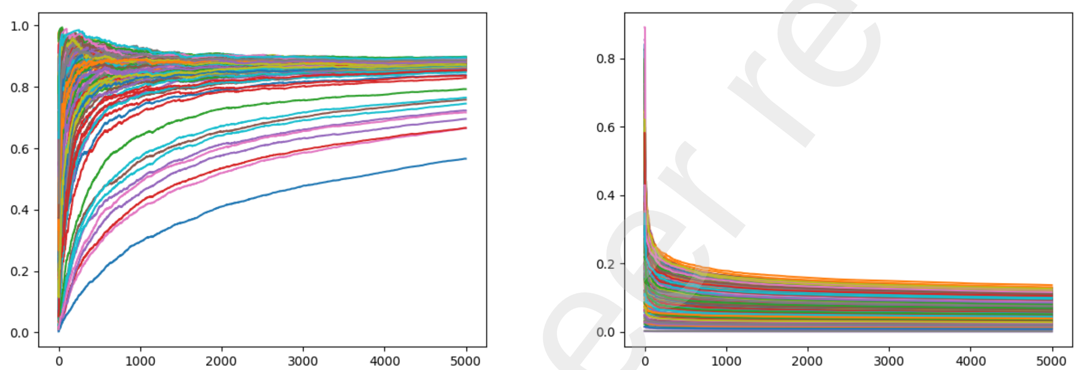
Figure 5 shows that when there is no communication between agents, they will learn the ambiguity of the evidence and eventually become self-doubtful. $m(\Omega)$ will go to q value for every agent. However, communication will immediately change the nature of this behavior. Even a frequentist approach will cause agents to become less doubtful about themselves. Eventually, $m(\Omega)$ will converge to 0 for every agent.

4.4 A political example

Finally, we study an example where $|\Omega| = 3$. Let us assume a political election with three parties: $\Omega = \{l, c, r\}$. In order to observe the impact of certainty, we study two types of agents. First, an agent can be the voter of a party with extreme certainty, e.g., $m(r) > 0.9$. Second, an agent can be indifferent between two parties while having a strong distaste for the third, e.g., $m(\{l, c\}) > 0.9$. We

Algorithm 5 Belief update under constant evidence

A network $\mathcal{G}$ is randomly generated via Barabasi-Albert algorithm;
for every time step t **do**
  Nature determines e_t:
  if $e_t = X \neq \Omega$ **then**
    $\mathbb{P}(e_t = X) = p_X * q$; $X \neq \varnothing$;
  else
    $\mathbb{P}(e_t = \Omega) = 1 - q$
  end if
  for every agent i **do**
    $m_t^i(Y)$ is computed by Eq. 10, $\forall Y \subseteq \Omega$;
  end for
end for



(a) No communication

(b) communication via frequentist update

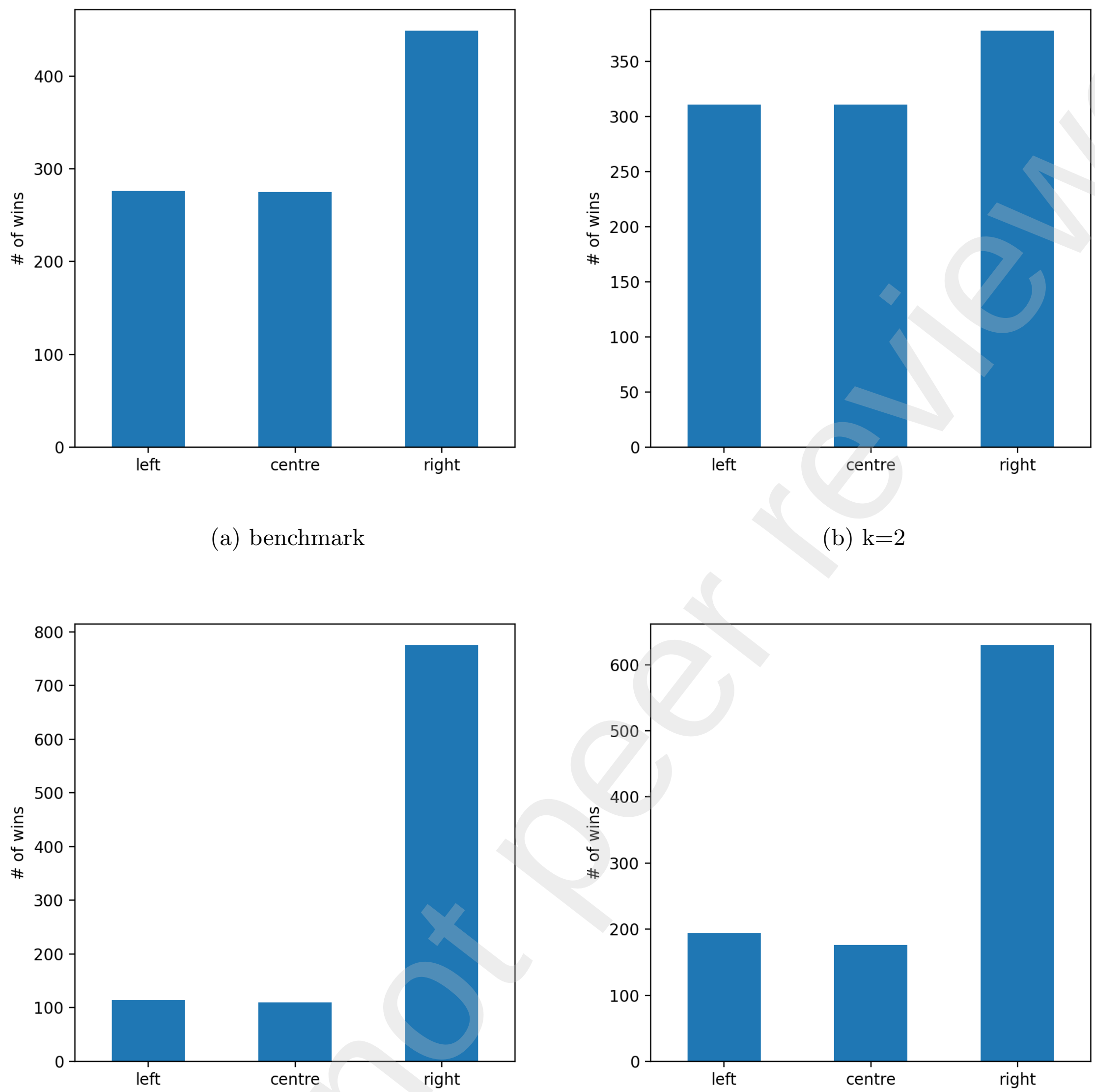
Figure 5: Figure 5a is created via Algorithm 5, while Figure 5b is created via Algorithm 2. Each line shows the recognition of ignorance of an agent ($m(\Omega)$). $q = 0.1$, meaning that agents receive almost only ambiguous evidence.

study under which circumstances the party with a certain voter base would prevail. For this section, we use the distributive update rule from Equation 6 since its advantage becomes apparent when we have more than two possible outcomes. Algorithm 6 summarizes the model.

Algorithm 6 Frequentist rule with three parties

A network $\mathcal{G}$ is randomly generated via Barabasi-Albert algorithm;
for every time step t **do**
  A proportion π of links is selected randomly;
  for every selected link ij **do**
    Agent i sends signal $X \subseteq \Omega$ to agent j with probability $m_{t-1}^i(X)$
    Agent j sends signal $Y \subseteq \Omega$ to agent i with probability $m_{t-1}^i(Y)$
    $m_t^i(Z)$ is computed by Eq. 10 with $e_t = X$ and m_{t-1}^i, $\forall Z$;
    $m_t^j(Z)$ is computed by Eq. 10 with $e_t = Y$ and m_{t-1}^j, $\forall Z$;
  end for
end for

Figure 6 shows that the right party wins most elections. The difference in initial ambiguity beliefs is apparent. More interestingly, we observe that the impact of increasing the starting age is exceptionally high. This effect is expected since our model increases *belief persistence* with the amount of evidence one receives. Therefore, left and centre-party voters tend to change their opinions since their initial beliefs are more ambiguous, while right-party voters eventually act as persuaders. We also observe



(a) benchmark

(b) k=2

(c) starting age = 100

(d) $m(\Omega)$ is very high

Figure 6: The effect of initial age, ambiguity, and network connectivity. In the benchmark model, k=1, starting age is 10, and initial $m(\Omega)$ is low for all agents.

that the more connected a network gets, the differences become less noticeable. This indicates that network bottlenecks work in favor of the party that is more certain of herself. Finally, we observe that initial ambiguity helps the party that is certain of themselves (see Grabisch et al. (2023) for an opinion dynamics paper that shows under which conditions polarization benefits the majority candidate).

5 Concluding remarks

Dempster-Shafer Theory in opinion dynamics had been relatively barren, with an emerging interest in the last decade (Lu et al., 2015; Huang et al., 2023). So far, to our knowledge, our paper is the first study that creates a frequentist model in DST. We study two approaches created in the works of literature of Bayesian and non-Bayesian models, using various update rules in DST. We show that DST is strong and versatile enough to generate various phenomena of social networks.

Climate change was a polarized topic of discussion in the mid-late twentieth century. However, the number of people who believe that climate change is real and that humans are the leading cause is increasing over time (Milfont et al., 2017). Ethical discussions like abortion (Adamczyk et al., 2020) or political discussions like fraud in elections, however, cause significant polarization in the population, which we do not observe vanishing. Our preliminary results predict this behavior: agents tend to form a consensus on issues with a constant source of meaningful, accurate information. Meanwhile, a social issue without a definitive true answer creates significant polarization. To summarize, nature eventually acts as a persistent persuader, which has been shown to prevail.

Furthermore, the ambiguity of the evidence changes behavior depending on whether there is communication or not. When there is no communication, agents realize that the evidence is mostly noise. However, communication polarises the community, and agents start believing they are entirely correct and there is no ambiguity. The combination rule might cause the violation of *belief persistence* (Ryan, 2002) when two agents with vastly different opinions collide. The proposed model is similar to subtle belief update methodologies observed in heuristic-based models, where agents take one step toward others instead of immediately meeting somewhere in the middle. Hence, this approach is immune to the violation of *belief persistence*.

A Application of Dirichlet distribution

The Dirichlet distribution is a generalization of the beta distribution to higher dimensions and is commonly used in Bayesian statistics for generating random probability vectors that sum to 1 (Ng et al., 2011; Robert et al., 2007). We use these distributions in this study for the construction of priors in Section 4. Formally, the Dirichlet distribution, denoted as $\mathrm{Dir}(\alpha_1, \ldots, \alpha_K)$ for a $K-1$ dimensional simplex, is a family of continuous multivariate probability distributions parameterized by a vector $\alpha = (\alpha_1, \ldots, \alpha_K)$ of positive reals.

$$f(x_1, x_2, \ldots, x_K; \alpha_1, \alpha_2, \ldots, \alpha_K) = \frac{1}{B(\alpha)} \prod_{i=1}^{K} x_i^{\alpha_i - 1} \tag{15}$$

where $x_1, x_2, \ldots, x_K$ are non-negative real numbers that sum to 1 ($\sum_i^K x_i = 1$), and $B(\alpha)$ is the multinomial beta function, which serves as a normalization constant to ensure that the total probability integrates to 1. The function $B(\alpha)$ is defined as:

$$B(\alpha) = \frac{\prod_{i=1}^{K} \Gamma(\alpha_i)}{\Gamma\left(\sum_{i=1}^{K} \alpha_i\right)} \tag{16}$$

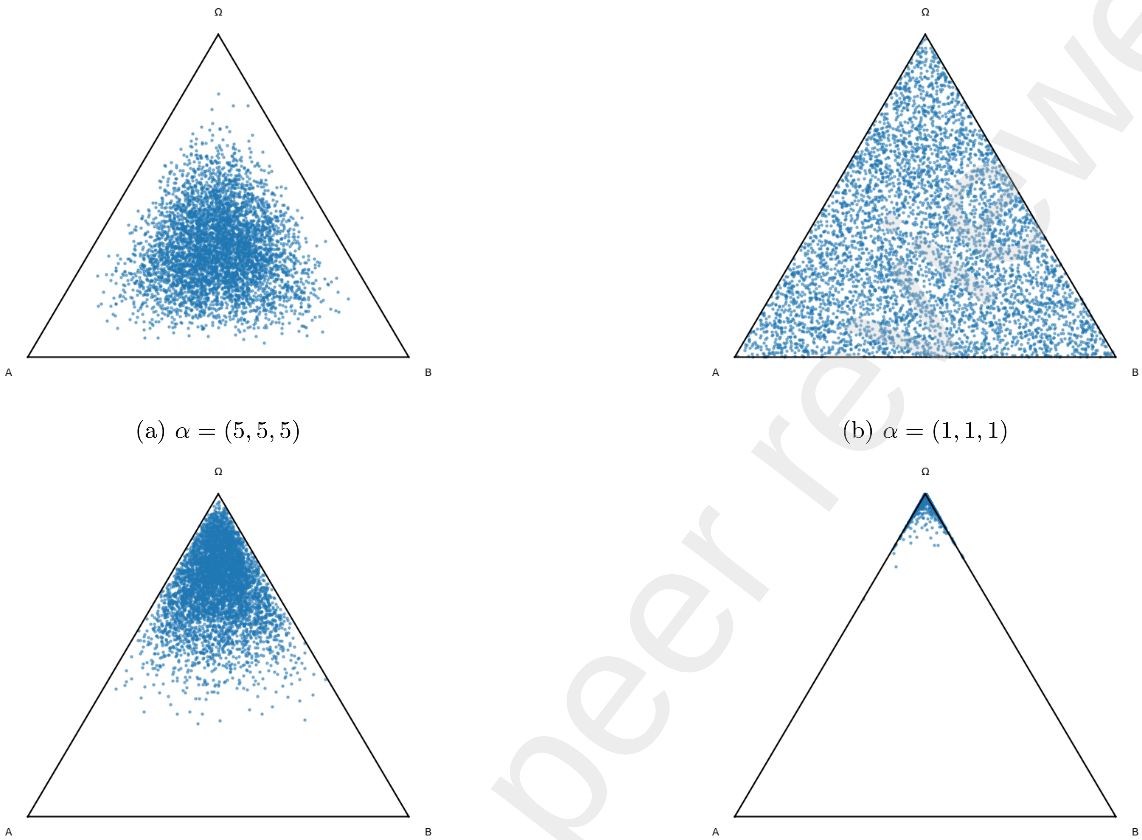
where $\Gamma(\cdot)$ denotes the gamma function. Its suitability for our model is based on two properties:

1. Conformity to normalization constraint: Given the requirement in DST that the sum of masses across all subsets of the frame of discernment must be unity, the Dirichlet distribution is apt due to its inherent property of generating vectors: $m_0 = (m_0(X_1), \ldots, m_0(X_K))$ where $\sum_{i=1}^{K} m_0(X_i) = 1$, $X \subseteq \Omega$.

2. Flexibility in encoding beliefs: The parameters α_i of the Dirichlet distribution directly influence the distribution's shape, allowing for modeling a wide range of initial belief states. This is vital in representing diverse agent perspectives in DST, where α_i essentially controls the agent's initial confidence in the corresponding outcome.

Let $\Omega = \{A, B\}$, in order to create a prior mass $m_0 = \{m(\varnothing), m(A), m(B), m(\Omega)\}$, we use a Dirichlet distribution with parameters $(\alpha_A, \alpha_B, \alpha_\Omega)$. Please note that we assume $\alpha_\varnothing = 0$ due to the closed-world hypothesis.

In Section 4, we first create agents randomly in Algorithms 1, 2, 3, 4, and 5. In these models, we have a power set $2^\Omega = \{\varnothing, \{\text{Climate change is real}\}, \{\text{It is not real}\}, \{\text{Either can be true}\}\} = \{\varnothing, \{A\}, \{B\}, \Omega\}$. We assume that an agent will always consider $\alpha = 1$ for all non-empty set events, e.g., $\alpha = (1, 1, 1)$ (we omit the case of the empty set for the sake of simplification). On the other hand, in Algorithm 6, we assume biased agents towards one of the outcomes. Let $\Omega = \{l, c, r\}$; for example, a voter who supports the right party with almost full certainty would have a prior selected from a distribution with parameters $\alpha = (0.1, 0.1, 20, 0.1, 0.1, 0.1, 0.1)$, while a voter who hates the right party while being indifferent between left and center parties would have a prior selected from $\alpha = (0.1, 0.1, 0.1, 20, 0.1, 0.1, 0.1, 0.1)$.



(a) $\alpha = (5, 5, 5)$

(b) $\alpha = (1, 1, 1)$

(c) $\alpha = (2, 2, 10)$

(d) $\alpha = (0.1, 0.1, 20)$

Figure 7: Each triangle represents a 2-simplex where each point inside is a vector (A, B, Ω), where $A + B + \Omega = 1$.

To illustrate how different Dirichlet parameters lead to different distributions, we generate four different plots in Figure 7. All parameters are equal in Figures 7a and 7b. It is possible to see that $\alpha = (1, 1, 1)$ from Figure 7b is equivalent to a multivariate uniform distribution, where all points are equally likely. In contrast, $\alpha = (5, 5, 5)$ from Figure 7a causes more central points to be selected. Then, we create two plots where the distribution favors one of the outcomes in Figures 7c and 7d. We observe that many agents from Figure 7c have beliefs that are still close to the center, while Figure 7d causes all agents to be biased. In this paper, we used parameters from Figure 7b for unbiased agents and those from Figure 7d for biased agents.

References

Acemoglu, D., Dahleh, M. A., Lobel, I., & Ozdaglar, A. (2011). Bayesian learning in social networks. *The Review of Economic Studies*, *78*(4), 1201–1236.

Acemoglu, D., & Ozdaglar, A. (2011). Opinion dynamics and learning in social networks. *Dynamic Games and Applications*, *1*, 3–49.

Acemoglu, D., Ozdaglar, A., & ParandehGheibi, A. (2010). Spread of (mis) information in social networks. *Games and Economic Behavior*, *70*(2), 194–227.

Adamczyk, A., Kim, C., & Dillon, L. (2020). Examining public opinion about abortion: A mixed-methods systematic review of research over the last 15 years. *Sociological Inquiry*, *90*(4), 920–954.

Allahverdyan, A. E., & Galstyan, A. (2014). Opinion dynamics with confirmation bias. *PloS one*, *9*(7).

Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *science*, *286*(5439), 509–512.

Baumann, F., Lorenz-Spreen, P., Sokolov, I. M., & Starnini, M. (2020). Modeling echo chambers and polarization dynamics in social networks. *Physical Review Letters*, *124*(4).

Broido, A. D., & Clauset, A. (2019). Scale-free networks are rare. *Nature communications*, *10*(1).

Chandrasekhar, A. G., Larreguy, H., & Xandri, J. P. (2020). Testing models of social learning on networks: Evidence from two experiments. *Econometrica*, *88*(1), 1–32.

Dominiak, A., & Lee, M. S. (2017). Coherent dempster–shafer equilibrium and ambiguous signals. *Journal of Mathematical Economics*, *68*, 42–54.

Dubois, D., & Denoeux, T. (2012). Conditioning in dempster-shafer theory: Prediction vs. revision. *Belief Functions: Theory and Applications: Proceedings of the 2nd International Conference on Belief Functions, Compiègne, France 9-11 May 2012*, 385–392.

Fryer Jr, R. G., Harms, P., & Jackson, M. O. (2019). Updating beliefs when evidence is open to interpretation: Implications for bias and polarization. *Journal of the European Economic Association*, *17*(5), 1470–1501.

Gilboa, I., & Schmeidler, D. (1993). Updating ambiguous beliefs. *Journal of economic theory*, *59*(1), 33–49.

Grabisch, M., et al. (2016). *Set functions, games and capacities in decision making* (Vol. 46). Springer.

Grabisch, M., Mandel, A., & Rusinowska, A. (2023). On the design of public debate in social networks. *Operations Research*, *71*(2), 626–648.

Grabisch, M., & Rusinowska, A. (2020). A survey on nonstrategic models of opinion dynamics. *Games*, *11*(4), 65.

Huang, Z., Li, W., & Yang, J. (2023). Belief updating under ambiguity: A numerical simulation analysis. *Economics Letters*, *233*, 111359.

Lu, X., Mo, H., & Deng, Y. (2015). An evidential opinion dynamics model based on heterogeneous social influential power. *Chaos, Solitons & Fractals*, *73*, 98–107.

Milfont, T. L., Wilson, M. S., & Sibley, C. G. (2017). The public's belief in climate change and its human cause are increasing over time. *PloS one*, *12*(3).

Mislove, A., Marcon, M., Gummadi, K. P., Druschel, P., & Bhattacharjee, B. (2007). Measurement and analysis of online social networks. *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*, 29–42.

Ng, K. W., Tian, G.-L., & Tang, M.-L. (2011). *Dirichlet and related distributions: Theory, methods and applications*. John Wiley & Sons.

Robert, C. P., et al. (2007). *The bayesian choice: From decision-theoretic foundations to computational implementation* (Vol. 2). Springer.

Ryan, M. J. (2002). Violations of belief persistence in dempster–shafer equilibrium. *Games and Economic Behavior*, *39*(1), 167–174.

Shafer, G. (1976). *A mathematical theory of evidence* (Vol. 42). Princeton university press.

Shafer, G. (1981). Jeffrey's rule of conditioning. *Philosophy of Science*, *48*(3), 337–362.

Shapley, L. S. (1953). *A value for n-person games*. Princeton University Press.

Smets, P., & Kennes, R. (2008). The transferable belief model. *Classic Works of the Dempster-Shafer Theory of Belief Functions*, 693–736.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world'networks. *nature*, *393*(6684), 440–442.